INVESTIGATION OF CARBON CATABOLITE REPRESSION IN *CLOSTRIDIUM BEIJERINCKII* NCIMB 8052

M. S. RAFUDEEN

A dissertation presented in fulfillment of the requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Molecular and Cell Biology

UNIVERSITY OF CAPE TOWN

November 2001

CONTENTS

ACKNOWLEDGEMENTS

First and foremost I wish to express my love and gratitude to my darling wife, Salma, for her constant support and encouragement during the course of my studies, and during this thesis in particular, I would never have been able to arrive so far without her by my side.

I would also like to express my sincere thanks and appreciation to my supervisor, Dr Sharon Reid, for her advice, invaluable guidance and encouragement throughout my stay in the Anaerobe unit.

There are a number of people I wish to thank, all past and present members of the Anaerobe unit, namely Garth, Joy, Nina, Careena, Anna, Marla, Wata, Anton, Keith and Anwar for their camaraderie in creating an enjoyable working environment and their willingness to provide assistance whenever it was needed. I would also like to thank my colleagues who were ex-members of the Marine microbe and Plant microbe units respectively, namely Declan and Craig.

A special thanks goes to Dr Neil Leat for proofreading this dissertation. I would also like to thank Dr Val Abratt for her insight when required and to Di James for technical assistance in nucleotide sequencing.

Lastly I express my sincere appreciation to my parents and family for their support and encouragement.

I acknowledge the financial support provided by the Foundation for Research and Development during the early part of this work.

ABBREVIATIONS

A	adenosine
Ap	ampicillin
A_x	absorbance at X nm
aa	amino acids
ATP	adenosine 5'-triphosphate
bp(s)	base pair(s)
BSA	bovine serum albumin
°C	celsius
C	cytidine
CBM	clostridial basal medium
Cm	chloramphenicol
CsCl	cesium chloride
cre	catabolite responsive element
CCR	carbon catabolite repression
Da	Daltons
DNA	deoxyribonucleic acid
dco	double cross-over recombination
ds	double stranded
EDTA	ethylenediaminetetra-acetic acid
Em	erythromycin
F	farad
G	guanosine
g	gram
_g	centrigugal force
hr	hour(s)
HMW	high molecular weight
IPTG	isopropyl -D-thiogalactopyranoside
kan	kanamycin
kb	kilobases
kD	kilodaltons
LA	Luria agar
LB	Luria broth
m	milli
M	molarity
mcs	multiple cloning site
mins	minutes
ml	millilitre
MW	molecular weight
ng	nanograms
nm	nanometers
nt(s)	nucleotide(s)
°C	degrees Celcius
OD	optical density
ORF	open reading frame
ori	origin of replication
p	plasmid
r (superscript)	resistance
rc	rolling-circle
RNA	ribonucleic acid
rpm	revolutions per minute
s	seconds
SD	shine-dalgarno
sco	single cross-over recombination
T	thymidine
Tc	tetracycline
UV	ultra violet

V	volt
vol	volume
W	Watt
w/v	weight/volume
X-gal	5-bromo-4-chloro-3-indolyl-D galactopyranoside
α	alpha
β	beta
Δ-	delta
	micro
$\bar{\lambda}$	lambda
()	plasmid carrier state
+(superscript)	presence of
-(superscript)	absence of

ABSTRACT

The substrate basis for the industrial acetone-butanol-ethanol (ABE) fermentations, has been agricultural products rich in starch or sucrose, and employed taxonomically distinct amylolytic and saccharolytic solventogenic clostridial strains respectively. There is evidence to suggest that the utilization of these substrates is subject to carbon catabolite repression. In Gram-positive bacteria, carbon catabolite repression is controlled by a global regulatory mechanism, central to which is an imperfect palindromic sequence, the *cre* element, which is recognized by a protein of the GalR-LacI family, the CcpA protein. A *ccpA* homologue, *regA*, has been previously identified in *C. acetobutylicum* NCP262 and successfully complemented a *B. subtilis ccpA* mutant strain. The sucrose operon from *C. beijerinckii* NCIMB 8052, *scrARBK*, has been characterized at the physiological and genetic level with the ScrR repressor found to negatively auto-regulate the operon. There was also physiological evidence that the sucrose operon was subject to carbon catabolite repression and a *cre* element was identified within the operon. The primary aim of this study was to first identify the *ccpA* homologue in *C. beijerinckii* and to establish whether a relationship exists between the homologue and the sucrose operon in order to elucidate the carbon catabolite repression mechanism.

The *C. acetobutylicum regA* gene was used as a probe in hybridization analysis to isolate a gene from a chromosomal gene library of *C. beijerinckii* NCIMB 8052. This gene was named *regB* and was 996 bp in length and encoded a protein of 332 amino acids with a calculated MW of 39 840 Da. Phylogenetic, sequence and structural analyses confirmed that the RegB protein belonged to the CcpA subfamily within the GalR-LacI family of transcriptional regulators. The RegB protein retains conserved key structural and functional amino acid residues of the CcpA subfamily, which indicates conformational and functional identity with CcpA. This suggested that RegB could act as a global regulator of CCR in *C. beijerinckii*. A gene encoding a protein with significant identity to isoleucyl-tRNA synthetase proteins was found upstream of *regB*. The gene immediately downstream of *regB* encoded a protein of 147 amino acids and showed identity with the MarR family of transcriptional regulators. Downstream of the putative MarR family transcriptional regulator was a truncated gene, which encoded a

protein with 30% amino acid identity with an unknown, uncharacterized conserved membrane protein (BAB05882) found in *B. halodurans*.

Attempts were made to inactivate the *regB* gene of *C. beijerinckii* using two suicide inactivation systems. The first was based on the pMTL30 inactivation system which relies on conjugation. A second alternative inactivation system for *C. beijerinckii* based on the thermosensitive pG^+host system was developed. This system required electroporation of plasmid DNA into *C. beijerinckii*, followed by growth of the transformants at a non-permissive temperature to prevent plasmid replication and to force integration to take place. Southern hybridization analysis revealed that the *regB* gene of *C. beijerinckii* was disrupted using the pMTL30 suicide system and that the pMTL30RegB construct was integrated in the *C. beijerinckii* chromosome. Subjection of the *C. beijerinckii regB* mutant to sporulation led to the subsequent loss of the mutant upon germination. The development of new inactivation vectors using the pG^+host system was also not successful in producing a *C. beijerinckii regB* mutant by either disruption or deletion of the *regB* gene. Though integration events occurred, Southern hybridization analysis revealed that the integrants retained the wild-type *regB* gene, and that there was vector re-arrangement in *C. beijerinckii*.

Glucose repression of the *C. beijerinckii scrARBK* operon was retained in a *C. beijerinckii scrR* mutant grown in glucose and sucrose, indicating that the *scrARBK* operon is subject to glucose repression in addition to negative regulation by ScrR. Nucleotide and primer extension analysis of the *scrARBK* operon identified a single transcriptional start at an adenine nucleotide, 44 bp upstream of the translational start of the operon, confirming the promoter region, translational start and *cre* (*cre*I) element which had been previously identified. A second imperfect palindromic sequence (*cre*II), $[A_{12}T_{11}T_{10}A_9T_8A_7G_6T_5T_4T_3T_2C_1/G'_1A'_2A'_3A'_4A'_5T'_6A'_7T'_8A'_9A'_{10}A'_{11}T'_{12}]$ was identified further upstream, 139 bp before the transcriptional start. Comparison of the *scrARBK cre*I element to published *cre* elements from other Gram-positive bacteria found that it had the highest G+C content, predominantly in its flanking sequences.

The RegB protein was successfully expressed in *E. coli* and purified using a histidine affinity column. We demonstrated that overexpression of RegB using plasmid pCTC1 in *C. beijerinckii* resulted in an increase in CCR efficiency of sucrose hydrolase activity, in the early exponential *C. beijerinckii* growth phase. This provided a physiological link between RegB activity and CCR regulation of the sucrose operon and was consistent with the role of RegB as a CcpA homologue.

In the light of the new evidence we have provided we propose that the *regA* and *regB* genes are indeed homologues of the *ccpA* gene found in many Gram-positive bacteria and that the nomenclature of *regA* and *regB* should be changed to *ccpA*.

CHAPTER ONE

GENERAL INTRODUCTION AND LITERATURE REVIEW

CONTENTS

CHAPTER 1

1.1 GENERAL INTRODUCTION

Clostridium beijerinckii NCIMB 8052 is a mesophilic, gram-positive, obligately anaerobic, spore-forming bacterium that produces solvents such as acetone, butanol and ethanol from carbohydrate sources (Verhasselt and Vanderleyden, 1993). This strain is closely related to the solventogenic clostridial strains that were used in industrial acetone-butanol-ethanol (ABE) fermentations. In order to understand *C. beijerinckii* NCIMB 8052 and its relationship with the solventogenic clostridial strains, an overview will be given of the history of the ABE fermentation and the different carbohydrate substrates used in the process. The general relationship between members of the clostridial solventogenic group will be reviewed. This will be followed by a review of bacterial catabolite repression with emphasis on catabolite control of the *C. beijerinckii* NCIMB 8052 sucrose system. For brevity, *C. beijerinckii* NCIMB 8052 will be hereafter referred to as *C. beijerinckii*.

1.1.1 The History of the ABE Fermentation Process

The outbreak of the First World War served as a catalyst for the industrial use of solventogenic clostridia to produce acetone and butanol from grain sources (Jones and Woods, 1986). In 1910, the company of Strange & Graham Ltd. (United Kingdom) sought the services of ChaimWeizmann in order to investigate the production of butanol and isoamylalcohol by microbial fermentation (Morris, 1993). Between 1912 and 1914 Wiezmann isolated a starch-fermenting clostridial culture, initially designated BY, which could produce acetone from maize mash. This culture was later named *Clostridium acetobutylicum* and served as the bacterium on which Weizmann patented the ABE fermentation process to produce solvents. Initially these ABE biotransformations used potatoes as a substrate but later maize was used which increased solvent production. With advent of the First World War in 1914, the ABE fermentation process became

crucial for the production of acetone, which was required in the manufacture of munitions. After the war an American company, Commercial Solvents Corporation, acquired the patent rights for the ABE fermentation process.

In 1922, Robinson carried out a study on the fermentation of various carbohydrates in the production of solvents (Robinson, 1922). His results indicated that glucose, fructose, mannose, sucrose, lactose, starch and dextrin were completely consumed while galactose, xylose, arabinose, raffinose, melezitose, inulin and mannitol were partially utilized and trehalose, rhamnose, melibiose and glycerol were not fermented. Subsequently other researchers showed that xylose, arabinose and carbohydrates occurring in dairy and wood wastes could also be completely fermented by solvent producing clostridia (Nakhmanovich and Shcheblykina, 1959; Compere and Griffith, 1979).

In the mid-1930s, starch based ABE fermentations were replaced by fermentations that used molasses as a substrate due to an abundant supply of molasses, its low cost and difficulties with starch as a substrate. It was thought that this rich sugar substrate would permit fermentation of carbohydrates at higher concentrations and have other technical advantages over grain sources (Hastings, 1978). Starch-fermenting clostridial strains could not be used in the molasses-based fermentation process with the result that new saccharolytic, solvent-producing strains with different phenotypic characteristics were isolated. By 1935 cultures were isolated which could ferment 6.5% sugars to produce 2% solvents, this halved distillation costs and allowed the use of the cheaper molasses substrate. These new strains were developed and patented by various companies producing solvents, with the result that each patented strain was assigned a novel species name. The nomenclature adopted was not systematic and there was no clear understanding of the relationships between the solvent-producing strains. With the expiry of the patent in 1935, the ABE fermentation process became available worldwide

and was the major supplier of industrial solvents till the 1950s and 1960s (Jones and Woods, 1986).

Two factors, however, led to the slow demise in the 1950s of ABE fermentations using solvent producing clostridial strains; firstly, the increasing price of carbohydrate substrates and secondly, competition from the petrochemical industry. The People's Republic of China is currently the only country in which ABE fermentations are operated. There has been renewed interest in solvent-producing clostridial strains due to various factors. The depletion of natural oil reservoirs coupled with concern over the environmental problems created by the petrochemical industry, makes solventogenic clostridia attractive due to the fact that they can produce solvents from safe, renewable resources (Woods, 1995; Dürre, 1998). Furthermore, economic factors are also attractive due to high oil prices and the current surplus of agricultural wastes and byproducts that could be utilized as inexpensive fermentation feedstocks.

The ability of solventogenic saccharolytic clostridia to ferment many different carbohydrates has allowed several schemes to be proposed which would allow production of solvents from alternative low-cost substrates (Annous and Blaschek, 1991). The cost of the traditional fermentation substrate accounts for 60% of the production cost in producing solvents by fermentation (Ross, 1961). Other carbohydrate substrates that have been looked at for fermentation were Jerusalem artichokes (Reynolds and Werkman, 1934), cheese whey (Welsh and Veliky, 1984), apple pomace (Voget et al., 1985) and algal biomass (Nakas et al., 1983).

The carbohydrate present in Jerusalem artichokes occurs as short oligomeric fructans which must be hydrolyzed prior to fermentation. The hydrolysates also had to be supplemented with maize or soy meal with the addition of ammonia as a nutritional supplement (Marchal et al., 1985). Cheese whey contains lactose, which could be used in fermentations and in 1983 it was estimated that 10 billion pounds of cheese whey is

wasted every year by dumping (Moreira, 1983). Although cheese whey can be used in ABE fermentation, it is a poor substrate due to its low lactose content (4-5%) and because there is incomplete utilization of lactose when compared to starch and molasses. Apple pomace is rich in carbohydrates of which glucose makes up 23% and sucrose, 10%. Voget et al., (1985), found that the yield of butanol from apple pomace fermentations was in the range of 1.9% to 2.2% with approximately 80% of the sugars consumed. Nakas et al. (1983), found that 16g per liter of solvents could be obtained by *C. pasteurianum* fermentation using algal biomass supplemented with 4% of glycerol. However, the solvent mixture did not contain acetone but was rich in butanol and 1,3 propanediol. Although no salt inhibition was observed, any large-scale algal cultivation may depend on finding salt-tolerant, solvent-producing clostridia.

Cellulolytic substrates have also been investigated for use in the fermentation process. These substrates include lignocellulose (Mes-Hartree and Saddler, 1982), sulfite waste liquors (Wiley et al., 1941) and hydrolysates (Langlykke et al., 1948). Lignocellulose consists of hemicellulose and cellulose. Xylose is the major constituent of hemicellulose and it was shown that *C. acetobutylicum* could use the sugars present in hemicellulose to produce solvents at between 8 to 17g/liter (Yu *et al.*, 1984a).

Sulfite waste liquors from the pulp and paper industry contain glucose, xylose and arabinose, which could be used for fermentation. However sulfur dioxide, lignin and calcium have to be removed from the effluent in order to obtain a good solvent yield. Fermentation of a sugar mixture made to simulate sulfite liquor effluent produced 0.36g of solvents per gram of sugar with almost complete utilization of the sugar (Wayman and Yu, 1985).

Acid hydrolysis of cellulolytic material could serve as substrate for fermentation. It has been shown that hemicellulose extracted by water, could be hydrolyzed by acid or enzymes to release sugars. These sugars were utilized by *C. acetobutylicum* and did not

contain inhibitory substances in the hydrolysates (Saddler et al., 1983; Yu et al., 1984b). Recently domestic organic waste has been investigated as a substrate for the production of acetone, butanol and ethanol by solventogenic clostridia (Claasen et al., 2000; Lopez-Contreras et al., 2000).

Recent advances in the field of molecular genetics and biotechnology raise the possibility of creating recombinant solventogenic strains with increased solvent production potential, amplified substrate degradation capabilities and an extended spectrum of substrates. This could in theory be achieved because the solvent-producing clostridial group contains a variety of species each with different characteristics and capabilities. However in order to utilize these features an understanding of the taxonomic relationships between clostridial species is required.

1.1.2 Taxonomy of Solventogenic Clostridia

Taxonomic studies in the 1920's and 1930's on solvent producing strains which preferentially used starch as a substrate, led investigators to conclude that all these microorganisms were closely related and belonged to a single species, namely *Clostridium acetobutylicum* (Weyer and Rettger, 1927; McCoy and McClung, 1935). A clostridial strain, ATCC 824, which was isolated in 1924, was selected as the type strain of the species.

Taxonomic studies performed on saccharolytic solvent producing strains based on cell wall composition and DNA homologies resulted in strains being assigned to two distinct groups, *C. butyricum* and *C. beijerinckii* with strains labeled *C. acetobutylicum* showing a low level of identity with either group (Cummins and Johnson, 1971). After further extensive taxonomic studies it has been found that the most commonly used solvent-producing strains differ substantially in their physiological and genetic characteristics (Johnson et al., 1997; Johnson and Chen, 1995; Keis, et al., 1995).

The solvent-producing clostridial strains that were studied most intensively: *C. acetobutylicum* ATCC 824 (ATCC *C. acetobutylicum* type strain); *C. acetobutylicum* DSM 792 (DSM *C. acetobutylicum* type strain); *C. beijerinckii* NCIMB 8052; *C. saccharoperbutylacetonicum* N1-4 and *C. acetobutylicum* NCP 262 appear not to constitute a homologous group and may represent several different species (Wilkinson et al., 1995a).

When *C. beijerinckii* was compared to other solvent-producing clostridial strains, the differences included the ability to sustain solventogeneis in continuous culture (Kashket and Cao, 1995), the pH at which solventogenisis is initiated (Holt et al., 1984), restriction endonuclease systems (Mermelstein and Papoutsakis, 1993a; Lee et al., 1992), genomic relatedness (Johnson and Chen, 1995) and restriction endonuclease profiles and genome size (Wilkinson and Young, 1993; 1995).

The taxonomic and phylogenetic study by Keis et al., (1995), on 55 solvent-producing clostridial strains using biotyping, DNA fingerprint analysis and 16S ribosomal RNA sequence analysis identified four taxonomic groups. They also concluded and confirmed previous results that solvent-producing clostridia isolated and patented from the mid-1930s did not constitute a homologous group and proposed that each group represents a distinct species.

The relationships between the four clostridial taxonomic groups can be seen in Figure 1.1 below. Taxonomic group 1 consists of amylolytic strains and is represented by *C. acetobutylicum* ATCC 824. Taxonomic groups 2, 3 and 4 consist mainly of saccharolytic strains and are represented by *C. acetobutylicum* NCP262, *C. saccharoperbutylacetonicum* N1-4a and *C. beijerinckii* NCIMB 8052 respectively.

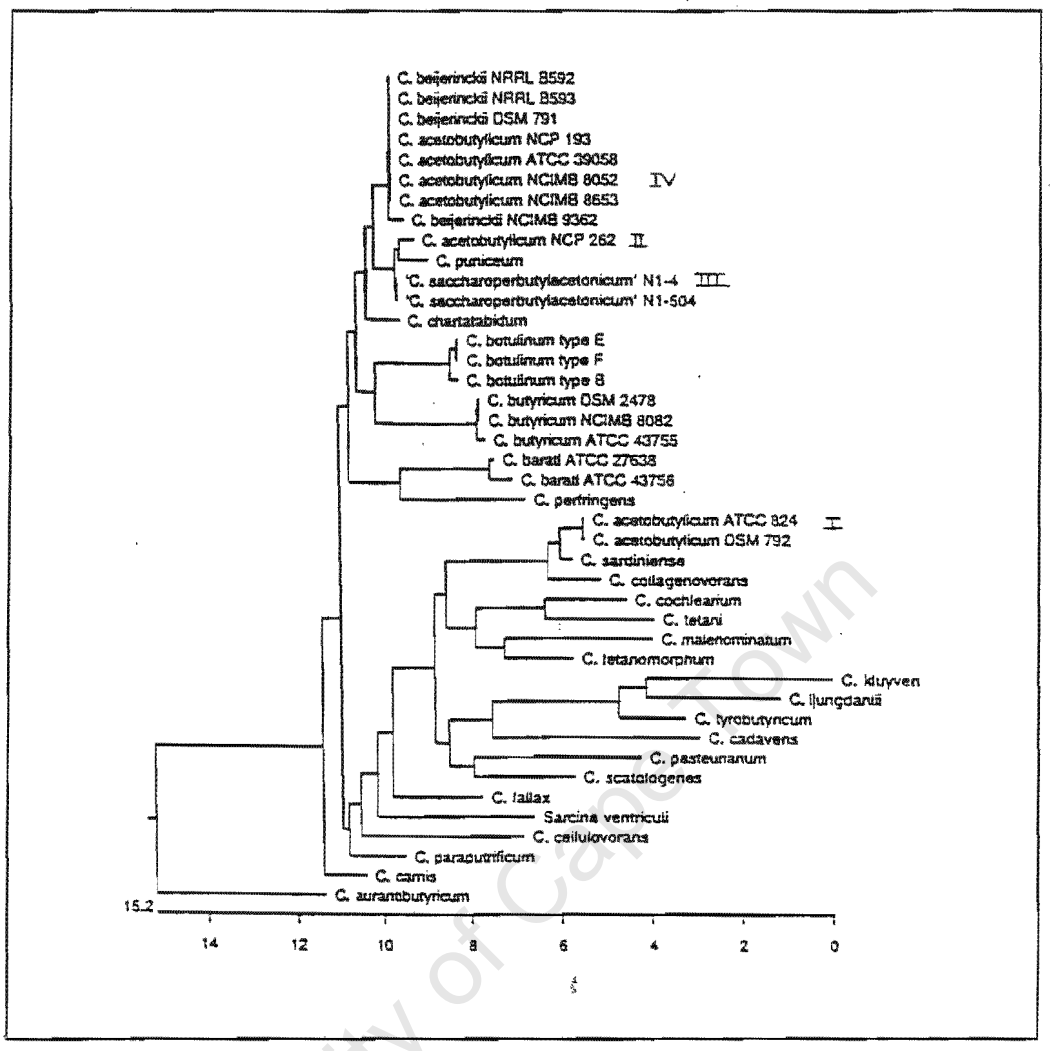

Figure 1.1: Unrooted phylogenetic dendogram showing the taxonomic positions of the solvent-producing strains *C. acetobutylicum* ATCC 824, DSM 792, NCP 193, NCIMB 8052 (*C. beijerinckii*), NCIMB 8653 and ATCC 39058, *C. beijerinckii* NRRL B593 and "*C. saccharoperbutylacetonicum*" N1-4 and N1-504, based on the partial 16S rRNA gene sequences examined (positions 830 to 1383; *E. coli* numbering). The scale bar indicates evolutionary distances. Adapted from Keis et al. (1995).

The phylogenetic tree illustrates the relationships between the various solvent-producing clostridial species, which were isolated over time based on their different substrate requirements for ABE fermentation.

The clarification of solventogenic clostridial taxonomy allows the genetic relationships between members of this group to be understood and also enables a more intelligent and structured approach in genetic analyses.

Shaheen et al. (2000) utilized representative strains from each of the sub-groups within the four taxonomic groups, including *C. beijerinckii*, for a comparative study of fermentation performance on different media under standardized conditions. This provided useful comparative data for the selection and development of strains for use on various commercial fermentation substrates. Knowledge of the taxonomic relationships between the clostridia will allow better development of commercial substrates for the production of solvents and will facilitate the selection of specific solventogenic clostridia for industrial use.

1.2 GENETICS OF SOLVENTOGENIC CLOSTRIDIA

Genetic analyses of saccharolytic clostridia have depended on mutagenesis, construction of clostridial vectors and the development of vector transfer methods.

1.2.1 Direct and Indirect Mutagenesis

The earliest and most basic approach used to understand clostridial genetics was the use of mutagenesis and the identification of mutants. A thorough examination of mutagenesis, mutagenic agents and clostridial mutations has been described respectively by Sebald (1993) and Jones (1993). The isolation and induction of clostridial mutants initially centered on the selection and the improvement of industrially important strains. Most studies have used physical and chemical mutagenic agents since the mode of action and the nature of mutations produced have been well characterized (Jones, 1993).

Bowring and Morris (1985), investigated the induction of stable chromosomal mutants in *C. acetobutylicum* by the use of direct and indirect mutagens. Direct mutagens cause

mutations by mispairing mechanisms involving either template or nucleotide precursors whereas indirect mutagens act by inducing a postreplication system that is error prone. They found that direct mutagens such as ethyl methane-sulfonate (EMS) and N-methyl-N'-nitro-N-nitrosoguanidine (MNNG), produced a significant increase in the relative induced mutation frequency of auxotrophic and antibiotic resistant markers. The indirect mutagens such as UV radiation and mitomycin C were ineffective. Subsequent studies have shown that for some clostridial species, UV radiation was successful in enhancing the mutation rate whereas MNNG was ineffective, thus there appears no correlation between the clostridial species and the mutagenic agent (Sebald, 1993). Through the use of mutagens, a variety of clostridial mutants have been isolated such as auxotrophic mutants, antibiotic-resistant mutants, granulose mutants, capsule mutants, solvent production mutants, phage resistant mutants and sporulation mutants (Jones and Woods, 1986). These mutants in turn have been used as genetic markers for analyzing gene function or for practical applications (Jones, 1993).

With the use of direct or indirect mutagens, it was difficult to ascertain whether a resultant mutant phenotype was caused by a specific mutation in a genetic region or was due to multiple mutations in different regions. This technique therefore does not allow for refined genetic analyses of clostridial genes. In order to study detailed gene function in solventogenic clostridia, an understanding of clostridial vectors and gene transfer methods is required.

1.2.2 Clostridial Vector Components

Plasmid or bacteriophage vectors which replicate in *Clostridium* are crucial elements for genetic analysis of solventogenic clostridia. It is necessary therefore to give an overview of some of the important clostridial vector components.

1.2.2.1 Clostridial plasmids and replicons

Plasmids of various sizes have been reported in *C. acetobutylicum*, *C. beijerinckii*, *C. butyricum*, *C. perfringens* and *C. saccharoperbutylacetonicum* (Jones and Woods, 1986; Young et al., 1989). Gene transfer was found only to occur between *C. perfringens* strains whereas the saccharolytic clostridia seemed to lack indigenous conjugative plasmids. In addition, plasmids from *C.perfringens* code for antibiotic resistance, bacteriocins, caseinase and other functions whereas plasmids from saccharolytic clostridia appear to be cryptic (Jones and Woods, 1986).

Conjugative transfer of several broad-host range plasmids from enterococci and streptococci to *C. acetobutylicum* has been reported (Yu and Pearce, 1986), and is summarized in Table 1.1 (adapted from Young et al., 1989).

Table 1.1: Conjugative transfer of enterococcal and streptococcal plasmids to *C. acetobutylicum*.

PLASMID	PROPERTIES	DONOR(S)[a]
pAMβ1	26.5 Kbp, Em^r	1,2,4,5
pIP501	35.0 Kbp, Em^r, Cm^r	1,5
pJH4	39.0 Kbp, Em^r, Cm^r, Km^r	1,5
pVA797	30.7 Kbp, Cm^r	2,3
pVA797::Tn917	36.0 Kbp, Cm^r, Em^r	1

[a]Donor organisms: 1, *Enterococcus faecalis*; 2, *Streptococcus lactis*; 3, *Streptococcus sanguis*; 4, *Bacillus subtilis*; 5, *C. acetobutylicum*.

Adapted from Young et al. (1989)

Plasmid pAMβ1 and pIP501 are broad-host range plasmids isolated from enterococci; which mediate resistance to the macrolides, lincosamides and streptogramin B antibiotic group (MLS^R). Of the two plasmids, the conjugative R factor, pAMβ1, has proven to be the most useful in developing clostridial vectors due to its promiscuous nature among gram-positive bacteria (Horaud et al., 1985; Swinfield et al., 1990). Transfer of pAMβ1 to *C. acetobutylicum* provided a functional resistance gene and replication origin which were subsequently used in the construction of *C. acetobutylicum* vectors for the

introduction of genetic information into *C. acetobutylicum* (Young, 1993a; Minton et al., 1993a). The pAMβ1 and pIM13 replicons, although derived from *E. faecalis*, are also termed clostridial because they function in that host.

Plasmid pAMβ1 can mobilize nonconjugative plasmids to several different gram-positive organisms, including *C. acetobutylicum* (Oultram et al., 1987). In the case of *C. acetobutylicum* this transfer occurs via the formation of plasmid co-integrates. Oultram et al. (1987) inserted the *ermAM* gene from pAMβ1 into a small nonconjugative, integrational plasmid lacking gram-positive replication functions. When transformed into a *B. subtilis* strain harbouring pAMβ1, it became established as a co-integrate molecule which could be transferred to and maintained in *C. acetobutylicum*. The cumbersome experimental procedure does not allow this procedure to be widely employed for gene-transfer experiments.

Replicon probe vectors have been used to identify replication functions of other clostridial plasmids. Some of the plasmids that were studied using this approach were pCB101, pCB102 and pCB103 from *C. butyricum* and pCPA1 from *C. paraputrificum* (Minton et al., 1993b). The broad-host range streptococcal plasmid, pWV01 was also shown to replicate in *C. beijerinckii* (Williams et al. 1990a).

Gruss and Ehrlich (1989) have shown that many of the small plasmids indigenous to gram-positive bacteria form a highly interrelated family known as "ss DNA plasmids". These plasmids all replicate via a single strand intermediate, using the rolling circle mechanism of replication. A plasmid-encoded Rep protein initiates DNA replication at a fixed DNA position termed the "plus" origin. On the basis of homology between the Rep proteins and the "plus" origins of replication the "ss DNA plasmids" fall into at least three subfamilies. The first family is represented by pC221 and pT181 (staphylococcal plasmids), the second family by, pE194 and pLS1 (staphylococcal and

streptococcal plasmids) and the last family is represented by pC194 and pUB110 (*S. aureus* plasmids).

Analyses of the replication regions of some of the clostridial plasmids indicate that they do fall into one or more of the "ss DNA plasmid" subfamilies but also possess several unique characteristics (Minton et al., 1993a).

In contrast, analysis of the replicon of pAMβ1 has shown that it undergoes unidirectional theta replication as opposed to the common strategy of rolling circle replication (Bruand et al., 1991). The different plasmid replicons and the associated replication mechanisms play an important part in clostridial vector development and stability.

1.2.2.2 Clostridial Selective markers

None of the indigenous plasmids that have been isolated from solventogenic clostridia have been found to specify antibiotic resistance and this has led to the use of selective markers from heterologous sources. Most of these selective markers belong to the MLS^R antibiotic group (Papoutsakis and Bennett, 1993). The antimicrobial activity of the MLS^R group is due to inhibition of ribosome-dependent protein synthesis.

The macrolide, erythromycin, is the most commonly used selective marker for the selection of clostridial transformants. For the construction of clostridial vectors, the *ermAB* gene of plasmid pAMβ1 is usually employed due to its extensive physical and functional characterization (Brehm et al., 1987; Swinfield et al., 1990). Erythromycin decomposes at pH values around or below pH 5.0 into a form which does not have antimicrobial activity (Mermelstein and Papoutsakis, 1993b). In *C. acetobutylicum* fermentations, solvent production is most substantial at pH values below 5.3 (Jones and Woods, 1986). If clostridial plasmids carrying the erythromycin marker are used in these

fermentations, plasmids will not be selectively maintained over time as erythromycin decomposition follows decreasing pH levels. Physiological studies with clostridial vectors require maintenance of selective pressure because plasmid free cells would ultimately overtake the recombinant culture if the antibiotic selection becomes unstable. It has also been observed in certain cases that *C. acetobutylicum* ATCC 824 plasmid vectors containing cloned homologous genes are considerably more unstable than the vector alone (Lee et al., 1993).

Mermelstein and Papoutsakis (1993b) sought to overcome these fermentation problems in recombinant *C. acetobutylicum* ATCC 824 by using alternate readily available and affordable MLS antibiotics that would maintain selective pressure in low pH at appropriate concentrations that would not inhibit solvent formation. They tested a number of macrolides of which clarithromycin, midecamycin and tylosin were found to provide selective pressure for MLS^R plasmid maintenance at low pH values and there was no antibiotic degradation by cell free fluids from late stationary phase. Of the three antibiotics, clarithromycin did not inhibit solvent formation in low pH fermentations and was found to be useful for plasmid maintenance in several low pH batch fermentations (Mermelstein et al, 1993).

Although the chloramphenicol and tetracycline resistance genes have been successfully employed in *C. perfringens*, the use of these selective markers in *C. beijerinckii* has proved to be problematic. It has been observed that anaerobic clostridia frequently become resistant to chloramphenicol due to enzymatic inactivation of the antibiotic by either reduction of the aryl group or O-acetylation (Staudenbauer and Dubbert, 1993). The use of thiamphenicol as an alternative to chloramphenicol to circumvent this problem has only had limited success (Minton et al., 1993b). The high number of *C. beijerinckii* cells, which are naturally resistant to tetracycline in a given population, makes this selective marker unsuitable for vector construction (Minton et al., 1993a).

1.2.2.3 Vector Stability

A crucial feature of any plasmid vector is the ability for that vector to maintain genetic and structural integrity during and after recombinant manipulation. Vector stability is dependent on three parameters, namely, structural stability, segregational stability and stability determinants.

Clostridial vectors based on replicons from the "ss DNA plasmid" family which use the rolling circle mode of replication have the tendency to be easily perturbed during vector construction or cloning and the use of the these plasmids in alternative hosts often leads to structural or segregational instability (Gruss and Ehrlich, 1989). Insertion of certain DNA fragments into rolling circle (rc) plasmids also tends to lead to accumulation of high molecular weight molecules (HMW), i.e. linear head to tail repetitions of the plasmid which alter the conformation and copy number of the plasmid (Gruss and Ehrlich, 1988). This makes clostridial vector design using "ss DNA" plasmid replicons difficult because vector stability can be easily compromised by minor factors such as positioning of cloning sites or change of plasmid host. However, this instability feature of rc plasmids makes them useful for chromosomal gene inactivation as a result of single or double crossover recombination between the chromosomal gene and a cloned, mutated plasmid-borne derivative of the gene.

Plasmids that replicate using the unidirectional theta mechanism such as pAMβ1 have been shown to result in 1000-fold greater structural stability in *B. subtilis* over a "ss DNA plasmid" (Janniére et al., 1990). In addition, the pAMβ1 replicons in clostridial vectors allowed efficient cloning and maintenance of large DNA fragments. Bi-functional *E. coli*/*Clostridium* shuttle vectors based on pAMβ1 replicons were also structurally and segregationally stable in most hosts (Oultram et al., 1988a).

Plasmid copy number and expression of cloned genes can also affect vector stability. Mermelstein and Papoutsakis (1993a) found that plasmids pSYL2, pFNK1, pFNK3,

pFNK5 have different copy numbers in *C. acetobutylicum* ATCC 824 and also different segregational stabilities. It was found that both plasmid copy number and segregational stability affected vector stability adversely. Plasmids pFNK3 and pFNK5, which carry fermentative genes had lower copy numbers and were segregationally less stable than pFNK1 from which they were derived.

In *C. acetobutylicum*, segregational instability and not structural instability has been a problem observed with vectors based on pIM13, pSC86, pCB101, pCPA1 and pAMβ1 (Minton et al., 1993b). In order to obtain segregational stability, stability determinants have been made use of in the construction of cloning vectors. Plasmid pAMβ1 carries a *res* gene and two ORFs, designated H and I, which shared identity with a recombinase and a DNA topoisomerase I respectively (Swinfield et al., 1991). The insertion of the two ORFs into clostridial vector pMTL500E resulted in significantly improved segregational stability in *C. acetobutylicum* and *B. subtilis*. This stability element could be utilized in the construction of stable clostridial vectors.

Blaschek and White (1995) characterized an M13-like genetic system for clostridia based on an *E. coli/C. acetobutylicum* pCAK1 phagemid expression vector. They previously identified a filamentous virus-like particle, CAK1, which was continuously extruded from *C. acetobutylicum* NCIB 6444 as single-stranded DNA during cell growth (Kim et al., 1990). The phagemid pCAK1 showed functionality in the *E. coli* host system by generating "ss DNA" without impairing cell viability and was also successfully introduced into *C. beijerinckii* and *C. perfringens* (Kim and Blaschek, 1993). The non-lytic characteristic of this genetic system could be used to construct versatile cloning vectors and may in future provide a more sophisticated tool for genetic analyses of clostridia.

1.2.3 Vector Transfer Methods

The development of wide-host range, non-specific gram-positive vectors is closely linked to vector transfer methods. There are different methods for the transfer of plasmid vectors into *C. beijerinckii*, each with different methodological and procedural advantages and disadvantages.

1.2.3.1 Transformation of Protoplasts

The first successful transformation of *C. acetobutylicum* protoplasts, was obtained by Reid et al. (1983) using phage CA1 DNA. There are numerous obstacles to transformation of protoplasts such as autolysin activity, nucleases and restriction endonucleases (Young et al., 1989). These obstacles coupled together with the cumbersome transformation procedure have prevented protoplast transformation from being widely used for the transfer of clostridial vectors.

1.2.3.2 Conjugative Transfer

Conjugation between bacteria is dependent on self-mobilizable plasmids that contain a transfer (*tra*) and mobilizable (*mob*) region that allows transfer and mobilization from one bacterium to another. Self-mobilizable plasmids from gram-negative bacteria belonging to incompatibility group P have a wide host range and their conjugation machinery shows a lack of specificity. This enables IncP plasmids (RP4, RK2 and R751) to transfer and mobilize nonconjugative plasmids from gram-negative bacteria to a wide range of organisms including gram-positive bacteria, yeasts and plants (Young, 1993a). IncP plasmid conjugative mobilization involves the interaction of a trans-acting protein complex with a cis-acting site on the plasmid denoted *oriT* (origin of transfer). A nick is made in the *oriT* region to generate a relaxosome complex after which a single DNA strand is transferred to the recipient starting at the 5' terminus nick site and accompanied by rolling circle replication of the plasmid in the donor (Willets and Wilkins, 1984). Thus IncP plasmids can mobilize other plasmids which have compatible

mob, *nic* and *oriT* sites but lack a *tra* region. This mobilization of nonconjugative plasmids can occur in three ways and two mechanisms involve the formation of co-integrate DNA molecules. Mobilization via co-integrate molecules can arise firstly as intermediates in the transposition of certain mobile genetic elements from one replicon to another and secondly, co-integrate molecules may be generated as a result of RecA-dependent recombination between identical sequences present on both replicons. Mobilization may also occur after the productive interaction of the origin-nicking complex, synthesized by a conjugative plasmid, with the *oriT* site on the unrelated co-resident replicon.

Trieu-Cuot et al. (1987) mobilized the nonconjugative shuttle plasmid pAT187 from *E. coli* to a number of gram-positive bacteria. Plasmid pAT187 has a pBR322 backbone, with pAMβ1 replicon functions, the *aphA-3* gene of a *Campylobacter coli* plasmid which confers kanamycin resistance and the *oriT* site from IncP plasmid RK2. Jennert et al. (2000) also used IncP-mediated conjugation to mobilize plasmids from *E. coli* to *C. cellulolyticum*.

Williams et al. (1990b) constructed a number of *E. coli*/*C. beijerinckii* shuttle vectors based on the pMTL20 backbone (Chambers et al., 1988; see Chapter 3 for further details). This was done by inserting the *oriT* segment of RK2 into these vectors in order to investigate how different gram-positive replication origins function in *C. beijerinckii*. Plasmid pCTC3 was constructed on the backbone of plasmid pGK13, which contained the replication origin of plasmid pWV01 of *S. cremoris* and the Em^r gene of plasmid pE194. Plasmid pCTC511 contained the replication functions of plasmid pCB101 from *C. butyricum* NCIB 7423 and the Em^r gene from plasmid pAMβ1. Plasmid pCTC40/41 was constructed by inserting the pAMβ1 replication origin into pMTL30/31. These plasmids including pAT187 and pCTC1 were transferred to *C. beijerinckii* by conjugative mobilization with an appropriate *E. coli* donor under anaerobic conditions. They found that all three gram-positive replicons functioned in *C. beijerinckii*. All the

pCTC plasmids transferred at frequencies higher than pAT187. The mobilization frequencies of plasmids pCTC40/41 were the same as plasmid pCTC1. Plasmid pAT187 appeared to replicate in *C. beijerinckii* at lower copy number than the pCTC plasmids. Plasmid pCTC511 was transferred to *C. beijerinckii* at a frequency slightly lower than those obtained for the various pCTC plasmids but replicated at a high copy number. Plasmid pCTC3 was mobilized at the same frequency as the other pCTC plasmids but it appeared that this plasmid replicates at lower copy number in *C. beijerinckii*. In addition plasmid pCTC40, which differed from pCTC41 in the orientation of the *oriT*, tended to act as a target for IS*1* insertion. The IS*1* transposition inserted into the same position within the pAMβ1 replication region. If *E. coli* containing pCTC40 was not kept at 37°C it was found to harbour enlarged derivatives when subcultured from the bench although this did not impair its ability to replicate in *C. beijerinckii*.

Plasmid transfer efficiency to *C. acetobutylicum* varies considerably with the donor organism (Oultram et al., 1987). In the case of plasmid pAM_1 this has been ascribed to the tendency of the plasmid to suffer large specific deletions in genetic regions required for conjugative plasmid transfer. Transfer frequencies were also found to vary from one experiment to another depending on the physiological state of the bacteria.

1.2.3.3 Electroporation

Oultram et al. (1988a) constructed a plasmid shuttle vector, pMTL500E for use in *C. beijerinckii* based on the pMTL20 backbone (Chambers et al., 1988; see Chapter 3, Figure 3.1) by incorporating the *erm* gene of pAMβ1 for selection and the pAMβ1 replication region, which promotes replication at high copy number in *C. beijerinckii* (Swinfield et al., 1990). Oultram et al., (1988a) successfully transformed *C. beijerinckii* with plasmid pMTL500E using electroporation. Minton et al. (1990) used the same electroporation methodology with certain modifications to investigate the ability of

chimeric plasmids carrying different gram-positive replicons to replicate in *C. beijerinckii*. They also found that transformation frequencies were reproducibly higher when plasmid DNA was prepared from *B. subtilis*, rather than *E. coli*. Lee et al. (1992) successfully electroporated *C. beijerinckii* using a different procedure, while Kim and Blaschek (1993) used a procedure based on a 10% PEG electroporation buffer solution with different electroporation settings. All three protocols however gave similar transformation frequencies.

The reproducibility of efficient electro-transformation in *C. beijerinckii* was, and remains, inconsistent (Minton et al., 1993b), although the factors contributing to this have not been determined. In *C. acetobutylicum* ATCC 824, the restriction endonuclease *Cac824*I is a major barrier to electrotransformation (Mermelstein et al., 1992). Mermelstein and Papoutsakis (1993a) protected plasmids which were to be electroporated with *in vivo* methylation in *E. coli* by the *B. subtilis* phage ϕ3T I methyltransferase. Three shuttle vectors pIMP1, pSYL2 and pSYL7 which were methylated in this manner, were successfully electroporated into *C. acetobutylicum* ATCC 824 and were not able to be introduced when unmethylated. Recently, Tyurin et al. (2000) successfully transformed *C. acetobutylicum* ATCC 824 by electroporation using high-voltage radio frequency modulated square pulses.

The knowledge of clostridial vectors and vector transfer methods provide a foundation for the following section which will focus on selected mechanisms of gene inactivation in gram-positive bacteria.

1.2.4 Selected Mechanisms of gene inactivation in Gram-Positive bacteria

Nonreplicative vectors are important tools for the study of gene function in most bacteria. These suicide vectors allow for inactivation of genes specifically or non-specifically depending on the nature of the suicide vector. In order for these suicide

vectors to be successful, high transformation efficiencies are required to allow detection of low frequency events such as transposition or recombination into the chromosome. In addition some of these suicide vectors have conditionally active replicons, which are restricted in their host range.

The pMTL30 vectors have successfully been used to inactivate genes in *C. beijerinckii* by homologous recombination. Wilkinson and Young (1994) cloned different sized regions of *C. beijerinckii* target DNA into pMTL30 and transferred them to *C. beijerinckii* via conjugation with *E. coli*. Since vector pMTL30 lacks a gram-positive replicon, only a Campbell-like recombination event between the target DNA and the homologous target in the genome will allow the plasmid to be maintained. Both single and amplified copies became established in the *C. beijerinckii* genome and there was no clear correlation between the size of the DNA regions and the frequency of plasmid establishment. Since Campbell-like integration events generate a duplication of target sequences, a reversal of the integration event was possible. Thus it was found that in the absence of erythromycin selection, erythromycin-sensitive recombinants segregated after several generations, albeit at a low frequency. Reid et al. (1999) disrupted the transcriptional repressor, *scrR*, and sucrose-6-phosphate hydrolase, *scrB* of the sucrose operon of *C. beijerinckii* using the pMTL30 vectors.

In *C. acetobutylicum* ATCC 824 the phosphotransacetylase, butyrate kinase and aldehyde/alcohol dehydrogenase genes were disrupted with the use of nonreplicative plasmids which were transferred via electroporation (Green et al., 1996; Green and Bennett, 1996).

Genetic manipulation of lactic acid bacteria has similar problems associated with it as when compared to clostridia in that there is no efficient genetic system for gene inactivation. In order to circumvent this problem, Maguin et al., (1992) uncoupled transformation and recombination. This was achieved by making use of a

thermosensitive (Ts) replicon. Therefore, even if the transformation frequency was low, a Ts replicon could first be established and propagated in a large bacterial population at the permissive temperature. A shift to the non-permissive temperature would prevent plasmid replication and allow the selection of low frequency events like transposition and recombination.

Maguin et al. (1992) constructed pVE6002, a replication-thermosensitive mutant of the broad-host range replicon, pWV01 originally isolated from *Lactococcus lactis* subsp. *cremoris*. Plasmid pWV01 replicates via the rolling-circle mechanism in a large number of gram-positive bacteria, including *C. beijerinckii* and in *E. coli* (see section 1.2.2.1). Replication of the pVE6002 plasmid was fully thermosensitive above 35°C in both gram-positive and gram-negative bacteria and the mutation conferring the thermosensitivity function was found to reside on a DNA fragment encoding RepA. Three derivatives of plasmid pVE6002 were made containing the multi-cloning site of the pBluescriptSK$^+$ plasmid and different combinations of the Em^r and Cm^r resistance markers.

The above study was used by Biswas et al. (1993) to create a high efficiency system for single and double-crossover homologous integration in gram-positive bacteria. They used *L. lactis* as the model system and plasmid pVE6004 (which was re-named pG$^+$host4) as the basis for the development of this gene inactivation and replacement system (See Chapter 3 for further details).

Law et al. (1995) also used a combination of Ts derived vectors to develop a system to generate chromosomal mutations in *L. lactis* and fast analysis of the targeted genes. In their system they combined the use of two pWV01 derivatives, namely the Ts plasmid pVE6007 which provided a pWV01 RepA in trans and a Ori$^+$ vector, pORI19 which cannot replicate in the absence of RepA. Transformation of *L. lactis* with a pORI19 bank of lactoccal chromosomal fragments at the permissive temperature allows

replication of several copies of a recombinant plasmid from the bank within a cell because of provision in trans of RepA-Ts from pVE6007. A non-permissive temperature shift results in the loss of pVE6007 and the integration of the pORI19 derivatives at high frequencies into the *L. lactis* chromosome. This system allowed for stable integrant formation at high efficiencies, identification of mutants at optimal growth temperature and easy recovery of the integrated plasmid from the chromosome. Previous studies (Williams et al., 1990b) have shown that the replicon of pWV01 replicates in *C. beijerinckii*, which suggests that the theremosensitive system, which is based on this replicon could possibly function in *C. beijerinckii*. The Ts vectors can be established at low temperatures in a large bacterial population before selection of integrants and can be used for bacteria that are poorly transformable. The low shutoff temperature of pG^{+}host plasmids allows the integration system to be used in bacteria, which have a limited temperature growth range. However plasmid replication, thermosensitivity, background integration and the level of drug resistance must be determined for each new host.

1.3 CARBOHYDRATE METABOLISM IN BACTERIA

Carbohydrate metabolism in gram-negative and gram-positive bacteria has been extensively reviewed (Stülke and Hillen, 2000; Stülke and Hillen, 1999; Saier, 1996; Saier et al., 1996; Saier et al., 1995, Hueck and Hillen, 1995). In the following sections, we will focus on the key mechanisms and components involved in carbohydrate metabolism in bacteria.

1.3.1 Carbon Catabolite Repression

In some bacteria, the regulation of carbohydrate utilization and metabolism is linked to two mechanisms. The first mechanism, carbon catabolite repression (CCR) is a regulatory mechanism whereby the cell coordinates the metabolism of carbon and energy

sources in order to maximize substrate efficiency and regulate secondary metabolic processes (Chambliss, 1993).

CCR is the phenomenon by which the presence of a preferred carbohydrate, usually glucose, represses the synthesis of enzymes required for the metabolism of other less rapidly utilized carbohydrates. This selective repression of enzyme synthesis allows priorities to be established by the cell with respect to the utilization of various carbon and energy sources. The synthesis of secondary metabolites is subject to CCR directly or indirectly. Readily metabolized carbohydrates also repress developmental pathways such as sporulation, synthesis of certain extracellular enzymes and toxin production.
The second regulatory mechanism of carbon source utilization is called induction and in most cases is coupled to CCR. In this mechanism, the bacterium will express enzymes that are required for utilization of specific carbohydrates when the nutrient is present or when more preferable carbohydrates are absent. Wastage of metabolic energy by the bacterium can therefore be avoided by the use of induction by a particular substrate in mixed carbohydrate environments.

Catabolite metabolism has been studied in both gram-negative and gram-positive bacteria, and although key differences between the two groups exist, an essential component common to both groups is the carbohydrate-transporting phosphotransferase system (PTS). Despite mechanistic and structural differences in the PTS system of both groups, one of the dominant features of the PTS system is its sensory and regulatory involvement in CCR.

The PTS system is a functionally complex and diverse system that transports sugars across the cell membrane, phosphorylates the transported sugars and ultimately determines the metabolic status of the bacterium. The regulatory consequences of components of the PTS system are cellular control of various physiological processes such as transport, metabolism, storage of carbon sources, chemotaxis, coordination of

carbon metabolism with nitrogen metabolism and oxidative versus fermentative carbohydrate utilisation (Postma et al., 1993).

The PTS consists of two energy-coupling soluble proteins, Enzyme I and a heat-stable protein, HPr, and various sugar-specific Enzyme II complexes, each of which comprises a membrane-bound permease. The Enzyme II complex in turn contains at least three protein domains termed IIA, IIB and IIC. In some permease complexes the domains are linked whereas in other complexes they compromise separate polypeptide chains. Two of these protein domains (IIA and IIB), act as a phosphoryl transfer chain that energizes carbohydrate uptake. The other domain spans the membrane. In the reaction scheme below, phosphoenolpyruvate (PEP) acts as a phosphate donor to the phosphoryl-transfer proteins, Enzyme I, HPr, IIA and IIB.

$$\text{PEP} \rightarrow \text{Enzyme I} \rightarrow \text{HPr} \rightarrow \text{IIA} \rightarrow \text{IIB} \rightarrow \text{sugar-}P \qquad \text{IIC}$$

Enzyme IIC is the membrane-embedded sugar permease that catalyses the transport of the sugar across the membrane as well as transferring the phosphate from the phosphorylated Enzyme IIB to the incoming sugar.

The relationship between the PTS system and CCR in bacteria differs amongst bacteria due to the presence or absence of structural and regulatory components. CCR in *C. beijerinckii* has been studied mostly at the physiological level with very little information available on the regulatory elements that form part of the CCR mechanism in this organism. In order to gain a better understanding of CCR in clostridial species and for comparative purposes, it will be useful therefore to review the basic CCR mechanism in *E. coli*, *Bacillus* and other low-guanine/cytosine (GC) Gram-positive bacteria.

1.3.2 Catabolite Repression in *E. coli*

Previously *E. coli* was used as the model system for the study of CCR in bacteria. The CCR model system was based on the *E. coli* lactose (*lac*) operon and focused on the interaction of glucose and lactose on CCR components. These studies suggested that CCR in *E. coli* is based on interference with a positive regulatory mechanism (Chambliss, 1993).

According to available data, a model of CCR in *E. coli* can be proposed. The glucose specific EIIA ($EIIA^{Glc}$) of the PTS encoded by *crr* is a key element for glucose repression of CCR sensitive genes (Stülke and Hillen, 1999). In the absence of glucose, phosphorylated $EIIA^{Glc}$ increases adenylate cyclase activity. The activity of adenylate cyclase converts ATP into cAMP with the liberation of pyrophosphate. The synthesized cAMP forms a complex with the catabolite repressor protein (CRP) and this complex binds to a specific site in the promoter region of CCR-sensitive genes or operons. This binding activates transcription of the CCR-sensitive promoters which have been induced. The consensus CRP-binding site, is generally located upstream of the -35 regions of promoters subject to CCR. The intracellular level of cAMP determines the binding of CRP at CRP-binding sites. Thus when glucose or another PTS sugar is present it will be taken up rapidly by the cell and will act as a sink for the transfer of phosphate through the PTS, and consequently the pool of $EIIA^{Glc}$ will consist predominantly of the non-phosphorylated form. In the absence of phosphorylated $EIIA^{Glc}$, adenylate cyclase activity will drop to a low basal level, thus causing a decrease in cAMP concentration to a level insufficient for complex formation with CRP. This will prevent CRP binding to CCR-regulated promoters even if these operons are induced. Biochemical studies have not managed to demonstrate regulation of adenylate cyclase by $EIIA^{Glc}$ and a study has suggested that other factors besides phosphorylated $EIIA^{Glc}$ contribute to adjusting cAMP levels (Notley-McRobb et al., 1997). One such factor is the PEP:pyruvate ratio which controls the phosphorylation state of $EIIA^{Glc}$ and thus the activity of adenylate cyclase (Hogema et al., 1998).

Non-phosphorylated $EIIA^{Glc}$ is also involved in inducer exclusion whereby glucose blocks the uptake of a number of carbohydrates. This prevents intracellular concentrations of these carbohydrates from reaching levels adequate for the induction of synthesis of enzymes involved in their metabolism. $EIIA^{Glc}$ binds to sugar permeases in the presence of the cognate sugar and also to glycerol kinase and inhibits the activities of the permeases (Stülke et al., 1998).

With respect to the *E. coli* lactose operon, it was assumed that cAMP-CRP dependent activation of transcription was the most important mechanism of CCR. Inada et al. (1996) have shown that cAMP levels are not correlated to the presence or absence of glucose and cAMP-dependent regulation of the *lac* operon results from inducer exclusion mediated by the PTS. This evidence therefore suggests that the mechanism of inducer exclusion plays a greater role than glucose repression of the *lac* operon. The mechanism of CCR and inducer exclusion in *E. coli* is shown in Figure 1.2.

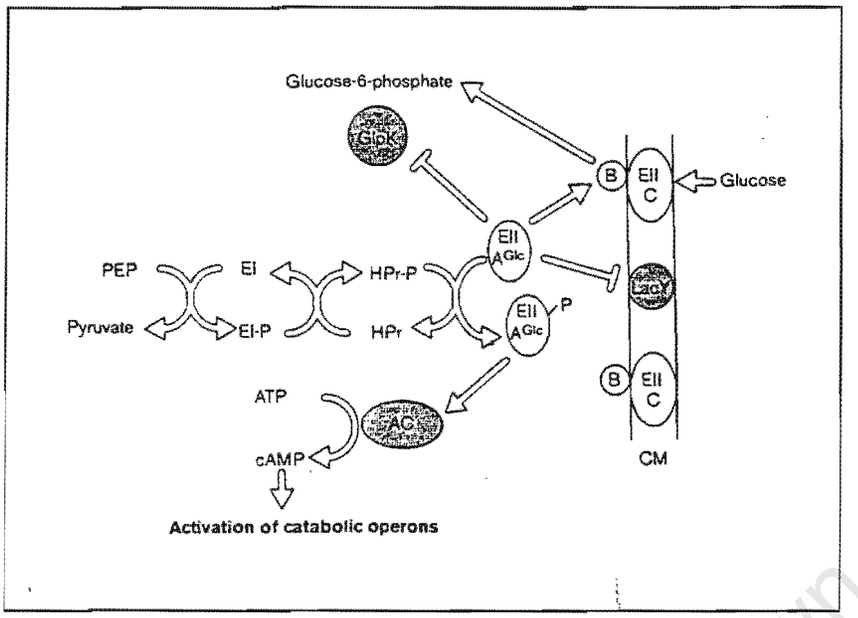

Figure 1.2: CCR in *E. coli* is mainly mediated by the glucose-specific EIIA of the PTS. The core of the PTS involves the flow of phosphate from PEP to the exogenous carbohydrate simultaneously with its transport. This flow is mediated by EI, HPr and carbohydrate-specific EII, which includes different subunits (A, B and C). In the presence of glucose, $EIIA^{Glc}$ binds and inactivates the lactose permease (LacY) and the glycerol kinase (GlpK). In the absence of sugars, phosphorylated $EIIA^{Glc}$ activates adenylate cyclase (AC) to result in Crp-mediated transcriptional activation of catabolic operons. CM, cell membrane. Adapted from Stülke and Hillen (1999).

A cAMP-independent catabolite repression mechanism was also found to occur in *E. coli* (Saier, 1996). In this mechanism the catabolite repressor/activator (Cra) protein [formerly designated the fructose repressor FruR] pleiotropically functions to control the direction of the metabolic carbon flux. The Cra protein controls the synthesis of numerous enzymes in a systematic way: Cra repressed enzymes catalyzing sugar fermentation (key enzymes of glycolysis, the Entner-Doudoroff pathway and enzymes initiating sugar catabolism) and Cra activated enzymes allowing substrate oxidation (key enzymes of glucogenesis, the Krebs cycle, the glyoxalate shunt and electron flow).

Cra is a homologue of the lactose repressor of *E. coli* (LacI) and is a member of the LacI-GalR family of transcriptional regulatory proteins with N-terminal helix-turn-helix DNA binding motifs and C-terminal ligand binding domains. The latter domains are

homologous to periplasmic sugar binding receptors such as those specific for arabinose, ribose and galactose. The tetrameric form of Cra binds asymmetrically to an imperfect palindromic consensus sequence (the ideal Cra operator). Operators that are activated by Cra are usually found upstream of the promoter and operators which are repressed by Cra are found downstream or overlapping the promoters. Cytoplasmic glycolytic intermediates bind to Cra resulting in Cra dissociating from the DNA, which causes either catabolite activation or repression (see Figure 1.3). Thus Cra mediates both catabolite repression and catabolite activation by a unified mechanism that is independent of cAMP.

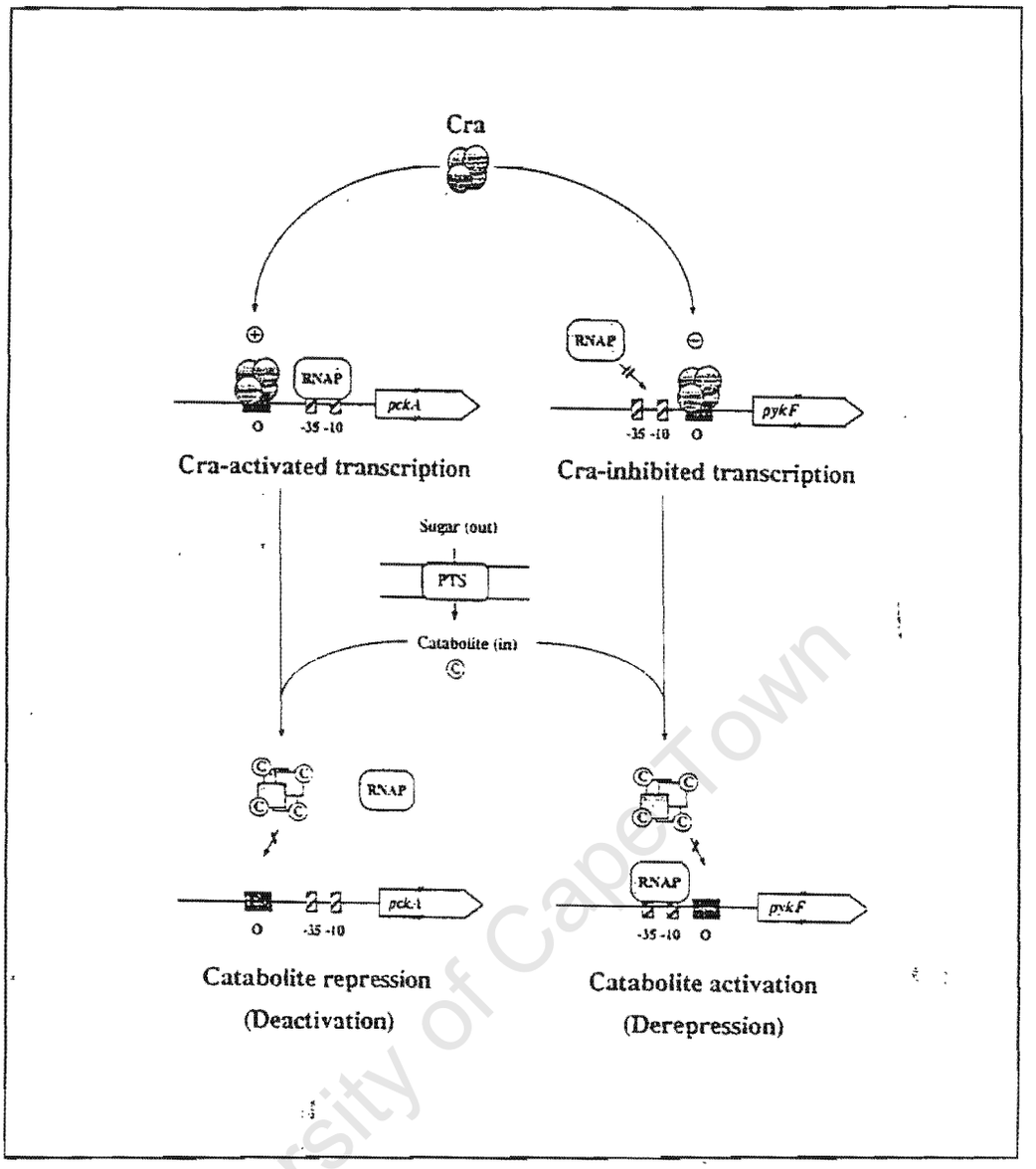

Figure 1.3: Model illustrating the sensory transduction pathway for Cra-mediated catabolite repression (left) and catabolite activation (right) in *E. coli*. In the unliganded form, Cra (formely designated FruR) binds to the operator (O) regions of the target operons. When the Cra operator is upstream of the RNA polymerase binding site, activation of transcription is observed. When the operator overlaps or is downstream of the RNA polymerase binding site, inhibition of transcription is observed. The presence of an exogenous sugar (e.g., glucose) results in uptake via the phosphotransferase system, producing sugar phosphates and cytoplasmic glycolytic catabolites (C). These catabolites (e.g. fructose-1-P; fructose-1,6-bisP) bind to the tetrameric Cra protein, causing it to dissociate from the DNA. Dissociation reverses the activating effect of Cra as in the case of the *pckA* gene encoding phosphoenolpyruvate carboxylase (catabolite repression; left), and reverses the inhibiting effect of Cra, as in the case of the *pykF* gene encoding pyruvate kinase (catabolite activation; right). Adapted from Saier (1996).

1.3.3 Catabolite Repression in *B. subtilis*

B. subtilis has been the model for the study of catabolite repression in gram-positive bacteria. Regulation of carbon catabolism in *Bacillus* species has been extensively reviewed (Stülke and Hillen, 2000; Fisher, 1991). Exposure of *B. subtilis* to conditions of nutrient depletion results in major changes in the programme of gene expression in the cell, which include competence for transformation, initiation of sporulation and synthesis of new products (Henkin et al., 1991). Sporulation is repressed by carbon-, nitrogen-, and phosphorus- containing metabolites during rapid growth (Chambliss, 1993). The mechanism by which gene expression is controlled in response to nutrient availability in *B. subtilis* differs fundamentally from that of the CCR mechanism in *E. coli*. In none of the gram-positive organisms does cAMP play a central role and CCR seems to be mediated by a negative regulatory mechanism.

Three components involved in CCR of gene expression in *B. subtilis* have been identified primarily from studies on α-amylase synthesis. These components are the CcpA, a GalR-LacI family regulator; the *cre* element, a cis site to which the CcpA binds and HPr, a PTS protein. Each component will be dealt with separately and a proposed model of molecular interaction between these components will be explained.

1.3.3.1 Catabolite Responsive Elements (*cre*)

Two independently isolated *B. subtilis* mutants which synthesize α-amylase in the presence of glucose have been isolated and characterized (Henkin et al., 1991). The mutations responsible for loss of glucose repression were designated gra-5 and gra-10. Both mutations were cis-acting and consisted of an identical single base-pair alteration close to the transcriptional start-point of the α-amylase structural gene, *amyE*. These mutations resulted in the derepression of *amyE* transcription in the presence of glucose. The sequence in the region near the site of the gra-5 and gra-10 mutations exhibited dyad symmetry that showed structural similarity to the operator regions of the *E. coli gal* and

lac operons. The sequences of *cre*, *galO* and *lacO* diverged only at positions 4 and 5 (see Table 1.2). The presence of an operator-like region at the transcription start site suggested that this operator-like region might be the site at which a repressor protein would bind in the presence of glucose.

Table 1.2: Comparison of the consensus *cre*, *galO* and *lacO* sequences.

	7	6	5	4	3	2	1
cre	T	G	T/A	N	A	N	C*
galO	T	G	T/G	A	A	N	C*
$lacO_{id}$	T	G	T	A	A	G	C*

Comparison of the left-hand sides of the consensus *cre* sequence with *galO* and ideal $lacO_{id}$. The asterisks indicate the centers of the palindromic symmetry. Positions are numbered. Adapted from Hueck & Hillen (1995).

Mutational analyses of the CCR-mediating sequence in *amyE* was carried out and a consensus sequence for this 14 bp palindromic catabolite-responsive element (*cre*) was proposed, Table 1.3 (Hueck and Hillen, 1995).

Table 1.3: Active *cre* sequences from selected gram-positive bacteria.

Organism	Gene	*Cre* sequence	TCS[a]	TLS[b]	Reference
B. subtilis	*acsA*	TGAAAGCGTTACCA	+37	+8	Grundy et al. (1994)
B. subtilis	*acuA*	TGAAAACGCTTTAT	-33	-61	Grundy et al. (1994)
B. subtilis	*amyE*	TGTAAGCGTTAACA	-3	-124	Weickert and Chambliss (1990)
B. subtilis	*gntR*	TGAAAGCGGTACCA	+141	+107	Miwa et al. (1997)
B. subtilis	*hutP*	TGAAACCGCTTCCA	+203	+170	Wray et al. (1994)
B. subtilis	*licS*	AGAAAACGCTTTCA	ND[c]	-33	Krüger et al. (1993)
B. subtilis	*xylA*	TGGAAGCGTAAACA	+134	+35	Kraus et al. (1994)
B. megaterium	*xylA*	TGAAAGCGCAAACA	+123	+35	Gösseringer et al. (1997)
S. xylosus	*xylA*	AGTAAGCGTTTACA	-42	-93	Sizemore et al. (1992)
B. subtilis	*ackA*	TGTAAGCGTTATCA	-64	-91	Grundy et al (1993b)
Consensus sequence[d]		TGWAARCGYTWWCA			Weickert and Chambliss (1990)

Adapted from Hueck and Hillen (1995). Deviations from the Weickert consensus sequence are underlined. Distances are given with respect to the 5'-nucleotide of the cre element.

a TCS, transcriptional start site.

b TLS, translational start site.

c ND, not determined

d Y represents C or T, R represents A or G, W represents A or T and N represents A, C, G or T.

A fusion of *cre*-containing DNA fragments to unrelated constitutive promoters conferred CCR on expression of indicator genes driven by these promoters (Hueck and Hillen, 1995). This result and the observation that *cre* elements were found in a number of operons involved in carbon catabolism indicated that *cre* sequences and hence the mechanism of CCR, may be globally distributed in gram-positive bacteria. It has been subsequently demonstrated that sequences similar to the *cre* consensus element mediate CCR in a number of other genes from *B. subtilis*, *B. megaterium, S. xylosus, L. casei, L. pentosus, Lactobacillus plantarum* and *S. mutans* (Stülke and Hillen, 1999).

The *cre* elements have been found in promoter regions, either downstream of the promoter within the coding region or upstream of the promoter (Hueck and Hillen, 1995; Krüger and Hecker, 1995). Repression may result from either prevention of transcription initiation, a transcriptional roadblock, or interference with interaction between RNA polymerase and an activator. The *cre* sites of the *amyE* and *acu* genes lie within or adjacent to the promoter region (Table 1.3) where the binding of a regulatory protein could interfere with transcriptional initiation (Hueck and Hillen, 1995). There is evidence that suggests that inhibition of transcription is achieved through the regulatory protein interacting with the RNA polymerase since CCR of *amyE* in strains carrying a mutation in *crsA* (encodes sigma factor sigma A) is relieved (Kim et al., 1998). The downstream *cre* sites located in the *B. subtilis hut* and *gnt* operons, result in RNA polymerase stalling at these sites and require the transcriptional repair-coupling factor Mfd to displace the stalled RNA polymerase (Zalieckas et al., 1998a). The Mfd protein promotes strand-specific DNA repair by displacing RNA polymerase stalled at the nucleotide lesion and directing the (A)BC exinuclease to the DNA damage site. Mutations in *mfd* partially relieve CCR at *cre* sites located downstream of the

transcriptional start site of the *hut* and *gnt* genes but do not affect CCR at *cre* sites located at the promoters. However, although the *acsA cre* element is also situated downstream of the transcriptional start site, *acsA* expression was not affected by Mfd modulation (Zalieckas et al., 1998b).

Furthermore, mutational analyses of nucleotides adjacent to the *acsA cre* site showed that higher levels of CCR were obtained by *cre* sites flanked by A+T-rich sequences than by G+C nucleotides indicating that the sequence context of the *cre* elements affects the level of CCR (Zalieckas et al., 1998b).

Miwa et al. (2000) characterized 126 putative and known *cre* sequences from the *B. subtilis* genome using a query sequence (WTGNAANCGNWNNCW, where N and W stand for any base and A or T, respectively) and analysed the sequences using *lacZ* fusions inserted into the *amyE* locus. There was a correlation between the partial palindromic nature of the *cre* sequence and the improved function of the *cre* element and for *cre* sequences located within coding regions of their target genes, the conserved bases were preferentially the third bases of codons where base degeneracy is allowed. However they also observed that low mismatching of the *cre* sequence in the same direction as that of the transcription of the target gene was more critical for their function than that in the inverse direction.

One of the key elements that interact with the *cre* element in CCR is the trans-acting protein CcpA, which will be discussed at in the following section. Computer analyses of the *B. subtilis* genome revealed that *cre* elements were found in most operons subject to CcpA-mediated glucose repression but not those subject to glucose activation suggesting that secondary transcription factors may regulate the latter genes (Moreno et al., 2001).

1.3.3.2 Catabolite Control protein (CcpA)

The existence of a *cre* sequence near the α-amylase promoter suggested the existence of a repressor molecule that would bind to the operator. Transposon mutagenesis was employed to identify genes potentially coding for such a repressor (Henkin et al., 1991). A promoterless *cat* gene was placed under the control of the α-amylase promoter and screened for trans-acting mutants capable of transcribing the amylase promoter in the presence of glucose and thus conferring CCR resistance. The mutations fell into four categories based on their locations on the *B. subtilis* genetic map. All mutants were similar in phenotype viz. derepression in presence of glucose, however there were mutants that remained sensitive to repression by glycerol suggesting that CCR of these two carbohydrate sources is exerted through a different pathway.

The above approach facilitated the isolation of a gene, *ccpA*, involved in the CCR of the α- amylase gene. A *B. subtilis* mutant WLN-29 (gra-26::Tn917*lac*) was isolated which was able to synthesize α- amylase in the presence of glucose. The α- amylase enzyme activity of *B. subtilis* strain 168 (wild-type), in the presence of rapidly metabolizable carbon sources was 15 to 30 % of the level obtained without the addition of excess sugars. In contrast, in cultures of the mutant strain WLN-29 (gra-26::Tn917*lac*), α-amylase activity was 47 to 75% of the unrepressed level even in the presence of most sugars except glycerol. The *ccpA* gene was cloned from wild-type *B. subtilis* and inserted into an SPβ specialized transducing phage vector for complementation analysis using the WLN-29 (gra-26::Tn917*lac*) mutant. Transductants had lost the Gra⁻ phenotype and α-amylase synthesis was repressed in the presence of glucose indicating that CCR was restored. The Gra⁻ phenotype was restored to WLN-29 (gra-26::Tn917*lac*) upon curing the SPβ prophage from the strain, confirming that the *ccpA* gene was involved in CCR.

The location of *alsA*, a gene involved in the regulation of acetolactate synthase activity was close to the map location of the *ccpA* gene. Therefore, acetoin production was examined in WLN-29 and an *alsA* mutant 1A147 (*alsA1*). Both the *ccpA* and *alsA*

mutations resulted in failure to produce acetoin in the presence of glucose, while acetoin production in the presence of glycerol was unaffected. Introduction of the SPβ prophage carrying the intact *ccpA* gene resulted in restoration of acetoin production in both the *ccpA* and *alsA* mutant strains, indicating that the *ccpA* and *alsA* genes were allelic. The *alsA1* mutation was due to a base transition in the ribosome-binding site preceding *ccpA*. Grundy et al. (1993b) also found that the stimulation of *ackA* expression by glucose was blocked in a CcpA$^-$ mutant indicating that CcpA induces *ackA* expression in contrast to α-amylase gene which is repressed in the presence of glucose.

Subsequently it was shown by many researchers that CCR in *B. subtilis* is deregulated in CcpA mutants (Hueck and Hillen, 1995). Most of the carbohydrate utilization pathways were subject to substrate-specific induction while global CCR was the superimposed regulatory mechanism. The enzymes were completely relieved of CCR in *ccpA* mutants, while only a few showed residual repression, which indicated that additional CCR mechanisms might be present for those enzymes.

The *B. subtilis* CcpA was homologous to sequences of the GalR-LacI family of transcriptional regulatory proteins (Nguyen and Saier, 1995). CcpA had 31% and 25% identity with GalR and LacI respectively. The region of greatest similarity between CcpA and this family resided in the amino-terminal half of the protein, containing the DNA binding domain of LacI. The amino-terminal portion of CcpA was predicted to form a α-helix-turn-helix structure characteristic of LacI and related repressor proteins. Only two residues in the recognition α-helices of CcpA, GalR and LacI were not identical (Hueck and Hillen, 1995). Comparison of number of CcpA-like sequences from gram-positive bacteria with regulators from the GalR-LacI family defined a CcpA subfamily based on extensive similarities found among the CcpAs (Kraus et al., 1998). The relief of CCR in *ccpA* mutants and the structural nature of the CcpA protein indicated that the *ccpA* product was a pleiotropic regulator of central carbon metabolism.

CcpA has also been shown to be involved in sporulation. Miwa et al. (1994) demonstrated that the *alsA1* mutation (discussed previously) affected sporulation. CCR of sporulation by CcpA protein further reinforced the idea that the regulator played an important role in coordinating global gene expression in response to rapidly metabolizable carbohydrates in the medium.

CcpA homologues have been isolated from numerous gram-positive bacteria (van Den Bogaard et al., 2000; Leboeuf et al., 2000; Mahr et al., 2000; Moon et al., 2000; Schick et al., 1999; Simpson and Russell, 1998; Lokman et al., 1997; Monedero et al., 1997; Egeter and Brückner, 1996; Hueck et al., 1995). As a part of a study on genes from *Clostridium acetobutylicum* NCP262 involved in the electron transfer system and that were able to activate the drug metronidazole in *E. coli*, Davison et al. (1995) isolated a catabolite repressor protein (RegA). Complementation studies of *regA* in *B. subtilis ccpA* mutants resulted in glucose repression indicating that *regA* had a similar role to *ccpA* in *C. acetobutylicum*. Furthermore, Küster et al. (1996) found immunological cross-reactivity of CcpA from *B. megaterium* to numerous gram-positive bacteria, which provided evidence that CCR in these bacteria may be exerted by mechanisms similar to the one found in *B. subtilis* and *B. megaterium*.

CcpA has been shown to bind to the *cre* elements with other co-factors involved in CCR (Fujita et al., 1995; Kim et al., 1995) and without co-factors (Kim et al., 1995; Ramseier et al., 1995). The nature of the interactions will be discussed in Section 1.3.3.4 below. Kim and Chambliss (1997) demonstrated that two monomers of the CcpA molecule contact the major groove of the *amyE cre* element, in each half of the region of dyad symmetry and on the same face of the DNA helix, which is typical of repressor-operator interactions. CcpA made contact with three phosphate groups at each end of the *cre* element and one or two phosphate groups near the dyad axis. The difference between the DNA contacts of CcpA and the GalR-LacI repressor family was the absence of strong CcpA contact near the dyad axis.

Several studies have examined the genes upstream and downstream of *ccpA* in certain Gram-positive bacteria. In *B. subtilis* and *B. megaterium*, two open reading frames (ORFs) are located downstream from and are probably co-transcribed with *ccpA* (Grundy et al., 1994; Hueck and Hillen, 1995). The deduced products of ORF1 and ORF2 are homologous to MotA and MotB membrane-associated proteins required for flagellar rotation in *B. subtilis* and *E. coli*. However null mutations in ORF1 and ORF2 resulted in no detectable phenotype with respect to CCR, growth rates on various carbon sources, motility, competence and sporulation (Grundy et al., 1993b). The role of ORF1 and ORF2 downstream from *ccpA* therefore remains unexplained. The sequential and structural conservation of these genes in *B. subtilis* and *B. megaterium* implies a selective pressure for maintaining an important function. Hueck and Hillen (1995) speculated that transmembrane sensing or membrane potential could play a role in CCR via the proteins encoded by ORF1 and ORF2. Mahr et al. (2000) compared the genetic structures of the *ccpA* regions of lactic acid bacteria, bacilli and *Staphylococcus xylosus* and found that the lactic acid bacteria conserved a [*pepQ*]-[*ccpA*]-[variable] gene order while the latter two conserved an [*aroA*]-[*ccpA*]-[variable]-[*acuC*] gene order. This result indicated that gram-positive bacteria might possess specific genetic structures for *ccpA* regions unique for the species.

In addition to loss of CCR, *B. subtilis ccpA* mutants exhibit a second phenotype, which is poor growth on minimal media with the growth defect less pronounced in complex medium (Miwa et al., 1994). Factors other than carbohydrate transport caused the growth defect since sugar-specific permeases do not depend on a functional CcpA and growth with different carbon sources was affected irrespective of the transport system involved. These *ccpA* mutants were auxotrophic for glutamate and required the addition of glucose, citrate or other tricarboxylic acid (TCA) cycle intermediates and ammonium (Wray et al., 1994). Deutscher et al. (1994) therefore suggested that high glycolytic activity is required to evoke CCR in *B. subtilis* and CcpA may exert its effect on a central pathway of carbon metabolism like glycolysis or the TCA cycle. However even

in the presence of citrate and glutamate, growth inhibition of the *ccpA* mutant by arabinose was observed (Faires et al., 1999). Growth inhibition was due to some unidentified intermediates of arabinose utilization since it was not observed in the presence of ribose and was probably not due to catabolism via the pentose phosphate shunt. A *spoOA* mutation results in *abrB* overexpression, which allows the growth defect of *ccpA* mutants to be circumvented. AbrB is a transition- state regulator capable of specifically binding to *cre* elements in vitro (Hueck and Hillen, 1995). An AbrB null mutation leads to more efficient CCR of some genes in *B. subtilis* (Strauch, 1995). Thus AbrB has been suggested to compete with CcpA (or other CCR-mediating proteins) for binding to *cre* sequences.

The CcpA protein may provide the link between carbon and nitrogen metabolism in Gram-positive bacteria via different mechanisms. Faires et al. (1999) presented data, which indicates that glucose transport was not impaired in the *B. subtilis ccpA* mutant strain. Isolated supressor mutants (*sgd*) that affected the growth behaviour but not CCR allowed one to distinguish between the two phenotypes of the *ccpA* mutant. They found the c*cpA* mutant growth deficiency to be similar to *glt* (glutamate) mutants affected in the central enzyme of ammonium assimilation, glutamate synthase. This phenotype was associated with the decrease of *gltAB* operon expression in the *ccpA* mutant strain. An *sgd* mutation, which restored growth of the *ccpA* mutant strain, concomitantly restored the expression of the *gltAB* operon. Therefore glycolytic activity was not the factor causing the *ccpA* mutant phenotype and the growth deficiency did not depend on the pathway used for carbon source catabolization. Analysis of the growth requirements of the *ccpA* mutant revealed that the mutant was auxotrophic for glutamate and had the same phenotype as *glt* mutants. They demonstrated that *gltAB* operon expression is induced by glucose and this induction depends on a functional *ccpA* gene though no *cre* could be identified in the upstream regions. The *sgd-1* mutation renders the expression of *gltAB* independent from the presence of glucose in the medium and from the CcpA protein. Using a *ccpA sgd-1* double mutant they found that inducible

expression of the *gltAB* operon did not restore the *ccpA* mutant growth defect. This suggests that there may be factor(s) positively controlled by CcpA, which are required for ammonium assimilation. Powell et al. (1995) showed how the PTS components connected carbon and nitrogen metabolism in *E. coli* and other gram-negative bacteria. The genes specified by the *sgd* mutations are possible candidates in the provision of this link in addition to CcpA.

Previous studies demonstrated the requirement for CcpA in the activation of fermentative pathways and the *ackA* gene (Grundy et al., 1993b). Nakano et al. (1997) showed that acetate and butanediol are fermentation products of *B. subtilis* when grown anaerobically in the presence of glucose and pyruvate. These findings indicate that CcpA may also be involved in anaerobic fermentative growth and thus CcpA is a central regulator of carbon, energy and nitrogen metabolism in *B. subtilis* and possibly other gram-positive bacteria (Tobisch et al., 1999).

A protein, CcpB with 30% identity to CcpA has been implicated in CCR of the *B.subtlis xyl* and *gnt* operons in parallel with CcpA when grown on solid media or in media with little agitation (Chauvaux et al., 1998). CcpB appeared to exert CCR by a mechanism dependent on HPr-Ser-P (discussed in 1.3.3.3.) and environmental conditions affected mediation by CcpB. Kraus et al. (1998) found that the CcpB sequence did not contain CcpA-specific residues and was not therefore a member of the CcpA sub-family.

Whole-genome analysis of *B. subtilis* found that 10% of all genes are regulated by glucose with repressed genes outnumbering activated genes and 80% of these genes depended on CcpA for regulation (Moreno et al., 2001). Furthermore they provided evidence for CcpA mediating glucose-independent activation or repression and that glucose might alter the direction or the intensity of either effect. The operons that were subject to CcpA-mediated CCR followed two distinct patterns where all the genes in the

operon were regulated in parallel (minor class) or where the gene encoding the solute binding receptor was preferentially regulated (major class). Several transcription factors were identified that regulated CcpA at the transcriptional level and genes subject to CcpA-independent CCR were identified and found to be primarily concerned with sporulation.

1.3.3.3 HPr

A primary co-factor of CcpA modulation is the protein HPr which is also a key PTS component. The HPr of several gram-positive bacteria can be phosphorylated at Ser46 by an ATP-dependent HPr kinase (Stülke and Hillen, 2000, Reizer et al., 1998). It is activated in the presence of fructose-1,6- bisphosphate and to a lesser extent, by other glycolytic intermediates and is inhibited by inorganic phosphate. An inorganic phosphate-activated phosphatase isolated from *Enterococcus faecalis* was shown to catalyze dephosphorylation of HPr-Ser-P from various organisms (Hueck and Hillen, 1995). The HPr-Ser- P form is present in cells with high glycolytic activity whereas HPr-Ser-P is dephosphorylated under starvation conditions. HPr-Serine46 phosphorylation therefore represents a switch responding to carbon-source availability and to the energy level of the cell. The HPr-Ser-46-Ala mutant where the Ser-46 is replaced by alanine, cannot be phosphorylated at this position and *B. subtilis* strains carrying this mutation are relieved of CCR of certain catabolic activities and of mannitol-PTS activity (Deutscher et al., 1994). This deregulation is caused neither by inducer exclusion nor by reduced uptake of sugars mediating CCR, and links CCR in *B. subtilis* to the PTS via HPr. These effects on sugar metabolism are mediated by direct interactions between HPr-Ser-P and the respective enzymes.

In order to ascertain whether protein-protein interactions exist between HPr-Ser46-P and CcpA, Deutscher et al. (1995) overexpressed and purified CcpA from *B. megaterium* with a His-tag fusion which was functional *in vivo*. The CcpA protein,

specifically retarded HPr-Ser46-P in the absence of DNA out of a mixture of unphosphorylated HPr, HPr-Ser-P, HPr-His-P and P-Ser-HPr-His-P. Addition of fructose 1,6 bisphosphate increased retardation of HPr-Ser46-P. Fujita et al. (1995) found that CcpA bound to the *B. subtilis gnt cre* sequence in an HPr-Ser46-P dependent manner *in vitro*. Ramseier et al. (1995) demonstrated that the interaction of CcpA with the *xyl cre* sequence was stimulated non-specifically by the presence of phosphate-containing compounds and negatively regulated by HPr-Ser-P. Furthermore, Jones et al. (1997) demonstrated that CcpA interacts with HPr-Ser46-P and these proteins formed a ternary complex with *cre* DNA with the binding site of CcpA encompassing both phosphorylation sites of HPr, allowing CcpA to recognize the phosphorylation state of HPr. Gösseringer et al. (1997) demonstrated that CcpA could bind to *cre* sites in the presence of glucose-6-phosphate. Studies by Kim et al. (1998) implicated NADP as a co-repressor of CcpA, where the NADP effect was primarily on stimulating the inhibition of transcription rather than on DNA binding by CcpA and suggested that CcpA interacted with the RNA polymerase. The effect of NADP was enhanced when combined with HPr-Ser46-P and synergistically stimulated DNA-binding affinity. The *B. subtilis crh* gene expressed a novel HPr-like protein also involved in CCR in which the Ser46 residue but not the His15 residue is available for phosphorylation (Cases and Lorenzo, 1998; Galinier et al., 1997). Crh is phosphorylated by HPr kinase in the presence of elevated concentrations of fructose 1,6-bisphosphate and it was shown that Crh-Ser-P stimulated binding of CcpA to the *xyn cre* element but had no effect on CcpA DNA binding indicating that it was another co-represser for CcpA (Galinier et al., 1999). It has been suggested that because Crh has only one phosphorylation site it might play a role in CCR through an ATP-dependent phosphorylation pathway separate from the PTS system (Cases and Lorenzo, 1998). Furthermore sequence conservation and negative trans-dominance studies on members of the CcpA subfamily indicated that other ligands might bind and exert their effects on CcpA (Kraus et al. 1998). All these results indicate that modulation of CcpA activity can be affected by different co-factors.

A different mechanism of catabolite regulation occurs in the absence of repressing sugars. Under these conditions HPr phosphorylates positively acting transcriptional regulators containing a PTS regulation domain (PRD) thereby stimulating their activity (Stülke et al., 1998). Specific regulators like LevR, LicT and LacT mediate the CCR of *B. subtilis* levanase, β-glucoside and *L. casei* β-galactosidase genes in *ccpA* mutant strains (Stülke and Hillen, 1999). These regulators whose activities are controlled by HPr- or EIIB- dependent phosphorylation contain the conserved PRD domain and act as transcriptional activators or antiterminators resulting in either induction or catabolite repression.

In the following section the interaction between the various CCR components will be discussed and a model proposed for gram-positive CCR.

1.3.3.4 Interaction of *cre*, CcpA and HPr

The components, which are central to CCR in *B. subtilis*, are the *cre* elements, CcpA, HPr (PTS) and Crh. CCR is affected by mutations in the *cre*, CcpA and HPr-Ser-P. These components are thus essential members of a pathway which signals the presence of a readily metabolizable carbon source to CCR of catabolic genes.

HPr from low-GC gram-positive bacteria can be phosphorylated/dephosphorylated at the His15 residue or the Ser46 residue, the former by PEP and enzyme I, and the latter by an ATP-dependent, metabolite-activated HPr kinase/phosphatase system (Martin-Verstraete et al., 1999, Jankovic et al., 2000). In order to regulate transcription of the controlled genes, CcpA needs to interact with a cofactor, which may be either glucose-6-phosphate, NADP, HPr-Ser46-P or Crh-Ser46-P. The generation of the cofactors and their subsequent interaction with CcpA link the metabolic status of the cell to the activity of the regulator. The two phosphorylation sites available on HPr allow many proteins, including CcpA, to recognize the phosphorylation status of HPr and thereby

affect the utilization of carbon sources at the level of their transport as well at the level of transcription of genes involved in their metabolism (Cases and Lorenzo, 1998). This also allows HPr to act as the mediator of inducer expulsion and inducer exclusion in gram-positive bacteria. The function of CcpA in catabolite repression in *B. subtilis* is therefore equivalent to that of the cAMP receptor protein Crp in *E. coli* and other enteric bacteria. A model of CCR in low-GC gram-positive bacteria is given in Figure 1.4.

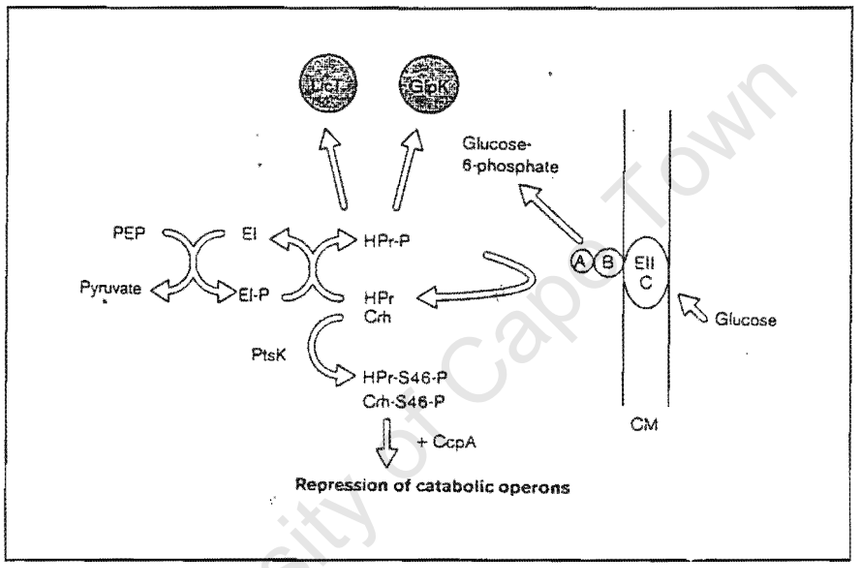

Figure1.4: CCR in bacilli and other low-GC gram-positive bacteria. The core of the PTS involves the flow of phosphate from PEP to the exogenous carbohydrate simultaneously with its transport. This flow is mediated by EI, HPr and carbohydrate-specific EII, which includes different subunits (A,B and C). The phophorylation state of HPr controls the expression of catabolic operons in response to repressing carbon sources. In the presence of glucose, HPr is phophorylated at Ser46 by HPr kinase (PtsK) and, along with phosphorylated Crh, acts as a co-repressor for CcpA. In the absence of glucose, HPr phosphorylated by Enzyme I can stimulate the activities of glycerol kinasė (GlpK) and operon-specific regulators such as LicT. CM, cell membrane. Adapted from Stülke and Hillen, 1999).

CcpA is able to sense glycolytic activity by at least two different mechanisms. The level of fructose 1,6-bisphosphate stimulates regulatory phosphorylation of HPr at Ser46 and subsequent binding to CcpA, and the other mechanism is when glucose-6-phosphate or NADP directly stimulates CcpA. In both mechanisms, the stimulated CcpA binds to

the *cre* elements. The absence of a readily metabolizable carbon source would lead to inorganic phosphate-stimulated HPr-Ser46-P-phosphatase activity resulting in dephosphorylation and dissociation of the complex with CcpA and ultimately relief from CCR.

Gösseringer et al. (1997) found that Hpr-Ser46-P and glucose-6-phosphate triggered different binding mechanisms of CcpA to *cre* elements in the *xyl* operon. The interaction triggered by glucose-6-phosphate was cooperative and involved three *cre* sequences in the *xyl* operon, probably brought in close proximity by DNA looping. The HPr-Ser46-P triggered interaction was non-cooperative and confined to a single *cre*. The two triggers were not simultaneously active and the pH of the buffer determined which of them activated CcpA. For pH values below 5.4, glucose-6-phosphate activation of *cre* binding was preferred and pH values above 5.4 favoured HPr-Ser46-P activation of *cre* binding. It appears that cooperative binding of CcpA to *cre* is confined to acidic conditions. Cooperativity of DNA binding requires oligomerization of the CcpA dimers, which are present at neutral pH. CcpA dimers form greater molecular mass aggregates at lower pH. The only member of the GalR-LacI family of bacterial regulators shown experimentally to form tetramers is the LacI repressor, and this is a property of the C-terminal region. CcpA from *B. subtilis* or *B. megaterium* does not contain such a tetramerization domain. Therefore it was the first member of this protein family for which aggregation to higher oligomers was demonstrated. CcpA has the capability for this binding mode, but it is not clear whether pH values affect binding *in vivo*. *Bacillus* species, maintain an intracellular pH of between 7.5 and 8.0. *B. subtilis* growing on glucose shows a temporary but drastic decrease in external pH down to values around pH 5.0 caused by the production and excretion of organic acids. Decreases of internal and external pH values were also reported for several gram-positive bacteria growing under anaerobic conditions. Gösseringer et al. (1997) speculated therefore that the low internal pH values might occur transiently in bacillus species. A proposed model of CcpA function in the CCR of *B. subtilis* operons containing *cre* elements is given below in Figure 1.5.

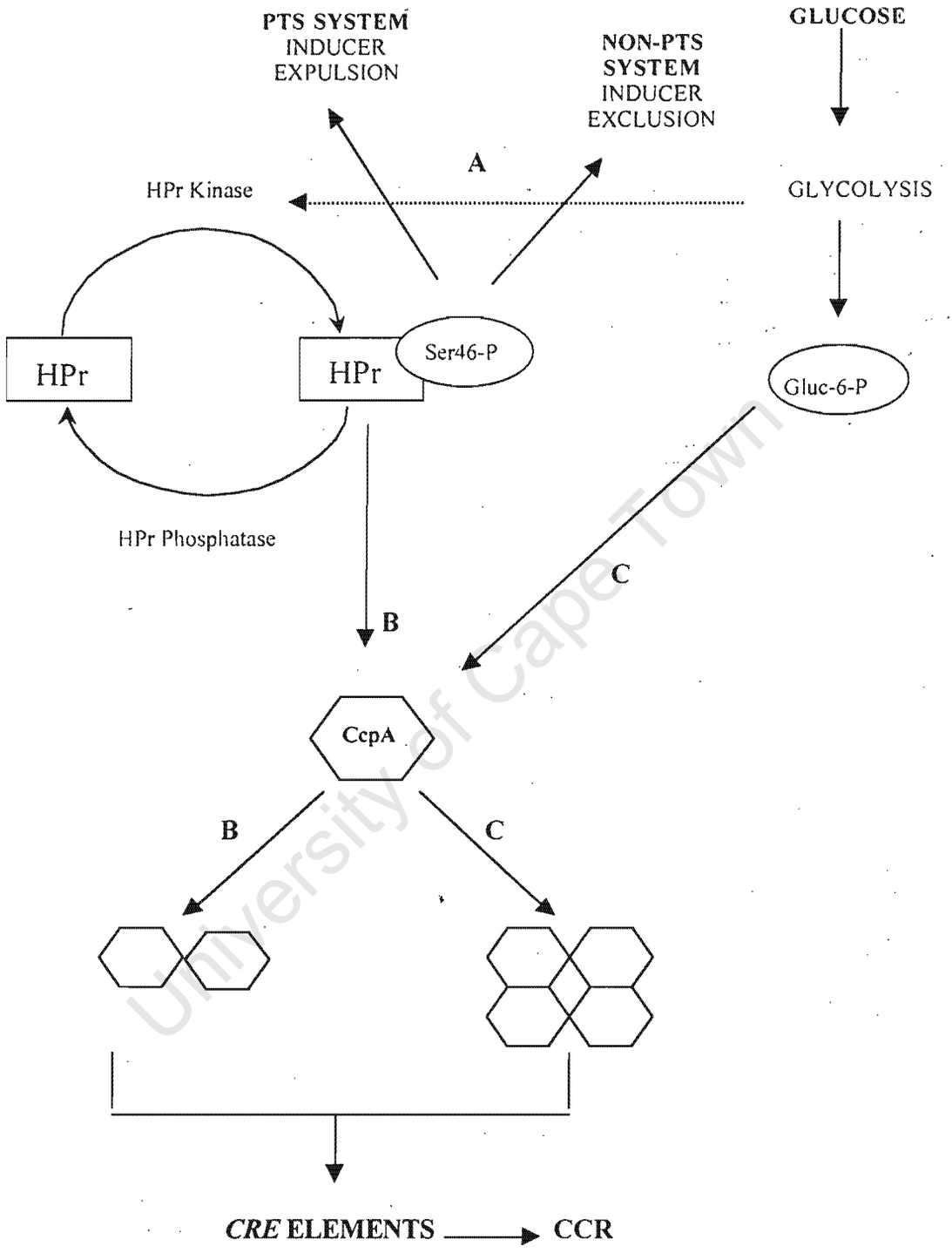

Figure 1.5: A proposed model of CcpA function in the catabolite repression mechanism in gram-positive bacteria based on the *xyl* operon of *B. megaterium* (Gösseringer et al., 1997). Glycolysis results in the accumulation of high concentrations of glycolytic intermediates such as glucose-6-phosphate and these intermediates stimulate/activate HPr kinase resulting in the phosphorylation of the Ser46 residue of

catabolism of sucrose intracellularly is the PTS system, whereby the sucrose is transported and phophorylated by an EII^{Scr} and results in the accumulation of intracellular sucrose-6-phosphate which is then hydrolysed further by sucrose-6-phosphate hydrolase to glucose-6-phosphate and fructose (Saier and Reizer, 1992). Tangney et al. (1998) demonstrated that in *C. beijerinckii* sucrose is also transported into the cell via a sucrose-specific PTS system and the process was subject to repression by glucose. The products of sucrose-6-phosphate hydrolysis were glucose-6-phosphate and fructose, which was subsequently phosphorylated by a fructokinase. Tangney and Mitchell (2000) detected a similar pathway to that in *C. beijerinckii* for sucrose metabolism in *C. acetobutylicum* ATCC824.

The components of the *C. beijerinckii* sucrose-specific PTS system were identified and characterized at the genetic level (Reid et al., 1999). The *C. beijerinckii* sucrose operon comprised four genes in the respective order: *scrA* encoding a sucrose-specific EII^{Scr} protein of the PTS; *scrR* encoding a transcriptional repressor of the GalR-LacI family; *scrB* encoding a sucrose hydolase and *scrK* encoding an ATP-dependent fructokinase. The *scr* genes were co-transcribed as a 5 kb mRNA transcript and constituted the *C. beijerinckii scrARBK* operon. The *scrARBK* genes were transcribed when *C. beijerinckii* was grown in sucrose and repressed in the presence of certain carbohydrates. The gene order in *C. acetobutylicum* ATCC824 was *scrTAKB* with *scrAKB* constituting an operon as all three genes were expressed as a single transcript (Tangney and Mitchell, 2000).

The *C. beijerinckii* ScrA protein had the highest identity to the EIIBC domains of the ScrA from *Streptococcus sobrinus* and conserved the proposed EIIB phosphorylation site and the GITE motif, which has been proposed to be involved in phosphate and/or substrate binding (Lengeler et al., 1994). The *C. beijerinckii* $EIIBC^{Scr}$ protein must interact with an EIIA domain, which could either be, a specific $EIIA^{Scr}$ domain or the $EIIA^{Glc}$ domain of the $EIICBA^{Glc}$ protein. The $EIIA^{Glc}$ domain in *B. subtilis* was shown to phosphorylate $EIIB^{Scr}$ (Sutrina et al., 1990). The *E. coli* $EIIA^{Glc}$ domain was also

required for EIIBC^{Scr} dependent sucrose uptake when the sucrose systems of *Klebsiella pneumoniae*, *Salmonella typhimurium* and *Vibrio alginolyticus* were cloned in *E. coli* respectively (Blatch et al., 1990; Lengeler et al., 1982; Sprenger and Lengeler, 1988). Tangney et al. (1998) demonstrated that there is no specific EIIA^{Scr} component in the cytosol of *C. beijerinckii* making it very likely that the IIA^{Glc} component fulfills this function in sucrose grown cells. In contrast the sucrose operon of *C. acetobutylicum* ATCC824 contained an EIIBCA protein in which all three functional domains are fused and the EIIA domain would allow sucrose transport via the PTS in this organism (Tangney and Mitchell, 2000).

The *C. beijerinckii* ScrR protein had the highest sequence identity to the *Staphylococcus xylosus* ScrR protein with the highest conservation being the N-terminal helix-turn-helix DNA-binding motif and regions identified with being involved in inducer binding (Gering and Brückner, 1996). Phylogenetic analyses confirmed that the ScrR proteins were more closely related to each other than to other members of the GalR-LacI family. A *C. beijerinckii scrR* mutant resulted in constitutive expression of the *scrARBK* operon under non-inducing conditions confirming that ScrR played a negative regulatory role (Reid et al., 1999). It was therefore proposed that ScrR was a transcriptional repressor, which most likely acts at the *scrA* operator sequence, and therefore the *scrR* gene was itself negatively autoregulated. In the absence of an *scrR* gene, the sucrose operon of *C. acetobutylicum* ATCC824 had an *scrT* gene and two RAT sequence terminators, which was evidence for an antiterminator-mediated regulatory mechanism, typical of the BglG family of regulators (Tangney and Mitchell, 2000). The BglG-Sac antitermination system has been well reviewed previously (Rutberg, 1997).

The *C. beijerinckii* ScrB protein showed homology to _-fructofuranosidases of the glucosyl hydrolase family 32 with highest identity to the sucrose-6-phosphate hydrolases, which confirms biochemical evidence that sucrose-6-phosphate is hydrolysed (Tangney et al., 1998). The *C. beijerinckii* ScrK protein showed homology

to the ribokinase/*pfkB* family fructokinases with conservation of regions possibly involved in ATP and fructose binding (Reid et al., 1999). Furthermore this was the first reported instance of a complete ribokinase/*pfkB* family fructokinase sequence associated with a sucrose utilization system in a gram-positive organism.

A *cre* element was identified downstream of the –10 region of the *C. beijerinckii* *scrARBK* operon, overlapping the transcriptional initiation site (Reid et al., 1999). This *cre* element differed by only 1 bp from the Weickert and Chambliss (1990) consensus *cre* sequence. A *cre* element, with one mismatch to the consensus *cre* sequence, was also identified overlapping the –10 region of the *C. acetobutylicum* ATCC824 *scrTAKB* operon (Tangney and Mitchell, 2000). This *cre* sequence had a TG palindromic center in contrast to the more common CG center found in *C. beijerinckii* and other *cre* elements identified in most other gram-positive bacteria (Zalieckas et al., 1998b). The *cre* sites are known to act as target sites for CcpA-Hpr complexes in *Bacillus* species (discussed in previous sections) and the identification of these sites in the sucrose operons of both clostridial species suggests that these operons are subject to CCR by mechanisms similar to those occurring in *Bacillus*. Furthermore ATP-dependent phosphorylation of HPr in cell-free extracts of *C. beijerinckii* has been demonstrated and a CcpA homolog, RegA, has been identified in *C. acetobutylicum* NCP262 indicating that all the components for CCR exist in clostridial species (Tangey and Mitchell, 2000; Davison et al., 1995).

1.5 AIM OF THIS STUDY

Catabolite repression of the sucrose operon has been observed in mixed sucrose/glucose batch cultures in which sucrose hydrolase and fructokinase activities were partially repressed but the mechanism facilitating the CCR has not been elucidated (Leat, 1997). Furthermore, the author identified a putative *cre* element in the operator region of the *scrARBK* operon indicating a possible mechanism involving a *C. beijerinckii* CcpA

homologue. The aim of this study was therefore to isolate and characterize a CcpA homologue from *C. beijerinckii* NCIMB 8052 by sequence analysis and gene inactivation and to investigate a physiological link between the CcpA homologue and CCR regulation of the sucrose operon.

CHAPTER TWO

CLONING AND SEQUENCING OF *REGB*, THE CCPA HOMOLOGUE IN *C. BEIJERINCKII* NCIMB 8052

CONTENTS

CHAPTER TWO

2.1 SUMMARY

The *regB* gene was isolated from a chromosomal gene library of *C. beijerinckii* NCIMB 8052 by hybridization using the *C. acetobutylicum regA* gene (Davison et al., 1995) as a probe. The gene was truncated and had to be reconstructed using a construct containing the C-terminal portion of the gene which had been identified by colony hybridisation. The gene was 996 bp in length and encoded a protein of 332 amino acids with a calculated MW of 39 840 Da. Sequence; phylogenetic and structural analyses confirmed that the RegB protein belonged to the CcpA subfamily within the GalR-LacI family of transcriptional regulators. The RegB protein retains conserved structural and functional amino acid residues of the CcpA subfamily, which indicates conformational and functional identity with CcpA. This suggests that RegB could act as a global regulator of CCR in *C. beijerinckii*.

A gene encoding a protein with significant identity to isoleucyl-tRNA synthetase proteins was found upstream of *regB*. The upstream gene encoded a protein with 30.06% identity and 51.45% similarity to the isoleucyl-tRNA synthetase of the Lyme disease spirochaete, *Borrelia burgdorferi*.

The gene immediately downstream of *regB* was 441 bp in length and encoded a protein of 147 amino acids with a calculated Mw of 17 640 Da. The encoded protein showed identity with transcriptional regulators of *L. lactis*, *B. subtilis*, *B. halodurans*, *Staphylococcus sciuri*, *Vibrio cholerae* and *E. coli*. Most of these transcriptional regulators belonged to the MarR family of transcriptional regulators. The functions regulated by the MarR family members include regulation of antibiotic resistance (Alekshun and Levy, 1997; Sulavik et al., 1997), negative regulation of sporulation and protease production (Koide et al., 1999) and anaerobic aromatic compound degradation (Egland and Harwood, 1999).

Downstream of the putative MarR family transcriptional regulator was a truncated gene, encoding a putative protein with 30% amino acid identity to an unknown conserved

protein (BAB05882) in *B. halodurans*. These proteins belonged to an uncharacterized membrane protein family, UPF0013 whose function is currently unknown.

2.2 INTRODUCTION

As a part of a study on the molecular characterization of *C. acetobutylicum* NCP262 genes involved in the electron transfer system and metronidazole sensitivity, Davison et al. (1995) cloned the gene encoding the catabolite repressor protein, RegA.

This *regA* gene (972 bp) encoded a protein of 324 amino acids (35600 Da). The ATG start codon was preceded 8 bp upstream by a Shine-Dalgarno sequence that resembled those reported for the *S. aureus* β-lactamase and *C. acetobutylicum* endoglucanase genes. The putative promoter showed homology to gram-positive promoters regulating the *C. acetobutylicum glnA* gene. Two stem-loop structures were identified downstream of the *regA* gene with eight thymine residues present after the first stem-loop structure. This structure is similar to factor-independent terminators. The dyad symmetry of stem-loop structures allows for the formation of a stable hairpin structure in the RNA transcript thereby slowing down the polymerase, whereas the thymine rich region allows for the formation of a rU-dA hybrid that facilitates release of the transcript.

Sequence comparison at the amino acid level indicated that the protein encoded by this ORF was most likely a DNA-binding protein. The deduced RegA protein had 40% and 38% amino acid identity to the *B. subtilis* and *B. megaterium* CcpA proteins respectively (Davison et al., 1995). The amino terminal region of RegA contained a helix-turn-helix structure which was similar to DNA-binding domains of other repressor proteins.

Since the *C. acetobutylicum regA* gene product showed the highest amino acid identity to the *B. subtilis ccpA* gene product, the ability of the *regA* gene to complement a *B. subtilis ccpA* mutant was examined. The *C. acetobutylicum regA* gene, contained on a 6.05 kb DNA fragment was inserted into SPβ phage vector, transduced into the *B. subtilis ccpA* mutant, WLN-26 and screened for α-amylase repression. Using a

qualitative plate assay, glucose repression was observed in the *B. subtilis ccpA* mutant complemented with *regA*. The *B. subtilis ccpA* gene also controls the production of acetoin in the presence of glucose by regulating the expression of the *alsA* gene responsible for the synthesis of acetolactate synthase (Grundy et al., 1993b). The *B. subtilis ccpA* mutant's ability to produce acetoin in a glucose medium was also restored upon complementation with *regA*.

Based on the above study, RegA may serve the same global regulatory role in *C. acetobutylicum* NCP262 as CcpA and its homologues in other gram-positive bacteria (Stülke and Hillen, 1999). However, genetic manipulation of *C. acetobutylicum* NCP262 is technically difficult with few of its CCR regulated genes analyzed at the molecular or physiological level.

C. beijerinckii however is amenable to genetic manipulation and operons that are hypothesized to be subject to CCR, such as the sucrose operon, have been cloned, sequenced and analysed (Reid et al, 1999; Tangney et al., 1998). This species therefore provides an ideal model for the study of CCR in saccharolytic clostridia. We therefore sought to find the homologue of RegA/CcpA in *C. beijerinckii* in order to investigate its possible effects on CCR. This chapter describes the steps taken to isolate a *regA* homologue from *C. beijerinckii* and also presents its subsequent sequence analysis.

2.3 MATERIALS AND METHODS

2.3.1 Bacterial strains, plasmids and culture conditions

All bacterial strains and plasmids are listed in Table 2.1

Cultures of *E. coli* were routinely maintained on Yeast Tryptone (YT) or Luria-Bertani (LB) medium (Sambrook *et al.*, 1989) at 37°C. YT and LB media were solidified with agar (1.5% w/v). Ampicillin (100ug/ml) was included in the media where appropriate.

The laboratory stock of *C. beijerinckii* NCIMB 8052 and *C. acetobutylicum* NCP262 was maintained in distilled water at 4°C. Both clostridial species were routinely cultured using Clostridial Basal Medium (CBM) (O'Brien and Morris, 1971). CBM liquid medium was cooled to approximately 50°C after autoclaving and transferred to an

anaerobic cabinet with a gas phase of 5% H_2, 10% CO_2 and 85% N_2 (Forma Scientifica Inc., Marietta, Ohio, USA). CBM agar plates, solidified with agar (1.5% w/v), were prepared under aerobic conditions and then pre-incubated under anaerobic conditions for at least 12 h before use. CBM broths were typically inoculated with 1/1000 volume of heat- shocked spores (70°C for 3 min) and cultured at 37°C.

Table 2.1: Bacterial strains and plasmids used in this study.

Name	Relevant Characteristics	Reference
Bacterial strains		
C. acetobutylicum NCP262	Wild-type	Jones et al. (1982)
C. beijerinckii NCIMB 8052	Wild-type	NCIMB
E. coli K12	Wild-type	Laboratoty stock
E. coli HB101	*thi-1, hsdS20(r_B-,m_B-), supE44, recA13*	Maniatis et al. (1982)
E. coli JM105	*Δ(lac-proAB) lac^q _(lacZ)M15*	Yanisch-Perron et al. (1985)
E. coli JM109	*Δ(lac-proAB)lac^q _(lacZ)M15, recA1*	Yanisch-Perron et al. (1985)
Plasmids		
pBluescriptSK$^+$	Ap^r, T3 and T7 polymerase promoters	Stratagene, La Jolla, California
pSK069	1.25 kb *ClaI/XbaI* DNA fragment containing *regA* cloned into pBluescriptSK$^+$	Davison S. (personal communication)
pECOB1	~8.5 kb *C.beijerinckii* DNA fragment in pECOR251	*C.beijerinckii* gene bank (Reid S., personal communication)
pSKB2	~2.76 kb *ClaI* DNA fragment containing a 5' truncated *regB* gene in pBluescriptSK$^+$	This study
pECOB3	~4.2 kb *EcoRI* derivative of construct pECOB1	This study
pSKB4	~4.0 kb *EcoRV* derivative of pECOB1 subcloned into pBluescriptSK$^+$	This study
PECORegB	~5.5 kb DNA fragment containing *regB* in pECOR251	This study
pSKA6	~1 kb *EcoRV/BamHI* DNA fragment from pECOB1 cloned into pBluescriptSK$^+$	This study

2.3.2 General DNA manipulation and extraction

Competent *E. coli* cells were prepared by rubidium chloride treatment as described by Armitage *et al.* (1988).

Large-scale extraction and purification of plasmid DNA from *E. coli* was carried out using the Nucleobond AX KIT (Macherey-Nagel, Germany) according to the manufacturers instructions. Small-scale extraction of plasmid DNA from *E. coli* was carried out by the alkali-hydrolysis method of either Ish-Horowicz and Burke (1981) or Zhou et al. (1990). Chromosomal DNA was extracted from *E. coli* according to the method of Wilson (1987).

DNA restriction endonucleases and T4 DNA ligase were obtained from Roche Diagnostics or Amersham and were used according to the manufacturers recommendations. Where DNA fragments with incompatible ends were to be ligated, protruding 5' and 3' ends were first converted to blunt ends, using Klenow fragment and T4 DNA polymerase as described by Sambrook et al. (1989). Both Klenow fragment and T4 DNA polymerase were obtained from Roche Diagnostics.

Electrophoresis of DNA was conducted in agarose gels (0.8%, w/v), using a Tris-acetate EDTA buffer as described by Sambrook et al. (1989).

2.3.3 *C. beijerinckii* and *C. acetobutylicum* genomic DNA extraction

One litre cultures of *C. beijerinckii* and *C. acetobutylicum* NCP262 were grown in CBM to an OD_{600} of 0.4 units and cells were harvested by centrifugation (7 min at 10 000 xg). Cells were resuspended in a 20ml volume of CBM containing: 10% (w/v) sucrose; 12.5mM $MgCl_2$; 12.5mM $CaCl_2$ and 5mg/ml lysozyme (Roche Diagnostics). After anaerobic incubation at 37°C for 1 h, SDS and EDTA were added to final concentrations of 2%(w/v) and 100mM respectively, bringing the total volume of the mixture to 27ml. A 20ml volume of hot phenol (50°C), equilibrated with 0.1M Tris-HCl (pH 8), was added and mixed gently. The aqueous phase was recovered by centrifugation (10 min at 12000 rpm) and extracted, firstly with one volume of chloroform: iso-amyl alcohol (24:1 v/v) and then with one volume of water-saturated

ether. RNA was removed using Ribonuclease A (Sigma Chemical Company, St. Louis, Missouri). Chromosomal DNA was precipitated using isopropanol (Sambrook et al., 1989). The DNA pellet was resuspended in distilled water and stored at 4°C.

2.3.4 Southern and Colony blots

Non- radioactive DIG-labelled probes were prepared using the Roche Diagnostics DIG random primer labelling kit (catalog number: 1175033). A 1.25 kb *Cla*I/*Xba*I fragment which contained part of the *regA* gene and a 380 bp *EcoR*V/*Xmn*I internal *regB* gene fragment were used appropriately as probes and labeled according to the manufacturers instructions. The 1.25 kb *Cla*I/*Xba*I *regA* probe was obtained after digestion of plasmid pSK069 and the 380 bp *EcoR*V/*Xmn*I *regB* probe was obtained after digestion of plasmid pSKA6.

For Southern blot analysis, DNA fragments were separated by agarose gel electrophoresis and capillary blotted onto positively charged Hybond N^+ nylon membranes according to the method of Reed and Mann (1985). Hybridization and detection of DIG-labelled probes was conducted according to standard protocols supplied by Roche Diagnostics.

For colony blots, colonies harbouring recombinant plasmids with *C. beijerinckii* insert DNA were duplicated onto Hybond N^+ nylon membranes placed on 2X YT solid medium. After overnight incubation the membranes were processed, hybridized and colonies detected according to standard protocols supplied by Roche Diagnostics.

2.3.5 Chromosomal walking and construction of size-selected gene bank

For chromosomal walking, selected restriction endonucleases were used to digest *C. beijerinckii* chromosomal DNA, which was probed with an internal *EcoR*V/*Xmn*I 380 bp *regB* probe after electrophoresis through an agarose gel. Restriction endonucleases that generated DNA fragments of suitable size, were selected for the creation of size-selected gene banks.

Size selected genebanks of *C. beijerinckii* were generated by digestion of *C. beijerinckii* chromosomal DNA (40ug) to completion with *Cla*I, and the resulting fragments were fractionated by electrophoresis in 0.8% agarose gel in Tris/acetate buffer. DNA fragments of the relevant size were recovered by the Geneclean (Q-biogene) procedure, ligated into pBluescriptSK$^+$ vectors, which were also restricted with *Cla*I, and then transformed into *E. coli* JM105.

2.3.6 Exonuclease III digestion, nucleotide sequencing and sequence analysis

Overlapping nested deletions were generated in DNA to be sequenced by Exonuclease III digestion (Henikoff, 1984). Double-stranded templates were sequenced by the dideoxy chain termination method of Sanger et al. (1977). Initially, sequencing was conducted using the Sequenase kit version (2.0) (U.S. Biochemical Corp., Cleveland, Ohio) and [^{35}S]-dATP. Subsequently, sequencing was conducted using the Sequitherm^{TM} kit (Epicenter Technologies, Madison, USA) using CY-5^{TM} labeled M13 primers and the sequence resolved using the ALFexpress automated DNA sequencer (Pharmacia).

Sequence analysis was performed using programs in the DNAMAN V4.0 (Lynnon Biosoft, Canada) and Genetics Computer Group (Devereux et al., 1984) sequence analysis packages. Amino acid identities and similarities between protein sequences were determined using DNAMAN V4.0 (Lynnon Biosoft, Canada). Multiple sequence alignments were conducted using the PILEUP component of GCG, ClustalW program (Thompson et al., 1994) and Genedoc V2.6 (www.psc.edu/biomed/genedoc). Protein hydrophobicity analysis was conducted according to the algorithm of Kyte and Doolittle (1982), as implemented in the DNAMAN sequence analysis package (Lynnon Biosoft, Canada). Phylogenetic analysis was conducted using the ClustalW program (Thompson et al., 1994), and the results were rendered using the TreeView software package (Page, 1996). The minimum free energy of RNA stem loop structures was determined using the FOLD component of GCG (Zuker and Stiegler, 1981).

Analyis of the unfinished genome of *C. acetobutylicum* ATCC824 was achieved using internet resources at www.cric.com/programs/sequence_data_clost.shtml and www.ncbi.nlm.nih.gov/Microb_blast/unfinishedgenome.html.

2.3.7 Preparation of total RNA from *C. beijerinckii*

C. beijerinckii cultures were harvested by centrifugation (5000 rpm for 5 min, 4°C) in 30ml culture volumes when the OD_{600} was between, 0.3 to 0.5. Cell pellets were then resuspended in 500µl aliquots of a 0.3M sucrose/0.01M sodium acetate solution (pH 4.5). In each case, a 500µl aliquot of a 2% SDS/0.01M sodium acetate solution (pH 4.5) was added and the mixture was maintained at 65°C for 1.5 min. Water-saturated phenol (1ml per sample) was added and the samples were mixed and maintained at 65°C for a further 3 min. The samples were then cooled at –70°C for 1 min. Aqueous phase recovery was achieved by centrifugation (10000 rpm for 5 min in an Eppendorf® 5415C microfuge). Two further hot phenol extractions were conducted. The RNA and DNA were precipitated by the addition of 1/10 vol of 5M sodium acetate, 3 vol of ethanol, and left at –70°C for at least 5 min or overnight. After centrifugation (10000 rpm for 10 min at 4°C) the pellet was resuspended in a 200µl vol of DNase buffer (20mM sodium acetate/ 10mM $MgCl_2$/10mM NaCl, pH 4.5) solution. The DNA was removed by the addition of DNase I (Roche Diagnostics) (30U per 200µl sample), and left at room temperature for 30 min. The remaining RNA was ethanol precipitated, resuspended in 50µl distilled water and stored at –70°C.

2.3.8 Primer Extension

RNA was extracted from *C. beijerinckii* grown in CBM with either glucose (1%) or sucrose (1%) serving as carbon sources. The RNA (100µg) was precipitated and resuspended in 100µl of HP buffer (40mM PIPES, pH 6.4; 1mM EDTA, pH 8.0; 400mM NaCl; 80% deionized formamide). The solution was transferred to a PCR tube and between 2 to 5 pmol of CY-5-labelled primer (5'-TTG ATA CTC CTG CTT CTT TTG C) was added. The mixture was heated at 95°C for 10 min after which the primer was annealed overnight at 43°C. The RNA with the annealed primer was precipitated by the addition of 300µl ice-cold water and 800µl of ice-cold ethanol, incubation at –20°C for 30 min and centrifugation at 10000 rpm for 12 min (4°C). The subsequent pellet was

washed with 70% ethanol and centrifuged at 10000 rpm for 8 min (4°C). The pellet was then dried and dissolved in 20µl RTB (4ul AMV Reverse transcriptase buffer, Promega; 1µl [10mM] dNTP, 1µl [40U] RNazin, Roche Diagnostics; 2µl [1mg/ml] Actinomycin D; 12µl sterilized water) and the mixture was kept at 42°C for 4 min. AMV Reverse Transcriptase (20U) (Promega, Madison, USA) was added, and the reaction incubated at (42°C) for 2 h. The reaction was stopped by the addition of 1µl 0.5M EDTA (pH 8.0). The RNA was removed by the addition of 1µl 10mg/ml Ribonuclease A (Sigma Chemical Company, St. Louis, Missouri) and left at 37°C for 30 min. The primer extension product was precipitated by the addition of 150µl TES (10mM Tris; 1mM EDTA; 100mM NaCl), 500µl ethanol, incubation at –20°C for 1 h and centrifugation for 12 min at 10000rpm (4°C). The pellet was washed in 70% ethanol, dried and resuspended in 5µl TE, to which 8µl of ALF stop buffer (Pharmacia) was added. The mixture was then heated at 95°C for 5 min before being resolved on an ALF automated DNA sequencer (Pharmacia) along with the products of the appropriate sequencing reactions.

2.4 RESULTS AND DISCUSSION

2.4.1 Southern hybridization analysis and isolation of the truncated *regB* gene

DNA hybridization analysis of *C. beijerinckii* chromosomal DNA using the *regA* gene from *C. acetobutylicum* NCP262 as a probe indicated that this species has a *regA* homologue (Figure 2.1). In addition, it appeared that the region of the chromosome on which the *regA* homologue of *C. beijerinckii* occurred had a different restriction enzyme profile to the corresponding region of *C. acetobutylicum* NCP262 due to the difference in fragment sizes. For *C. beijerinckii* the *PstI* and *ClaI/XbaI* fragments identified by hybridization were ~26 kb and 2.8 kb respectively, while in *C. acetobutylicum* NCP262 the fragments detected occurred at ~16 kb and ~1.25 kb for the same restriction enzyme digestions. This further emphasized the taxonomic difference between the two species. As expected, hybridization occurred to the positive control (lane 4) at ~1.25 kb while there was no hybridization to *E. coli* K12 chromosomal DNA (lanes 2 and 3). The non-specific bands that hybridized together with the 1.25 kb band in the positive control lane are most likely due to incomplete digestion of plasmid DNA, DNA entanglement and probe contamination with plasmid DNA.

In order to obtain the *regA* homologue in *C. beijerinckii*, a *C. beijerinckii* gene bank in plasmid pECOR251 (Reid, personal communication) was probed with the *regA* gene probe. One of the pools of the gene bank was found to harbor the *regA* homologue. This pool was transformed into *E. coli* JM109 and a colony blot was performed. Upon hybridization a positive clone was detected. Small-scale DNA extraction and a Southern blot analysis on this putative clone confirmed that the construct contained a *regA* homologue and the construct was named pECOB1 (results not shown).

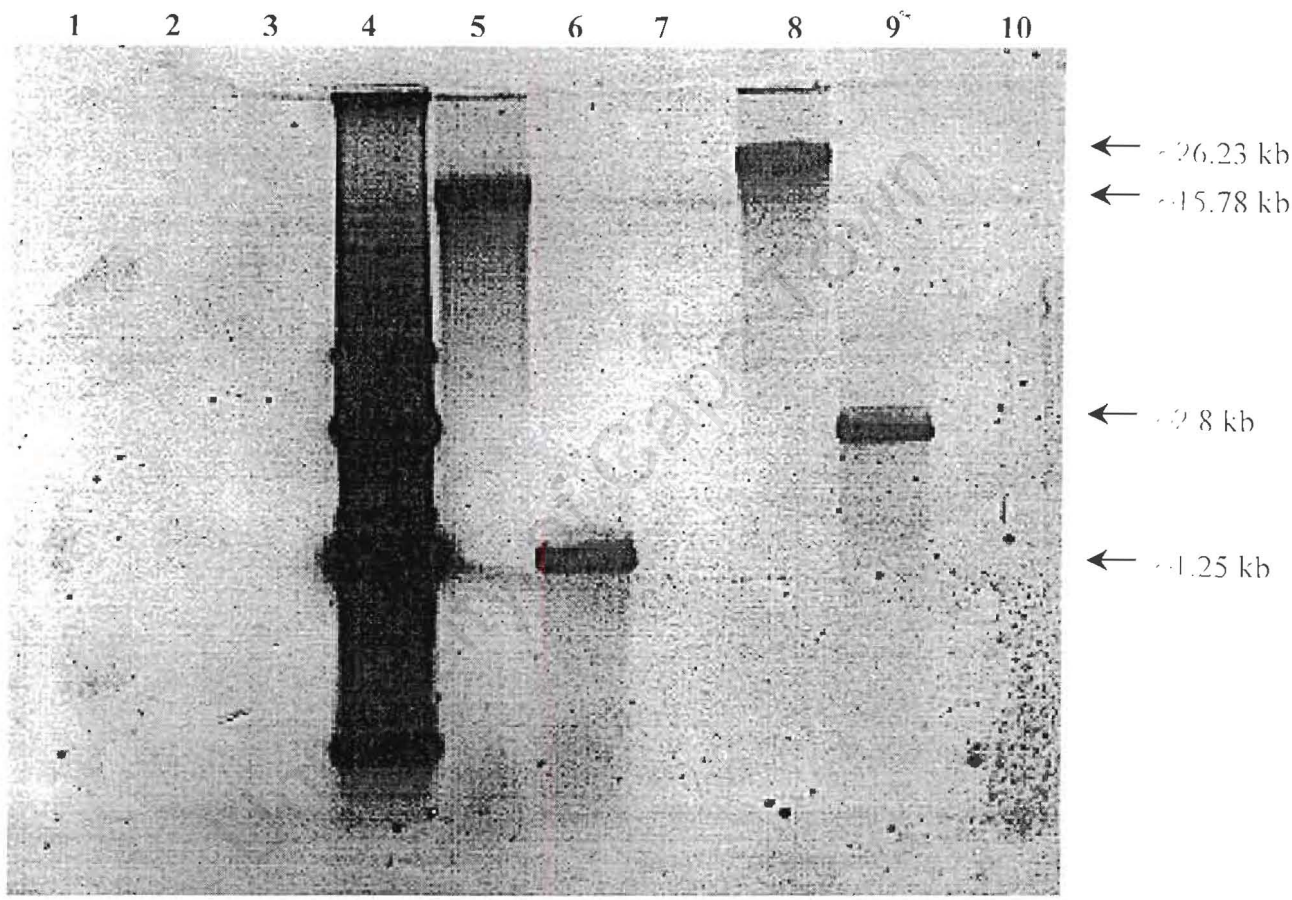

Figure 2.1: Southern hybridization analysis of *E. coli* K12, *C. beijerinckii* and *C. acetobutylicum* NCP262. Chromosomal DNA (50_g) from *E. coli* K12, *C. acetobutylicum* NCP262 and *C. beijerinckii* were digested with *ClaI*/*Xba*I and *Pst*I respectively. The *regA* probe was used throughout.

Lanes 2 and 3: *E.coli* K12, *Pst*I and *ClaI*/*Xba*I digestions respectively.

Lanes 5 and 6: *C. acetobutylicum* NCP262, *Pst*I and *ClaI*/*Xba*I digestions respectively.

Lanes 8 and 9: *C. beijerinckii*, *Pst*I and *ClaI*/*Xba*I digestions respectively.

Lanes 4: Plasmid pSK069, 5_g digested with *ClaI*/*Xba*I, (the positive control).

Lanes 1, 7 and 10: Lambda marker (λ DNA digested with *Pst*I).

2.4.2 Reconstruction of the truncated *regB* gene

Restriction endonuclease and southern hybridization analysis revealed that construct pECOB1 contained an 8.5 kb insert and that the *regB* gene lay near the 3' end of the insert. Further subcloning into pSKbluescriptSK$^+$ was done to localize the *regB* gene for preliminary sequence analysis. Sequence analysis revealed that *regB* showed homology to the previously cloned *regA* but that *regB* was truncated by 266 bp at its 3' end.

Chromosome walking was then employed to find the full-length clone. No appropriate restriction endonuclease digestions of the *C. beijerinckii* chromosome, which gave a suitable DNA fragment containing the entire ORF of *regB* could be identified. A size-selected gene library was therefore constructed using *Cla*I digested *C. beijerinckii* chromosomal DNA and pBluescriptSK$^+$ vector. Although this gene library would have the 5'-end of the *regB* gene deleted (~77 bp) it should still contain sufficient downstream region. A colony blot was performed on the size-selected gene library transformed into *E. coli* JM109, using the *regB* gene as a probe. A recombinant plasmid, pSKB2 was found which contained the downstream region of *regB* on a ~2.76 kb DNA fragment (results not shown).

The two truncated parts of *regB* were then reconstructed to obtain a full-length *regB* gene. Reconstruction of the full-length sequence of *regB* is shown in Figure 2.2.

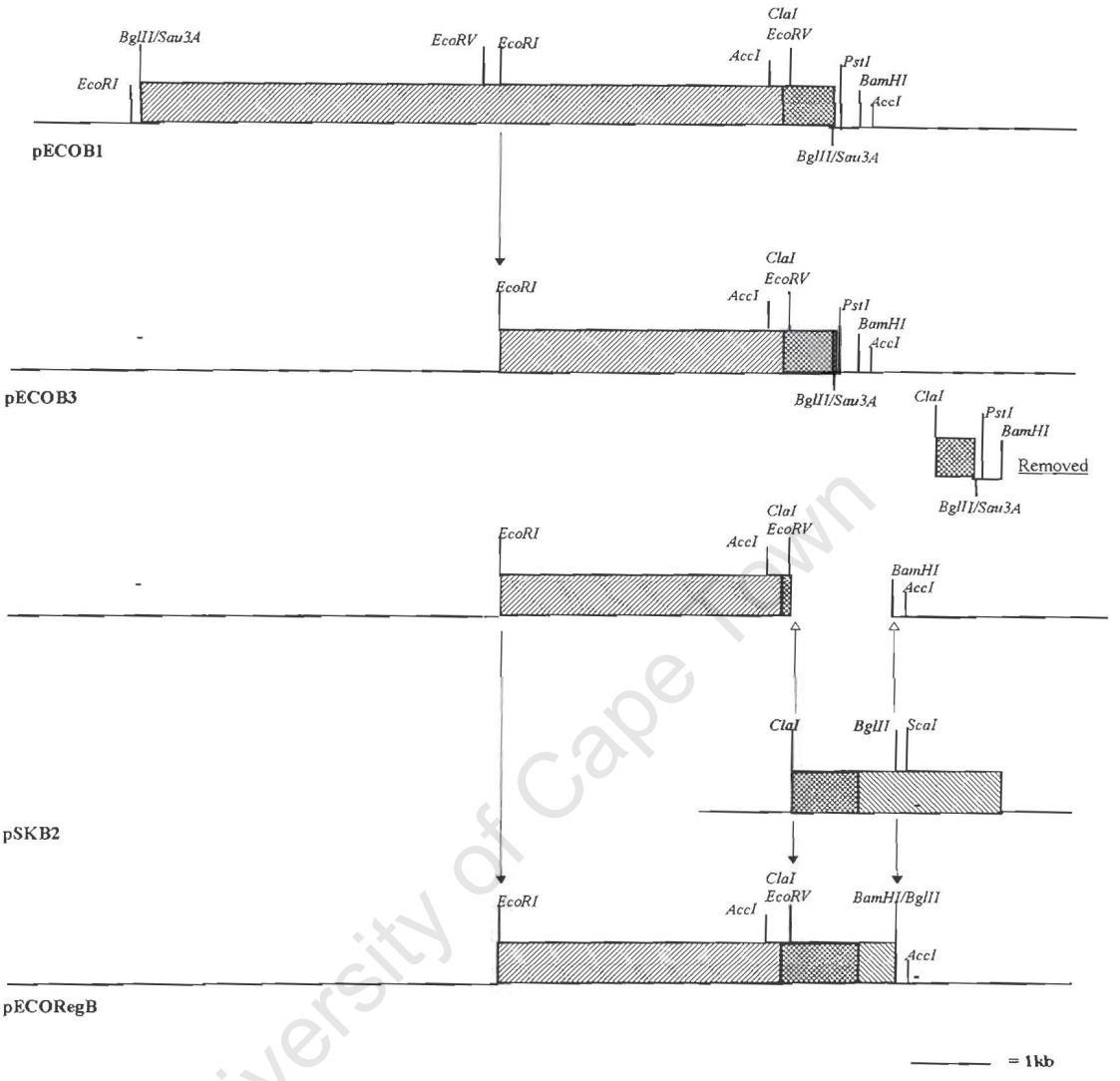

Figure 2.2: Reconstruction of *regB* and its flanking regions. Plasmid pECOB1 was digested with *EcoR*I, which removed a ~5.1 kb fragment and religated to form plasmid pECOB3. This construct contained the upstream truncated *regB* gene on a ~4.2 kb fragment. Plasmid pECOB3 was digested with *Cla*I and *BamH*I, which removed an internal part of the *regB* gene. Construct pSKB2 was digested with *Cla*I and *Bgl*II and the resulting ~1.205 kb *ClaI*/*Bgl*II fragment was subcloned into the previously digested *ClaI*/*Bam*HI pECOB3 plasmid. The resulting construct, plasmid pECORegB contained the entire *regB* ORF in pECOR251 on a ~5.5 kb fragment with adjacent upstream and downstream regions. Plasmid pSKB4 (not shown) contained the ~4.0 kb insert between the two *EcoR*V restriction sites on the pECOB1 insert subcloned into the *EcoR*V site of pBluescriptSK$^+$.

2.4.3 Determination of the transcriptional start of *regB*

C. beijerinckii cells grown were grown in glucose or sucrose in order to determine whether different carbon sources had an effect on transcription of the *regB* gene. A single transcriptional start site for *regB* was identified by primer extension and different carbon sources appeared to have no effect on *regB* expression (Figure 2.3). This result is similar to Miwa et al. (1997) and Heuck et al. (1995), which suggests that CcpA and its homologues from *Bacillus* are expressed constitutively. A global regulator such as CcpA needs to respond quickly to changes in the metabolic status of the bacterium. The constitutive expression of CcpA thus minimizes the unnecessary use of metabolic energy. In contrast, primer extension analysis of *ccpA* genes for lactic acid bacteria and *Staphylococus xylosus*, revealed two transcriptional start sites and there was evidence for autoregulation of CcpA expression in these bacterial species (Egter and Brückner, 1996; Mahr et al., 2000). They found that the expression levels of CcpA were slightly affected by different carbon sources, which suggests that this arrangement of two promoters ensures expression of CcpA under different environmental conditions and allows adjustment to a certain level of active CcpA. The position of the transcriptional start site of the *regB* gene will be discussed in the next section (2.4.4).

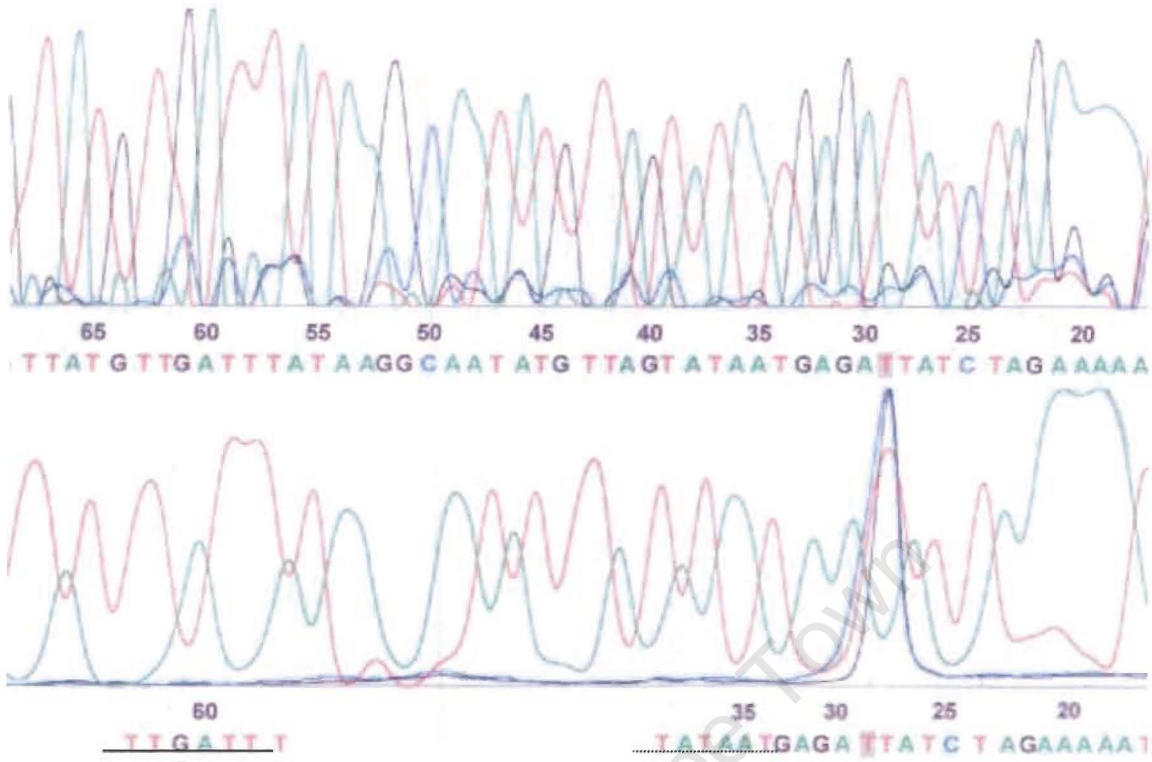

Figure 2.3: Mapping of the transcription initiation site of the *C. beijerinckii regB* gene by primer-extension analysis. [A], The DNA sequencing fluorogram corresponding to the *regB* gene region analysed in [B]. The same Cy^5 labelled primer was used as a primer for both sequencing and primer extension reactions. [B], The fluorogram of the primer extension product obtained using RNA isolated from *C. beijerinckii*. A and T sequencing reactions (green and red curves, respectively) were used as tracking lanes. Both reactions [A] and [B] were run in parallel and the sequence information is given below. The residue identified on the coding strand as the transcriptional start site is given underneath fluorogram [B] and is shaded. The –10 and –35 regions are indicated by a solid and dashed line respectively. Colour code- Green (Adenosine), Black (Guanosine), Blue (Cytosine) and Red (Thymidine).

2.4.4 Nucleotide and protein sequence analysis of *regB* and flanking regions

The genetic organization and restriction endonuclease profile of the *regB* gene and its flanking regions are shown in Figure 2.4. The nucleotide sequence of the full-length *regB* transcript and adjoining upstream and downstream sequences are shown in Figure 2.5.

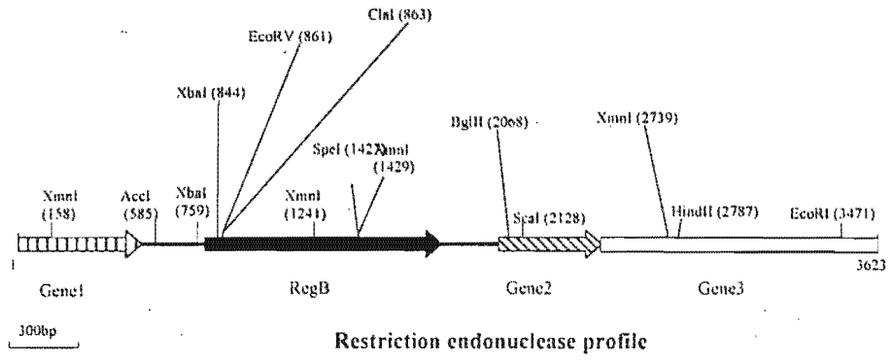

Figure 2.4: The genetic organization and restriction endonuclease profile of *regB* and flanking regions

```
   1  AACTTTGAAA TAAAACCTAA TCTACCTGTT ATAGGTAAGT TATATGGAAA
  51  ATTAATTCCT CAAATCAGAA AGGCAATTTC TGAAAAGAAC CAAATGGAAT
 101  TAGCTCAAAA GATTCAAAAT GGTGGAAGTG AAACTATAAT TGTTAATGAT
 151  ACAGAAATAG TTCTTACAAA TGAGAATCTT TTAGTTACAA TGCAAGGTTT
 201  AGAAGGTTAT GCATTCGCAG GAGAAGGTGA ACTTGGAGTT GTTTTAGATA
 251  CAACTGTAAC ACCAGAACTT CAAGAGGAAG GGCATGTAAG AGAAGTTATT
 301  TCAAAGATTC AAAATATGAG AAAAGATAAA GGATTTGAAG TTGCTGATAG
 351  AATAAATCTT TATGTATCAA ATAACGATAT GTTAATTGAT GTTATTAAGA
 401  AATTCGAACA AACAATAAAG AAAGAGACTT TAACTTGTGA AGTTCTTTAT
 451  AATCAAGAAT CTAATTATTC AGAAACAGTT ATAAATAGTG AAACATTAAA
 501  CATGAATGTT GAAGTTGTAA GATAAAAACT AAAATGACTA GGAATAAAAG
 551  ATTTTATTAA TATTCCTAGT CATTTTAATA TTTGTATACA TATTTAATAT
 601  GGTTTTCATA TATATCTCTT ATATAAGATT GTAAATTAGT TAATATGTTT
 651  ATAAATTTAA GATTAAAAAC TATGATTTTT GACAAAACTT ACAATAAAAT
                          -35                        -10
 701  AATAATTTAT TTTAAGTTAT GTTGATTTAT AAGGCAATAT GTTAGTATAA
                              SD
 751  TGAGATTATC TAGAAAAATT CAAAGGAGGA ATAATAATGG CTACTTCTAT
                                        RegB ------>
 801  TAAGGATGTT GCAAAAGAAG CAGGAGTATC AATTGCAACC GTTTCTAGAG
 851  TTTTAAATGA TATCGATGTA GTAAACGAAG ATACTAAGAA GAAAGTTTTA
 901  GATGCAATAA AAAAGCTTGG ATACAGACCA AATATAGTTG CAAGAAGTTT
 951  AAAAACTCAA AGAACAAAGA CAATAGGAAT TTTACTTCCT GATATATCAA
```

1001 ACCAGTTTTA CCCAGAAATT GTAAGAGGTG CTGAAGATGT TTCAAATATA

1051 TACGATTACA ATATAATACT TTGTAATTCA GACCTTGATA TAGAAAAGGA

1101 AAAAGAATAC TTAAGGGTAC TTAGAGAAAA AATGGTTGAT GGTGTAATTT

1151 ATATGAGTAG TTCTCTTAAT GAAGAAATAT TAGAACTAAT CAATGAATTA

1201 GATATAAAAA CTGTTTTAGT TGAAACTAAG GATAAAGAAG GAGTTCTACC

1251 AAGTGTAACA ATTGATAATG TTAAAGCTAG CTATGATAGT ACAAAGTTAT

1301 TAATTGAAAA GGGAATAAAA GAAATTGCCT TTATAGGAGC TGAAAAAGAG

1351 AGCAAAAATG CTTGGGGTGA TAGATACGTT GGTTATGAAA ATGCAATGAA

1401 GGAATCAGGA ATAGAAATAG ATCCTGAACT AGTTCATCTA AGCTCAATGA

1451 AAGTAAAGAG TGGATACGAG GGCATACAAA AGTTCATAAA GCAAAATAAG

1501 AAATTTAGAG GTGTTGTATG TGCTTCTGAT GATATTGCAA TGGGAGCAAT

1551 TAATGCACTA CGAGATAATA ACTTAGATAT TCCAAAAGAT GTAAGTGTAA

1601 TTGGATTTAA TGATAATTTT GCAGCTTCAA TTTTCTATCC AAAGATCACA

1651 ACAATTTCTC AACCAACTTA TGATATGGGA TCTGTTGCTA TGAGAATGCT

1701 TATAAAACTT TTAAATAAGA AAGAGCTAGA TGAGCCACAT TATGTTTTAG

1751 AGCATCAGCT AGTTGAAAGA GAAAGTACAG TT**TAATAG**AA AAATAATCCA

1801 ATTTTTACT TAGTTAGTAT AAATCTTAAA TGTATATGTT TATAGAGGAC

1851 TACTTCAAGA TAGTTTAGAG CTATTTGGGA GTAGTCCTCT ATTAGCGTTA

1901 AAGTTATAAT TAGTTGCGCA CGTAACTAAT TTATGCTAAT ATAGTTTGTA

-35 -10

1951 TGTAATTATT TGGATATGAA ATA***TGATT***AT TGTTAATACT AT***ATTAAT***AT

SD

2001 GATTAAGATA GG**GGAGAA**GT AAT**ATG**GGTT GCTATCAGCA TATAGGTAAG

Gene2 ⟶

2051 TATATTGGTG AAATACATAG ATCTAGCTAT ATGTATTTTG GCAAGAGATT

2101 TAGTAAATTT GGGATAGGGG CTGGTCAGTA CTTATTTCTT TTAAATCTTT

2151 ATGAAAATGA TGGTATAACA CAAGAAGAAT TGACGAAAAA GGTTAGATTA

2201 GATAAGGCGA CAACAGCTAG GGCAATAAAG AAATTAGAGG ATGAAGGTTA

2251 TGTAAGGAGA ATAAAAAAAG AAAGTGATAA GCGTGCATAT AGATTGGAAC

2301 TAACTGAGAA AGCAGAACAA ATTAAAGATG ATGTATATTC TATAATGAAT

2351 GAATGGGAAT CTGAGATTAG AATGTGTTTT ACTGATGAGG AATCTCAAGA

2401 ATTAATGAAT TTGTTGAATA AGTTATCTAA GAGTTCTTTA ATTAATAA**GG**

SD *

2451 AGGATATTCA TGAATAAGCA GAATAGATTA GGAGAAATGG GTGTAGGAAA

Gene3 ⟶⟶

2501 ATTGTTGCTT GAGTTTTCAA TCCCTGCAAT TATAGGAATG TTAGTTAATA

2551 CATTATATAA CATAATAGAT AGAATATTTA TTGGGCATAT AGAAGGCATA

2601 GGTGAACTCG CCATGGCTGG TGTTGGAATT ACAATGCCAA TAATGTTTAT

2651 AATTTTAGCA TTTGGTATGT TAATCGGAAT AGGTACAGCT ACAAAAGTGT

2701 CAATAAAGCT TGGAGAGCAT GATAAGGAAG GGGCAGAAAG GCTTCTAGGA

2751 AATGCGTTTA CATTAATAAT AATAATTGGT GTATTGTTAA CGTTGTTTGG

2801 TCATATATTT GCTAATCCGC TTTTGAAAGC ATTTGGAGCG AGTGAAAATA

2851 TTATAGGTTA TGGTGAAGAT TTTATAAGGG TAATTATTAG CGGTTGCATA

2901 TTTAATTTGA TGAGTTTCGG TTTGAATCAT TCAATTAGAA GCGATGGAAG

2951 CCCAAAAATA GCAATGGCGT CTATGCTAAT GAGTGCTATT ATAAACATCA

3001 TCTTAGATCC AATATTTATA TTCGGTTTAG GCCTTGGAGT TAAAGGAGGA

3051 GCCCTTGGAA CTGTTGTTGC GCAGACAATA AGCAGTATAT GGATTTTATA

3101 TTATTTCACA AAGGGATCTA GTGTACTTAA GATAAGAAGG AAAAATTTAA

3151 AATTGGAGAA GGACGCAGTT TTAAGTATTT TTGCTATAGG GGTCAGTCCA

3201 TTTAGCATGC AGCTTGCACA AAGTGTAGTT TCAAGTGATA CTAATAATTC

3251 ATTGCAGTCA TATGGTGGTG ATTTAGCCGT AGAGCAAATG ACTATAGTGA

3301 ACAGCTTAGC TATGATATTT CTAATGCCTA TATTTGGATT AAATCAGGGA

3351 CTTCAGCCTA TAATCGGTTA TAATTTTGGC GCTAGAAAAT ATGATAGAGT

3401 TAAGCAAGCG GTAAAATACG GGGTTATTAT AGCTACAATA ATAGTGGTAA

3451 TAGGATTTAT AATAATTGAG GGAATTCCAG AAACATTAGT TAAGATTTTT

3501 AATAATGAAT CATCGCTTAT AGAAGTAACT GCTCATGGAA TGAGAATATT

3551 TTTAATTATG CTGCCATTTA TAGGAGCACA AATTATAATT ACGAATTTCT

3601 TTCAATCAAT AGGAAGAGTA AAAA

Figure 2.5: Nucleotide sequence of the *C. beijerinckii* DNA fragment containing *regB* and its flanking regions. Two stem loop structures (double underlined) were identified, one between positions 530-546 and 561-577 and the second between positions 1836-1861and 1873-1898. Possible hairpin structure occurs at positions 1907-1915 and 1953-1962 (single underlined).

↓ = Transcriptional start; SD = Shine Dalgarno; * = end of gene

⟶ = full-length gene transcript; ⟶ = truncated gene transcript

The *regB* gene was 996 bp in length and encoded a protein of 332 amino acids with a calculated MW of 39 840 Da. The ribosome-binding site was located 7 bp upstream of the translational start (ATG-787bp). This Shine-Dalgarno sequence resembles those reported for the *S. aureus* β-lactamase (Mclaughlin et al., 1981), *C. acetobutylicum* endoglucanase (Zappe et al., 1988) and *regA* genes (Davison et al., 1995). The transcriptional start (756 bp) was located 31 bp from the translational start and the corresponding nucleotide was a thymine. The –35 region and –10 regions were 34 bp and 10 bp away from the start point of transcription, respectively. This promoter region showed homology to the promoter region of the *C. acetobutylicum glnA* and *regA* genes (Janssen et al., 1990; Davison et al., 1995). Downstream of the *regB* gene a region of dyad symmetry was identified between positions 1836-1861 and 1873-1898. Similar regions of dyad symmetry have been demonstrated to be factor-independent terminators (Brendal and Trifnov, 1984). The dyad symmetry downstream of the *regB* gene may allow for the formation of a stable hairpin in the RNA structure. Stable hairpin structures have been implicated in slowing down the RNA polymerase (Martin and Tinoco, 1980). Immediately downstream of this region of dyad symmetry between positions 1907-1915 and 1953-1962 is another possible hairpin structure.

Protein multiple sequence analysis showed that the RegB protein belonged to the GalR-LacI family of transcriptional regulators. The RegB protein sequence showed the highest identity with RegA of *C. acetobutylicum* and the CcpA homologues of *C. perfringens* and *C. acetobutylicum* ATCC824 with identities of 87%, 69% and 62% respectively. The RegA of *C. acetobutylicum* has been previously discussed while the *ccpA* gene product of *C. perfringens* was found to be necessary for sporulation (Knapp, R. et al, 2000, unpublished; Accession number, AAG25712). The CcpA homologue of *C. acetobutylicum* ATCC824 was identified by genome analyses of the unfinished *C. acetobutylicum* ATCC824 genome sequence using the RegB protein sequence as the query sequence (see 2.3.6.).

The RegB protein had identities ranging between, 38% to 35%, with most of the other CcpA homologues isolated from different gram-positive bacteria (see Figure 2.6). RegB was compared phylogenetically with the CcpA homologues (Figure 2.7). The

phylogenetic relationship resembles the taxonomic relationship between the various gram-positive species. Furthermore, phylogenetic comparison of the RegB with other selected protein members of the GalR-LacI family demonstrated that RegB fell within the CcpA subfamily. RegB specifically falls into a group with the CcpA homologues of *C. acetobutylicum* NCP262, *C. perfringens* and *C. acetobutylicum* ATCC824 indicating that the homologues most likely had a common protein ancestor.

Clostridium acetobutylicum ATCC824 (putative CcpA homolog, see 2.3.6); *Bacillus halodurans* (BAB06960); *Lactococcus lactis* (AAC96330); *Lactococcus lactis* subsp. *lactis* (AAK05746); *Lactococcus lactis* subsp. *cremoris* (CAB10075); *Thermoactinomyces* sp. E79 (AAC82367); *Bacillus subtilis* (P25144); *Bacillus megaterium* (P46828); *Listeria monocytogenes* (AAC82783); *Streptococcus mutans* (O07329); *Staphylococcus xylosus* (Q56194); *Lactobacillus pentosus* (AAD53119); *Enterococcus faecalis* (CAA09491); *Lactobacillus casei* (AAC46030).

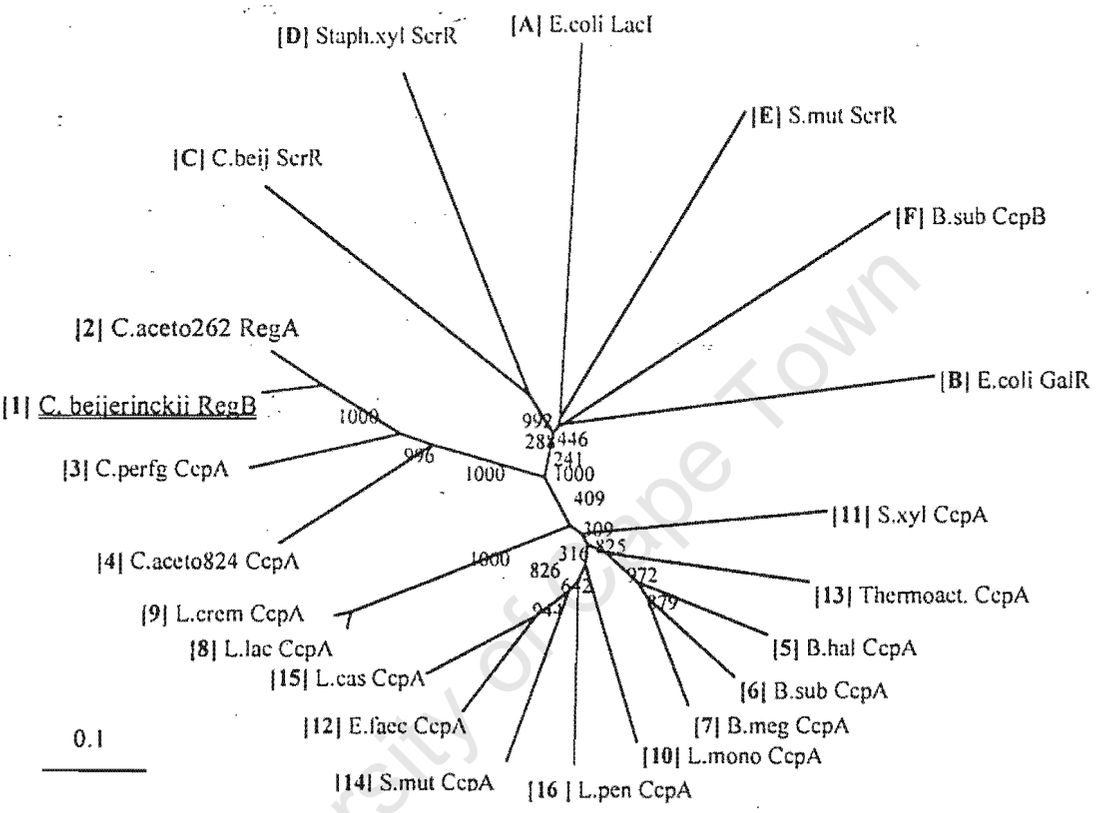

Figure 2.7: Phylogenetic tree for CcpA protein homologues and selected proteins of the GalR-LacI family. Branch lengths are approximately proportional to the amount of sequence difference. The scale of 0.1 indicates 10% amino acid sequence divergence. Bootstrap values were taken over 1000 trials. The CcpA homologues are: **[1]** *C. beijerinckii* (RegB), **[2]** *Clostridium acetobutylicum* NCP262 (Q4583); **[3]** *Clostridium perfringens* (AAG25712), **[4]** *Clostridium acetobutylicum* ATCC824 (putative CcpA homolog, see 2.3.6), **[5]** *Bacillus halodurans* (BAB06960), **[6]** *Bacillus subtilis* (P25144), **[7]** *Bacillus megaterium* (P46828), **[8]** *L. lactis* (*Lactococcus lactis*, AAC96330), **[9]** *Lactococcus lactis* subsp. *cremoris* (CAB10075), **[10]** *Listeria monocytogenes* (AAC82783), **[11]** *Staphylococcus xylosus* (Q56194), **[12]** *Enterococcus faecalis* (CAA09491), **[13]** *Thermoactinomyces* sp. E79 (AAC82367), **[14]** *Streptococcus mutans* (O07329), **[15]** *Lactobacillus casei* (AAC46030), **[16]** *Lactobacillus pentosus* (AAD53119). The selected proteins of the GalR-LacI family are: **[A]** *E. coli* LacI (P03023), **[B]** *E. coli* GalR (P03024), **[C]** *C. beijerinckii* ScrR (AAC99321), **[D]** *Staphylococcus xylosus* ScrR (CAA47969), **[E]** *Streptococcus mutans* ScrR (AAC31628), **[F]** *B. subtilis* CcpB (P37517).

Before a discussion on the features of the RegB protein is presented an overview will first be given on the general features of the CcpA protein. The CcpA protein referred to will be that from *B. subtlis* unless otherwise specified.

CcpA proteins in general are similar to other members of the GalR-LacI family in that they are comprised of a helix-turn-helix DNA-binding domain, a C-terminal effector domain and a dimerization domain. A comparison of 12 CcpA-like sequences with regulators from the GalR-LacI family defined a CcpA subfamily based on similarities among the protein sequences (Kraus et al., 1998). They included RegA of *C. acetobutylicum* in their analyses and though it showed the most deviation from the CcpA consensus, it still contains large conserved regions. Most of the identical residues were clustered in three blocks, two in the N-terminal 100 residues and one in the C-terminal 80 residues.

CcpA binding to the *cre* elements involves two monomers of the CcpA dimer, which contact the major groove at two sites on the same face of the DNA helix (Kim and Chambliss, 1997). Sequence similarities, limited proteolysis (Jones et al., 1997) and mutational data (Kraus and Hillen, 1997) indicated that CcpA, LacI and PurR share a common three-dimensional peptide fold. By using the protein backbone of PurR, the putative structure of the CcpA subfamily was analysed after the introduction of CcpA side-chains. The locations of several of the conserved residues were located in the DNA binding domain and in the co-repressor binding cleft between the N- and C-terminal subdomains of the protein core. However, most of the conserved residues formed a continuous patch with side-chains exposed to the surface. This was a CcpA-specific feature and all mutants in this region were defective in CCR and could not interact with HPr-Ser-P. These conserved residues with their side-chains exposed to the solvent are therefore thought to form the interaction surface for HPr-Ser46-P (Kraus et al., 1998).

It is known that besides HPr-Ser46-P and Crh-Ser-P, glucose-6-phosphate, fructose-1,6-bisphosphate and NADP are also effectors of CcpA (Stülke and Hillen, 1999). Jones et al. (1997) proposed that the CcpA C-terminal domain binding site on HPr-Ser-P encompassed both phosphorylation sites, Ser46 and His15 of HPr. This would allow

CcpA to recognize the phosphorylation state of HPr and thereby link the process of PTS sugar transport with CCR in gram-positive bacteria. Unlike HPr-Ser-P, NADP had little effect on triggering CcpA to bind DNA but increased the ability of CcpA to inhibit transcription which suggests that DNA binding and transcription inhibition are two distinct properties of CcpA. Furthermore mutations in CcpA separated growth effects from catabolite repression suggesting that CcpA discriminates between different signals and could respond either by altered co-factor binding or CcpA binding to other DNA sites (Küster et al., 1999a).

```
                                                                  ⇩
   1  MATSIKDVAK EAGVSIATVS RVLNDIDVVN EDTKKKVLDA IKKLGYRPNI

  51  VARSLKTQRT KTIGILLPDI SNQFYPEIVR GAEDVSNIYD YNIILCNSDL

 101  DIEKEKEYLR VLREKMVDGV IYMSSSLNEE ILELINELDI KTVLVETKDK

 151  EGVLPSVTID NVKASYDSTK LLIEKGIKEI AFIGAEKESK NAWGDRYVGY

 201  ENAMKESGIE IDPELVHLSS MKVKSGYEGI QKFIKQNKKF RGVVCASDDI

 251  AMGAINALRD NNLDIPKDVS VIGFNDNFAA SIFYPKITTI SQPTYDMGSV

 301  AMRMLIKLLN KKELDEPHYV LEHQLVERES TV
```

Figure 2.8: The protein sequence of RegB with key structural features indicated. Amino acids in italics and those underlined indicate a putative helix-turn-helix region and the periplasmic/sugar binding domain respectively. Both regions belong to the LacI family. [R] = located on N-terminal sub-domain. [N] = located in putative co-repressor binding cleft. ⇩ = Amino acid involved in DNA binding conformation of CcpA. Amino acids that are double-underlined are located on the CcpA specific putative HPr-Ser46-P interaction surface. [Aa] = Amino acids involved in postulated interaction between CcpA and Hpr-Ser46-P.

The amino acids of RegB between positions 5 to 71 had 62% identity with the helix-turn-helix region of the *E. coli* lactose operon repressor while amino acids between positions 63 to 282 had 30% identity with periplasmic binding proteins and the sugar-binding domain of the LacI family.

Comparison of RegB with the CcpA subfamily indicated that RegB shared similar key amino acids as the rest of the family (Figure 2.8.). The asparagine residue, Asn-110, is a CcpA-specific residue within a region of residues conserved among all members of the GalR-LacI family. It is located on the surface of the N-terminal subdomain of the core

protein on the opposite site of the postulated HPr-Ser46-P binding area. The side chain of this Asn-110 is directed towards the DNA binding domain of the other monomer. In LacI the respective residue opposite the DNA binding domain is His-112, which is a basic side chain. This residue contacts the binding domain but can be substituted by several other amino acids. In RegB, arginine (Arg-110) is the equivalent residue and is also a basic side chain.

The isoleucine (Ile-191) in CcpA was identified in the region of the co-repressor binding cleft between the N- and C-terminal subdomains. In PurR, threonine (Thr-192), an uncharged polar side chain, is the equivalent residue and makes direct contact with the co-repressor hypoxanthine. In RegB, asparagine, which is also an uncharged polar side chain, is found at position 191. Kraus et al. (1998) suggest that this specific sequence conservation supports the hypothesis that a ligand other than HPr-Ser46-P may bind and exert its effect on CcpA.

In CcpA, when Arg-47 is substituted by serine, the protein is completely defective in CCR. Arg-47 is located in the turn between a helix and the hinge-helix of the DNA-binding domain, and its side-chain points towards the N-terminal subdomain of the other monomer. Mutants in the corresponding amino acid position in LacI also yield inactive proteins, and the amino acid contacts the protein core. This mutation interferes with the DNA binding conformation of CcpA and confirms that HPr-Ser46-P interacts with the DNA binding confirmation of CcpA. In RegB, arginine is found in the corresponding position.

In CcpA, alanine (Ala-299), arginine (Arg-303) and tyrosine residues (Tyr-89, Tyr-295) are located in the putative CcpA specific surface and the side chains are solvent exposed. All mutants in these residues were defective in CCR and mutants for Tyr-295, Ala-299 and Arg-303 showed no interaction with HPr-Ser46-P. These mutations indicated that the ability of CcpA to perform CCR is correlated to HPr-Ser46-P interaction and the solvent exposed residues conserved in the CcpA subfamily form this interaction surface. These mutants also showed weak or no detectable retardation of *cre* DNA and there was no stimulation of *cre* binding by HPr-Ser46-P, which confirmed the

importance of this CcpA specific surface. In RegB, all of the specific residues identical to CcpA were found except for Ala-299, where Ser-299, an uncharged polar side chain, was in the corresponding position.

The charge distribution on the surface of CcpA is conserved and the postulated site of interaction between CcpA and HPr-Ser-46-P is surrounded by a number of lysine and arginine residues, which have basic side chains. These residues are the only positively charged region on the surface of CcpA; the majority of the surface is negatively charged. The positively charged residues at the binding site are important because CcpA binds phosphorylated HPr but not HPr alone. The postulated interaction surface is close to the DNA binding domain of the other monomer. Co-repressor binding so close to the DNA-binding domain could therefore directly force CcpA into the DNA-binding conformation. In CcpA, lysine residues (Lys-58, Lys-59 and Lys-90) and arginine residues (Arg-53, Arg-303 and Arg-47 of the other monomer) surround the interaction site. In RegB, the corresponding residues were Gln-58, Arg-59, Asp-90, Arg-53, Arg-303 and Arg-47, which makes this region also predominantly positively charged. Though Asp-90 is an acidic side chain, its charge contribution in such a positively charged region could be negligible, the RegB protein conformational structure could allow variance at the interaction site or the residue could contribute to the structural conformation of the binding site.

The sequence, structural and phylogenetic information all indicate that RegB is a CcpA homologue, belonging to the CcpA subfamily of the GalR-LacI family of transcriptional regulators.

The deduced amino acid sequence from the Gene1 ORF had the highest sequence identity with isoleucyl-tRNA synthetase proteins. This amino acid identity (30.06% identity and 51.45% similarity) was to the isoleucyl-tRNA synthetase (isoleucine-tRNA ligase) protein of the Lyme disease spirochaete, *Borrelia burgdorferi*. This protein identity was primarily confined to the end of the *B. burgdorferi* isoleucyl tRNA synthetase protein due to partial sequence data. Further preliminary and partial sequencing of construct pECOB4, which contained further upstream sequence of *gene1*,

also showed similarity to isoleucyl-tRNA synthetases (results not shown). The partial sequencing of *gene1* together with restriction map analysis and comparison with other isoleucyl tRNA synthetases indicated that *gene1* is approximately 3 kb long. The *B. burgdorferi* isoleucyl-tRNA synthetase ORF encodes 1042 aa which corresponds to a DNA length of 3.1 kb (Fraser et al., 1997). These results when taken together indicate that *gene1* most likely encodes an isoleucyl-tRNA synthetase that lies immediately upstream of *regB*. A region of dyad symmetry was identified between positions 530-546 and 561-577 downstream from the isoleucyl tRNA synthetase gene. This region of dyad symmetry is suggestive of factor-independent terminators and could allow for the formation of a stable hairpin structure in the RNA.

Davison et al. (1995) also found an isoleucyl tRNA synthetase gene upstream of *regA* in *C. acetobutylicum* NCP262. An isoleucyl tRNA synthetase with amino acid identity to one found in *Deinococcus radiodurans* (strain R1) was also identified immediately upstream of the *C. acetobutylicum* ATCC824 putative CcpA homologue by sequence analysis. The *C. perfringens ccpA* sequence data were unavailable and therefore no sequence analysis was possible of its flanking regions in order to determine whether the genetic organization was similar to that of RegB. In *B. subtilis*, a gene encoding DAPH synthase chorismate mutase (AAC00298.1) was found upstream of the *ccpA* gene.

The isoleucyl-tRNA synthetase of *B. burgdorferi* belongs to the class-1 aminoacyl-tRNA synthetase group and is located in the cytoplasm. Isoleucyl-tRNA synthetase is involved in protein biosynthesis and catalyzes the de-phosphorylation of a L-isoleucine-tRNA(Ile) into L-isoleucyl-tRNA(Ile) with the liberation of AMP and pyrophosphate (Voet and Voet, 1995).

Gene2 was 441 bp in length and encoded a protein of 147 amino acids with a calculated MW of 17 640 Da. No termination signals could be detected for this gene and the stop signal occurred immediately within the adjacent downstream gene. The Gene2 ORF encoded a protein with identity to transcriptional regulators of *Lactococcus lactis*, *B. subtilis*, *B. halodurans*, *Staphylococcus sciuri*, *Vibrio cholerae* and *E. coli* (Figure 2.9). Nearly all of these transcriptional regulators belonged to the MarR family of

transcriptional regulators. The protein identities of the encoded *gene2* product was 31%, 28%, 26%, 23%, 21% and 20% to the *L. lactis*, *B. subtilis*, *B. halodurans*, *E. coli*, *Staph. sciuri*, *V. cholerae* regulators respectively.

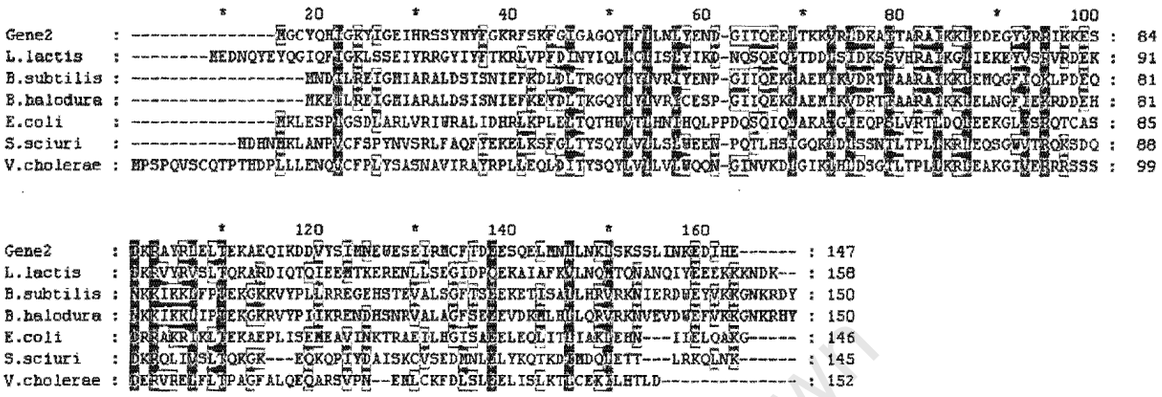

Figure 2.9: Multiple sequence analysis of Gene2 with *L. lactis subsp. lactis* [AAK04840, conceptual translation], *B. subtilis* [P37503, hypothetical transcriptional regulator in the *cotF-tetB* intergenic region], *B. halodurans* [BAB06796, transcriptional regulator MarR family], *E. coli* (P55740, SlyA, transcriptional regulator MarR family), *S. sciuri* [CAA73494, ORF145) and *V. cholerae* (group O1 strain N16961) [B82389, transcription regulator MarR family].

Presently no function has been linked with the above regulators except for the *E. coli slyA* gene product. In *E. coli*, the *slyA* gene product is a transcriptional activator that activates a cryptic haemolysin and thereby induces haemolytic activity (Oscarsson et al., 1996).

The nucleotide sequence, ORF145 was identified in the *mecA* region of methicillin-susceptible *S. sciuri* subsp. *carnaticum* K11 (Wu et al, 1998). The *mecA* region showed 85% nucleotide sequence identity to a segment of DNA in the *S. aureus mecA* hypervariable region. This finding was significant because of the high degree of DNA sequence similarity and the importance of the hypervariable region in the transposition of *mec* DNA. The 3' sequences flanking *mecA* and the DNA sequence in the hypervariable region are believed to be foreign DNA for *S. aureus* (Archer and Niemeyer, 1994). The ORF145 was located in the hypervariable region between *mecA* and an IS-like element IS*431*. This polypeptide was homologous to the N terminus of the glycerophosphoryl diester phosphodiesterase (UgpQ) of *E. coli*. The presence a

ugpQ-like gene together with the homolog of *mecA* in *S. sciuri* subsp. *carnaticum* strain K11 supported the hypothesis that these genetic elements may be evolutionary relatives or precursors of the genetic determinant of methicillin resistance in *S. aureus*. The function of glycerophosphoryl diester phosphodiesterases is the hydrolysis of deacylated phospholipids to sn-glycerol-3-phosphate and the corresponding alcohols (Tommassen et al, 1991). There are two isozymes of this enzyme, namely GlpQ and UgpQ. Both sequences are related; the former is located in the periplasm and the latter is cytosolic.

The *marR* gene of *E. coli* encodes a negative regulator of the *marRAB* operon, a regulatory locus controlling an adaptational response to antibiotics and other environmental hazards (Alekshun and Levy, 1999). MarA, a transcriptional activator encoded within the *marRAB* operon, regulates the expression of multiple genes on the *E. coli* chromosome. An inducer of these genes, 2-hydroxybenzoate (salicylate) binds to MarR and inhibits its ability to bind to the *marAB* operator region thereby allowing transcription of the *marAB* operon.

Many structurally dissimilar chemicals affect MarR activity in whole cells and the MarR family is identified as a group of regulatory factors whose activity is modulated in response to environmental signals in the form of phenolic compounds. It is also postulated that naturally occurring toxic compounds may be the effector molecules for this family of regulators (Miller and Sulavik, 1996). This family contains a conserved sequence of amino acid residues consistent with a similar tertiary structure for all members. The range of functions regulated by MarR family members are varied and besides regulation of antibiotic resistance in *E. coli* and *S. typhimurium* (Alekshun and Levy, 1997; Sulavik et al., 1997) they are involved in negative regulation of sporulation and protease production in *B. subtilis* (Koide et al., 1999) and anaerobic aromatic compound degradation in the facultative anaerobe *Rhodopseudomonas palustris* (Egland and Harwood, 1999). Members of the MarR family generally act in a negative fashion and are inactivated by small-molecule effectors. However MarR regulators have been found which act solely as positive regulators or have both positive and negative regulatory effects on regulatory effects on expression (Koide et al., 1999). BadR of *R. palustris* (Egland and Harwood, 1999), NhhD of *Rhodococcus rhodocrous* (Komeda et

al., 1996) and SlyA of *E. coli* (Oscarsson et al., 1996) all activate rather than repress gene expression. BadR regulates anaerobic benzoate degradation in response to benzoate as a carbon source, NhhD is involved in the gene cluster of nitrile hydratase and the function of SlyA was previously described. MexR acts as both a positive and negative regulator of expression of the multidrug resistance operon *mexA-mexB-oprM* in *Pseudomonas aeruginosa* (Poole et al., 1996). No work has yet been reported on the mechanism responsible for transcriptional activation by any of these proteins.

The Shine-Dalgarno sequence and translational start of *gene3* occurred within the *gene2* nucleotide sequence (2 bp away from stop signal). No consensus promoter region could be detected for this gene and the gene was truncated at the 3' end. It is possible that *gene2* and *gene3* are both transcribed from the promoter region of *gene2*, and form part of an operon. Alternatively a weak promoter region for *gene3* may be present within *gene2*. This is indicative of a gene whose encoded protein is required in small amounts in the bacterium. Since no termination signals could be detected for *gene2*, it tentatively supports the idea that both genes form part of an operon. However since *gene3* is truncated it could not be determined whether termination signals occur at the end of this gene. Gene3 had 30% protein identity with an unknown conserved protein (BAB05882) in *B.halodurans*. Protein alignments of Gene3 indicated identity with an uncharacterized membrane protein family, UPF0013. The function of this protein is currently unknown.

It was found that the genetic organization downstream of *regB* when compared to the downstream region of the putative *ccpA* gene homologue from *C. acetobutylicum* ATCC824 was significantly different. Analysis of the region immediately downstream of the *C. acetobutylicum* ATCC824 *ccpA* homologue revealed the respective ORFs with the following protein identities; 29% identity to a disease resistance protein from *Aradopsis thaliana* (AC082643), 36% identity to a decaprenyl cis transferase (YteR-A69991) from *B.subtlis*, 39% identity to a LacI transcriptional regulator (F72282) from *Thermotoga maritima* and 56% identity to a deoxy-gluconate oxidoreductase (P50842) from *B. subtilis*.

Mahr et al. (2000) found that when comparing genetic structures of *ccpA* regions, certain taxonomic groups possess a characteristic gene order. Lactic acid bacteria possessed the *pepQ-ccpA*-variable gene order whereas bacilli and *Staphylococcus xylosus* possessed the *aroA-ccpA*-variable-*acuC* gene order. It was of interest therefore to determine whether the clostridia possess a unique *ccpA* gene order. Based on a comparison of the *C. beijerinckii*, *C.acetobutylicum* NCP262 and *C. acetobutylicum* ATCC824 *ccpA* regions we tentatively propose a isoleucyl tRNA synthetase-ccpA-variable gene order for certain clostridia. This conservation of gene order among groups of gram-positive bacterial species reflects the taxonomic relationships between the group members and is most likely a product of the group's overall evolution.

2.5 CONCLUSION

We have successfully cloned and reconstructed the *regB* gene of *C. beijerinckii*. Sequence analysis revealed that the *regB* gene shares identity with the previously characterized *regA* of *C. acetobutylicum* and the protein encoded by *regB* belongs to the GalR-LacI family of transcriptional regulators.

The sequence, structural and phylogenetic analyses of RegB clearly demonstrated that RegB is a CcpA homologue. The RegB protein had the closest phylogenetic relationship with the CcpA homologues from *C. acetobutylicum* NCP262, *C. perfringens* and *C. acetobutylicum* ATCC824, which was a reflection of their taxonomic relationship.

The genetic organization of the *regB* gene on the chromosome was different to that of other previously identified gram-positive CcpA homologues. This difference in genetic organization was attributed to the taxonomic differences between the bacterial species. However there is evidence to suggest that the *ccpA* region of clostridia may possess a characteristic *ccpA* gene order as reported for bacilli and lactic acid bacteria (Mahr et al., 2000). An isoleucyl-tRNA synthetase, which had identity with the same gene from *B. burgdorferi* was found upstream of *regB* gene and belonged to the class-1 aminoacyl-tRNA synthetases. In *C. acetobutylicum* NCP262 and *C. acetobutylicum* ATCC824 an isoleucyl-tRNA synthetase was also found upstream of the *ccpA* gene homologues. No sequence information was available for the downstream region of *C. acetobutylicum*

NCP262 and for the flanking regions of the *C. perfringens ccpA* gene for purposes of genetic comparison. Southern hybridization analysis indicated that *C. acetobutylicum* NCP262 has a different restriction endonuclease profile to that of *C. beijerinckii*. We found through analysis of the *C. acetobutylicum* ATCC824 genome that the genetic organization downstream of the putative *ccpA* was different to that of the *regB* downstream region. Based on these results we propose that certain clostridia possess a characteristic *ccpA* genetic structure, namely the isoleucyl-tRNA-synthetase-ccpA-variable gene order. Future work could use RT-PCR analysis in order to verify and compare the operon structure of the *ccpA* regions from the various clostridial species.

Immediately downstream of the *regB* gene was a putative transcriptional regulator, which belonged to the MarR family of transcriptional regulators (Sulavik et al., 1995). Adjacent to this putative transcriptional regulator was a gene, which had identity with an uncharacterized membrane protein family, UPF0013, whose function is unknown. No promoter region could be identified for this uncharacterized membrane protein and its translational start was within the putative transcriptional regulator. In order to determine whether the putative regulator and membrane protein are transcribed from the same promoter, which is suggestive of an operon, further downstream sequencing and northern blot analysis is required.

The conservation in the RegB protein of key structural and functional amino acid residues within the CcpA subfamily indicates conformational similarity with CcpA. This conformational identity with CcpA homologues strongly suggests functional identity as well. The *C. acetobutylicum* NCP262 *regA* gene was capable of restoring glucose repression of α-amylase activity as well as acetolactate synthase activity in a *B. subtilis ccpA* mutant (Davison et al., 1995). RegA and RegB therefore probably serve the same global regulatory roles in CCR in *Clostridium* species as the CcpA homologues do in other gram-positive bacteria.

In order to demonstrate the functional identity of RegB at the physiological level and to better understand CCR in *C. beijerinckii*, inactivation of the *regB* gene was attempted.

CHAPTER THREE

INACTIVATION OF THE REGB GENE OF *C. BEIJERINCKII* NCIMB 8052

CONTENTS

CHAPTER THREE

3.1 SUMMARY

Attempts were made to inactivate the *regB* gene of *C. beijerinckii* using two suicide inactivation systems. The first was based on the pMTL30 inactivation system which relied on conjugation (Wilkinson and Young, 1994). An alternative inactivation system, based on the thermosensitive pG$^+$host system, (Biswas et al., 1993) was developed for *C. beijerinckii*. This required electroporation of plasmid DNA into *C. beijerinckii*, followed by growth of the transformants at a non-permissive temperature to prevent plasmid replication and to force integration to take place.

Key factors affecting both systems were the reproducibility of the individual DNA transfer techniques and *C. beijerinckii* strain degeneration (Kashket and Cao, 1995). Attempts to reduce the effects of strain degeneration by growth at lower temperatures, and in the case of conjugation, to initiate the procedure near the end of the exponential phase, did not remove strain degeneration entirely but helped to reduce the frequency of the condition. We attempted to increase the conjugation frequency in order to concomitantly increase the recombination frequency by using heat treatment based on a study by Schäfer et al. (1990). Subjection of *C. beijerinckii* cultures to a brief heat treatment resulted in only a 50-fold difference between the conjugation frequencies of heat-treated cells and the untreated cells and this was not a considered a significant difference when compared to previously published values. The results obtained from the electroporation procedure were variable and inconsistent as has been reported previously (Minton et al., 1993b).

Southern hybridization analysis revealed that the *regB* gene of *C. beijerinckii* was disrupted using the pMTL30 suicide system and that the pMTL30RegB construct had integrated in the *C. beijerinckii* chromosome. Subjection of the *C. beijerinckii regB* mutant to sporulation led to the subsequent loss of the mutant upon germination. The development of new inactivation vectors using the pG$^+$host system was also not successful in producing a *C. beijerinckii regB* mutant by either disruption or deletion of the *regB* gene. Though single cross-over and double cross-over integration events

occurred, Southern hybridization analysis revealed that the Em^r ex-conjugants retained the wild-type *regB* gene, and the pGAhost2 vector had re-arranged in *C. beijerinckii*.

3.2 INTRODUCTION

Deregulation of CCR in *ccpA* mutants has been observed for many gram-positive species (Küster et al., 1999a; Kraus and Hillen, 1997; Heuck and Hillen, 1995). These *ccpA* mutants have helped researchers to identify the components and to understand the mechanisms of CCR in most gram-positive bacteria (Küster et al., 1999b). Furthermore these CcpA mutants have allowed molecular and physiological analysis of specific carbohydrate genes (Schmiedel et al., 1997; Gosalbes et al., 1997; Pujic et al., 1998).

In order to understand the effects of CCR on the *C. beijerinckii* sucrose utilization system, the mechanisms facilitating CCR in *C. beijerinckii* needed to be understood. Sequence and structural analysis had indicated that RegB belongs to the CcpA subfamily of the GalR-LacI family of transcriptional regulators. The creation of a *regB* mutant would therefore facilitate the analysis of CCR mechanisms in *C. beijerinckii* and would specifically permit determination of the potential role of CCR in the regulation of the *scrARKB* operon by RegB.

3.2.1 Gene inactivation systems for *C. beijerinckii*

In order to inactivate genes in *C. beijerinckii*, the pMTL30 inactivation system (Wilkinson and Young, 1994), was available, which relies on conjugation with a bacterial donor for suicide vector transfer to *C. beijerinckii* (See Chapter 1). The vector is unable to replicate in *C. beijerinckii* and integration is therefore achieved by homologous recombination. The plasmids pMTL20 and pMTL21 are analogous to the pUC plasmids but lack a functional *nic* site and therefore cannot be mobilized even in the presence of trans acting mobilization proteins (Chambers et al., 1988, see Figure 3.1). Three different gram-positive antibiotic resistance genes were inserted into a unique site on these vectors, namely the Em^r gene of plasmid pAMβ1, the Cm^r gene of plasmid pC194 and the Tc^r gene of plasmid pJIR71 with the resultant plasmids being pMTL20/21E, pMTL20/21C and pMTL20/21T respectively (Minton et al., 1988).

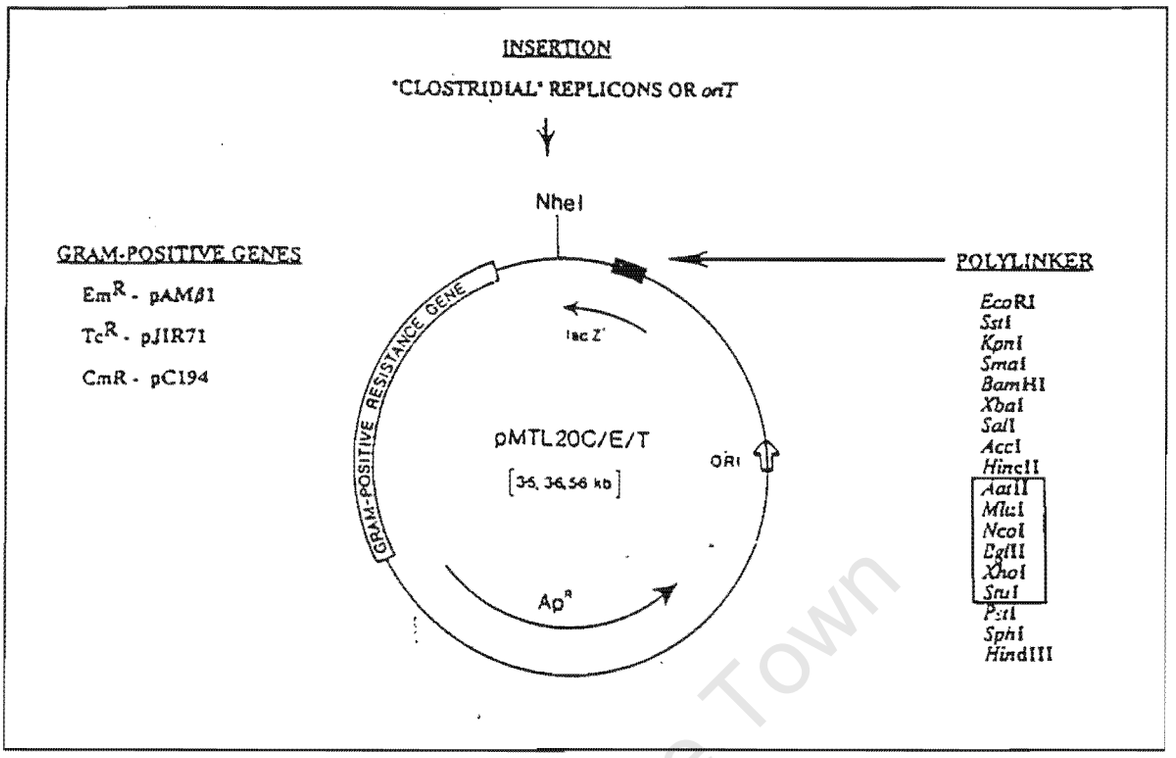

Figure 3.1: The pMTL gram-positive cloning vectors. Restriction fragments carrying the antibiotic resistance genes were a 1.12 kb *Hha*I fragment (pAM_1 Em^r), a 1.07 kb *Hpa*II fragment (pC194 Cm^r), and a 2.9 kb *Sst*I/*Pst*I fragment (pJIR71 Tc^r). Additional cloning sites (to those in pUC18) are boxed. The indicated *Nhe*I site represents the position at which the pAMβ1 replicon was inserted in the derivation of pMTL500E, and the oriT fragment of RK2 used to generate pMTL30/pMTL31. Adapted from Minton et al. (1990).

The *oriT* fragment of plasmid RK2 was inserted in either orientation into pMTL20E to form pMTL30/31 in order for the plasmid to be mobilized from *E. coli* cells to *C. beijerinckii* (Williams et al., 1990b). These plasmids lack a gram-positive replication origin and cannot therefore be established in *C. beijerinckii*, but can be transferred to *C. beijerinckii* from *E. coli* via conjugation. The pMTL500E shuttle vector (Oultram et al., 1988a) contained the pAMβ1 replication origin and the insertion of the *oriT* segment resulted in a high-copy number vector pCTC1.

Wilkinson and Young (1994) successfully used the pMTL30 vectors to inactivate the *gutD* and *spo0A* genes in *C. beijerinckii* by homologous recombination. In this system, the target DNA recombines with the homologous chromosomal region via a Campbell-like recombination mechanism. Reid et al. (1999) used the same system to successfully

disrupt the sucrose transcriptional repressor, *scrR* and sucrose-6-phosphate hydrolase, *scrB*, of the sucrose operon of *C. beijerinckii*. The bacterial donor for the conjugation of these suicide vectors was *E. coli* CA474, which is an *E. coli* HB101 derivative (Table 3.1). A temperature sensitive λ prophage is present in this *E. coli* strain and the lysogen is induced at temperatures of 37°C and above. The use of this *E. coli* donor helped in screening for transconjugants since the donor background was reduced at the non-permissive temperature.

Wilkinson and Young (1994) compared the recombination frequency of *C. beijerinckii* target DNA ranging in sizes from 0.34 to 3.8 kb in length. Both single and amplified copies became established in the *C. beijerinckii* genome and there was no clear correlation between the size of the DNA regions and the frequency of plasmid establishment. They obtained only a four to seven-fold higher frequency with the 3.8 kb segment as compared with the smaller segments. It appeared therefore that other factors such as chromosomal location, DNA base composition and sequence affect the recombination frequency. The plasmids were stably maintained in the chromosome and the frequency of non-replicative plasmid establishment ranged between 2.6×10^{-7} to 1.7×10^{-6} transconjugants per recipient. Campbell-like integration events generate a duplication of target sequences and a reversal of the integration event is possible in the presence or absence of erythromycin selection.

As an alternative inactivation system for *C. beijerinckii*, a new vector based on that of Biswas et al. (1993) was developed. These authors had created the high efficiency pG$^+$host system for single and double-crossover homologous integration in Gram-positive bacteria. The gene inactivation and replacement system was based on plasmid pVE6004 (renamed pG$^+$host4). In order to measure single crossover integration (sco) frequencies along the *L. lactis* chromosome, random chromosomal fragments ranging from 0.9 to 1.4 kb were cloned into pG$^+$host4 derivative and then established in *L. lactis* at 28°C, after which erythromycin resistant integrants were selected at 37°C. The integration frequency per cell (ipc) was found to be between 10^{-2} to 10^{-7} and the ipc of pG$^+$host5 without insert was between 10^{-6} and 10^{-7}. The chromosomal inserts could be classified into two groups. In group I, the lengths of homology of the chromosomal

inserts were the same and therefore variations in the ipc were due to the nature of the homologous insert or local DNA conformation surrounding the fragment rather than to its size. For group II, the integration frequency was found to be very close to the background level and these integrations were found to be nonspecific. These thermosensitive vectors replicate via the rolling circle mechanism and could accumulate high-molecular-weight (HMW) molecules, which would alter the plasmid conformation and in turn have an affect on the recombination frequency. They found that two plasmids produced HMW but as both plasmids belonged to Group I, it suggests that this plasmid form does not strongly affect the intermolecular recombination frequency. Analysis of the integrants revealed that 'sco' homologous recombination occurred in eight of the ten plasmids of Group I. They suggest that the random integration of the Group II plasmids indicated that homologous recombination was a lethal event for these plasmid constructs.

An examination was made of the relationship between the ipc and the lengths of homology between 356 and 2 552 bp. A log-linear relationship was observed between the insert size and frequency of integration with 330 bp being the threshold level for homologous integration. Recombination frequencies reached a plateau for homologous segments larger than 2.5 kb. Factors other than length such as HMW production, structural instability and copy number appear to be important since the ipc frequencies for the same 3.9 kb segment in two orientations on the vector differed by 18-fold. The pG^{+}host thermosensitive system was shown to work in *B. subtilis* and *L. lactis*, indicating that this system could work in other gram-positive bacteria as well.

Noirot et al. (1987) reported that an active rolling circle replicon inserted into the chromosome stimulates homologous recombination between flanking repeated sequences 20 to 450 times. Biswas et al. (1993) made use of this stimulatory effect of rolling circle replication to develop a protocol for 'dco' gene replacement. Strains containing plasmids integrated by sco at the non-permissive temperature using harbour copies of the plasmid flanked by directly repeated sequences. When these strains are shifted to the permissive temperature, the ensuing replication of the plasmid will strongly stimulate a second recombination event resulting in high-frequency excision of

the replicon, giving rise to either a parental or 'dco' chromosomal structure. In this study, they used two derivatives of pG$^+$host4, which carried either a contiguous, or two non-contiguous chromosomal segments. The non-contiguous segments were separated by a tetracycline marker. The *L. lactis* strains carrying these plasmids were grown either at the non-permissive (37°C) or permissive temperatures (28°C). In the case of the plasmid carrying the tetracycline marker, recombination resulted in replacement of the chromosomal gene by the modified segment on the plasmid. This allowed them to obtain 50 to 98% gene replacement when tetracycline selection was used for the replaced gene. The same protocol was followed with the plasmid with the contiguous chromosomal segment, but without tetracycline selection. In this case 1 to 40% replacement was obtained when no selection was applied, which demonstrated the feasibility of gene replacement without antibiotic selection. There was no 'dco' recombination in the absence of replication for both plasmids.

The above pG$^+$host thermosensitive system had never been utilized in *C. beijerinckii*. However the pG$^+$host thermosensitive system is based on the replicon of pWV01 which was shown to facilitate plasmid replication in *C. beijerinckii* (Williams et al., 1990a; Minton et al., 1990). This suggested that the pG$^+$host theremosensitive system could possibly function in *C. beijerinckii*. However, the thermosensitivity of *C. beijerinckii* and the background integration levels of the pG$^+$host vectors were unknown factors. We therefore undertook a study to develop a thermosensitive inactivation system for *C. beijerinckii* based on the pG$^+$host vectors, in order to construct a *C. beijerinckii regB* mutant.

3.2.2 Factors affecting DNA transfer systems in *C. beijerinckii*

3.2.2.1 Strain Degeneration

C. beijerinckii strain degeneration is the spontaneous loss of ability to produce solvents and form spores, and has been analysed by Kashket and Cao (1993; 1995). This typically occurs when *C. beijerinckii* is subcultured from batch culture, as opposed to germination from heat-treated spores. Strain degeneration in *C. beijerinckii* can be caused by a number of factors such as excessive acidification of the culture, a

consequence of culture longevity, mutations in genes involved in solvent production or selective growth advantages of degenerate strains.

The most common mechanism of strain degeneration is thought to occur during exponential growth. *C. beijerinckii* ferments glucose to acetic acid at an uncontrolled rate, so that, during rapid growth, the rate of acid production exceeds the rate of induction of the solventogenic pathway enzymes. The resultant drop in the pH of the medium to bactericidal levels prevents the cells from initiating solventogenesis or sporulation and results in cell death. Thus the physiological state of *C. beijerinckii* appears to be poised either to produce excess acids or to initiate solventogenesis depending on small differences in the growth rate. There is evidence that this behavior is subject to a global regulatory region on the *C. beijerinckii* chromosome. In continuous culture, degenerate cells continue to divide while non-degenerate cells stop dividing and differentiate. Kashket and Cao (1995) therefore reduced the growth rate, which reduced the rate of acidification and thereby allowed a successful switch between acidogeneis and solventogenesis.

Clostridial degeneration is thus the result of progressive overgrowth by phenotypic variants, which are deficient to varying degrees in solvent and spore formation. However over-acidification is merely one cause of strain degeneration since degeneration also occurs under pH-controlled conditions (Meinecke et al., 1984). Furthermore, Kashket and Cao (1995) found that degeneration occurred within a relatively few generations in liquid and solid media. They also found that each *C. beijerinckii* colony was a heterogenous mixture of cells with respect to degeneration with the rate of degeneration rapid enough for each colony to have an individual composition of degenerate and non-degenerate cells. The higher the proportion of degenerate cells, the more rapid the rate of subsequent culture degeneration.

The progressive changes in colonial morphology during the process of degeneration and the four different types of colonies for *C. acetobutylicum* ATCC 824 and ATCC 4259 that arise have been described previously (Woolley and Morris, 1990; Adler and Crow, 1987). An important finding from this study was that colonies that were intermediate

between degeneracy (low solvent producers) and non-degeneracy (high solvent producers), contained spores that were as heat-resistant as those of high solvent producers. The consequence of this observation was that an inoculum must be selected from the appropriate morphological colony type in order to obtain non-degenerate cells and the process of degeneration was step-wise and irreversible. Kashket and Cao (1995) linked similar specific morphological characteristics of *C. beijerinckii* colonies to their degenerate or non-degenerate status. They identified three morphological type colonies for *C. beijerinckii*, the first type had a spore containing center and a thin outgrowth consisting of degenerate, non-viable cells. the second type was similar to type one but did not show a degenerate outgrowth and the third type consisted entirely of degenerate cells. Both groups reported that degenerate cells are at a selective growth advantage to non-degenerate cells when growing in colonies on agar plates.

The physiological status of *C. beijerinckii* during the process of DNA transfer may therefore have a direct or indirect impact on the ability to obtain transconjugants due to the degeneracy within the colony.

3.2.2.2 Conjugation and Electroporation

Conjugation has previously been used to successfully transfer plasmids from bacterial donors to *C. beijerinckii* (Williams et al., 1990b; Oultram et al., 1987, 1988b). *C. beijerinckii* has also been transformed with plasmid constructs using electroporation (Oultram et al., 1988a; Lee et al., 1992; Blaschek and White, 1995).

C. beijerinckii is thought to lack an active restriction enzyme system and is therefore thought to be more amenable to DNA uptake than *C. acetobutylicum* NCP262 or ATCC824 (Richards et al., 1988). However there may be other systems or mechanisms present in *C. beijerinckii* that reduce the effectiveness of DNA transfer in this species. Schäfer et al. (1990) reported that they obtained a high conjugal frequency for vector transfer from *E. coli* to gram-positive coryneform bacteria using heat treatment. They incubated the *Corynebacterium glutamicum* recipients for 9 min at 48.5°C before the conjugation procedure and found that the normal conjugation frequency per donor cell from *E. coli* to *C. glutamicum* increased from 10^{-6}-10^{-7} to 10^{-2}-10^{-3} after heat treatment.

They also found that the enhancing effect of heat-treated recipient cultures declined within 5 hrs after heat treatment and that temperatures above 50°C led to lower transfer frequencies due to decreased cell viability. This suggested that heating blocks the restriction system of recipient cells either directly by inactivating temperature sensitive restriction endonucleases or indirectly through SOS repair as a consequence of DNA damage. This was also confirmed by high transfer frequencies to a restriction deficient mutant of *C. glutamicum*. The mechanisms of how an impairment of the restriction system leads to increased mating efficiency remains unclear since it is generally believed that single stranded transfer is insensitive to restriction in the recipient.

There is no report on the physiological or molecular effect of heat on *C. beijerinckii* cultures, however, clostridial strains, which are subjected to heat, low pH or other stresses undergo the heat shock/stress response (Bahl, 1993). In *C. acetobutylicum* DSM 792 a temperature upshift (30°C to 42°C) led to enhanced synthesis of heat shock proteins and decreased synthesis of other cellular proteins. We therefore sought to subject *C. beijerinckii* recipients to heat treatment prior to conjugation with *E. coli* donors in order to increase the conjugation frequency and concomitantly increase the pMTL30 suicide vector integration frequency.

The variability and reproducibility of the electroporation procedure for *C. beijerinckii* has been reported previously and factors contributing to this inconsistency remain unknown (Minton et al., 1993b). The three electroporation protocols that were available to transform *C. beijerinckii* were those of Oultram et al. (1988a) with modifications by Minton et al. (1993b), Lee et al. (1992), Kim and Blaschek (1993) and Blaschek and White (1995). The procedures of Minton et al. (1993b) and Lee et al. (1992) had a similar electroporation buffer composition, but differed with respect to the electroporation settings and the amount of DNA transformed per cell volume. The procedure of Kim and Blaschek (1993, 1995) differed not only with regard to the electroporation settings and in the amount of DNA transformed per cell volume but utilised a 10% PEG electroporation buffer solution. All three protocols gave similar transformation frequencies, which varied between 10^2-10^3 transformants/ug of DNA.

We therefore based our experimentation on the *C.beijerinckii* electroporation protocol of Minton et al. (1993b).

All the above factors were taken into consideration when attempting to obtain an efficient plasmid delivery system which would yield a high frequency of transformants. This would increase the chances of homologous recombination and of isolating a *regB* mutant.

3.3 MATERIALS AND METHODS

3.3.1 Bacterial strains, plasmids and culture conditions

All bacterial strains and plasmids are listed in Table 3.1

Cultures of *E. coli* were routinely maintained on Yeast Tryptone (YT) medium (Sambrook et al., 1989) at 37°C or 30°C according to requirements. The YT media was solidified with agar (1.5% w/v). Ampicillin (100μg/ml) and kanamycin (50μg/ml) were included in the media where appropriate.

The laboratory stock of *C. beijerinckii* NCIMB 8052 was maintained in distilled water at 4°C. *C. beijerinckii* was routinely cultured using Clostridial Basal Medium (CBM) (O' Brian and Morris, 1971), pH 6.3. CBM liquid medium was cooled to approximately 50°C after autoclaving and transferred to an anaerobic cabinet with a gas phase of 5% H_2, 10% CO_2 and 85% N_2 (Forma Scientifica Inc., Marietta, Ohio, USA). CBM agar plates, solidified with agar (1.5% w/v), were prepared under aerobic conditions and then pre-incubated under anaerobic conditions for at least 12 hrs before use. CBM broths were typically inoculated with 1/1000 volume of heat-shocked spores (70°C for 3 min) and cultured at 37°C or 30°C. Spore stocks of *C. beijerinckii*, were made by sub-culturing an overnight culture on either CBM or GCBM agar plates both of which contained 2% Starch and 2% Sucrose in order to obtain single colonies and then leaving the colonies for one week at 30°C till the spores were formed.

Growth defects associated with *ccpA* mutants were relieved by the addition of glutamate or glutamine (Faires et al., 1999). For the conjugation procedures specifically, the CBM (pH 6.3) was modified by the addition of glutamine (0.02M) after the medium was

cooled to approximately 50°C. The modified medium is hereafter referred to as GCBM. *C. beijerinckii* was maintained at 30°C in the anaerobic cabinet. Where appropriate, erythromycin (10μg/ml) was added to GCBM broths and agar plates after the media had cooled to 50°C.

Table 3.1: Bacterial strains and plasmids used in this study.

Name	Relevant Characteristics	Reference
Bacterial strains		
C. beijerinckii NCIMB 8052	Wild-type	NCIMB
E. coli CA474	*F⁻ Thi-1, _cI857, hsdS20(r^-_B, m^-_B), supE44 rpsL20(Sm^r) recA13*	Wilkinson and Young (1994)
E. coli JM105	*Δ(lac-proAB) lac^q _(lacZ)M15*	Yanisch-Perron et al. (1985)
E. coli JM109	*Δ(lac-proAB)lac^q _(lacZ)M15, recA1*	Yanisch-Perron et al. (1985)
E. coli JM110	*Δ(lac-proAB)lac^q _(lacZ)M15, dam, dcm*	Yanisch-Perron et al. (1985)
Plasmids		
pBluescriptSK+	Ap^r, T3 and T7 polymerase promoters	Stratagene, La Jolla, California
pUC19	Ap^r *lacZ*	Roche Diagnostics
pCTC1	Ap^r Em^r; Tra^- Mob^+	Williams et al. (1990a)
pMTL30	Ap^r Em^r; Tra^- Mob^+	Williams et al. (1990a)
R702	Kmr; Tra^+ Mob^+	Wilkinson and Young (1994)
pMTL500C	Ap^r Cm^r	Minton et al. (1990)
pBR322	Ap^r Tc^r	Bolivar et al. (1977)
pG^+host4	Ts derivative of pGK12, Em^r	Maguin et al. (1992)
pECORegB	~ 5.5 kb DNA fragment containing *regB* in pECOR251. Total size is ~ 8.7 to 8.8kb	See Chapter 2
pSKB2	~2.76 kb *Cla*I DNA fragment containing a 5' truncated *regB* gene in pBluescriptSK⁺	See Chapter 2
pSK2C	860 bp sequenced by Exonuclease III	Chapter 2
pQE0RegB	*regB* PCR product subcloned into the *BamH*I/*Pst*I site of pQE30	See Chapter 4.3.8
pSKRegB	2125 bp *Acc*I pECORegB fragment subcloned into pBluescriptSK⁺	See Chapter 4
pCTCRegB	*Xho*I/*Sac*I pSKRegB insert subcloned into	See Chapter 4.

	pCTC1	
pUC19RegB	*BamHI/PstI regB* fragment subcloned into pUC19	This study
pUCRegBCm	963 bp Cm^r gene subcloned into *SpeI* site of *regB* contained on pUC19RegB	This study
pMTL30RegBCm	1953 bp *BamHI/PstI* pUCRegBCm fragment subcloned into pMTL30	This study
pMTL30RegB	583 bp *XbaI/SpeI regB* pUC19RegB fragment subcloned into pMTL30	This study
pGAhost	2.297 kb band containing the pBR322 ori and Amp^r gene subcloned into pG+host4 *EcoRV/EcoRI* sites	This study
pSKnull	860 bp *KpnI/BglII* pSK2C insert subcloned into pSKB2 with the 1205 bp *KpnI/BamHI* fragment removed	This study
pGAhost1	380 bp *regB EcoRV/XmnI* fragment subcloned into *SmaI* of pGAhost	This study
pGAhost2	1822 bp *KpnI/StuI* pSKnull insert subcloned into pGAhost	This study

3.3.2 General DNA manipulation and extraction

See Chapter 2.3.2

3.3.3 Small-scale *C. beijerinckii* genomiç DNA extraction

C. beijerinckii 10ml CBM cultures were grown overnight and cells were harvested by centrifugation (5 min at 3000 μg [Hettich EBA 3S bench-top centrifuge]) in Sterilin™ tubes in the anaerobic unit. Cells were resuspended in a 200μl of CBM containing: 10% (w/v) sucrose; 12.5mM $MgCl_2$; 12.5mM $CaCl_2$ and 5mg/ml lysozyme (Roche Diagnostics). After incubation at 37°C for 1 hour, cells were analyzed microscopically for protoplast formation. If greater than 50% protoplast formation was seen, 20μl SDS and 50ul EDTA were added to a final concentration of 2%(w/v) and 100mM respectively. An equal volume of hot phenol (50°C), equilibrated with 0.1M Tris-HCl (pH 8.0), was added and mixed gently. The aqueous phase was recovered by centrifugation (10 minutes at 12000 rpm) and extracted, firstly with one volume of

chloroform: iso-amyl alcohol (24:1 v/v) and then with one volume of water-saturated ether. RNA was removed by adding 3µl Ribonuclease A (10mg/ml) (Sigma Chemical Company, St. Louis, Missouri). Chromosomal DNA was precipitated using isopropanol (Sambrook et al., 1989). The DNA pellet was resuspended in distilled water and stored at 4°C.

3.3.4 Southern blots

Non- radioactive DIG-labelled probes were prepared using the Roche Diagnostics DIG random primer labelling kit (catalog number: 1175033). A 380 bp *Eco*RV/*Xmn*I, or 188 bp *Xmn*I internal *regB* fragment, the pGAhost vector, the *erm* gene from plasmid pGAhost and the *erm* gene from plasmid pAMβ1 were used as probes.

The 380 bp *Eco*RV/*Xmn*I and 188 bp *Xmn*I *regB* probes were obtained after digestion of plasmid pSKA6 with the respective enzymes. The pAMβ1 erythromycin gene was obtained from plasmid pMU11328 (Achen et al., 1986). This plasmid was digested with *Ava*I and *EcoR*I to yield a ~2.4 kb fragment which contained the erythromycin gene. Plasmid pGAhost was digested with *Sac*I in order to obtain a ~ 1.180 kb fragment containing the *erm* gene.

For Southern blot analysis, DNA fragments were separated by agarose gel electrophoresis and capillary blotted onto positively charged Hybond N+ nylon membranes according to the method of Reed and Mann (1985). Hybridization and detection of DIG-labelled probes was conducted according to standard protocols supplied by Roche Diagnostics.

3.3.5 Conjugation

Stage 1: Inoculation

E. coli CA474 colonies containing plasmid R702 and transformed with the respective plasmid constructs were inoculated into 5ml YT broth containing kanamycin (50µg/ml) and ampicillin (100µg/ml). After overnight growth at 30°C, 0.5 ml of the respective cultures were inoculated into 50ml YT broth with the aforementioned antibiotics at the required concentration and left overnight at 30°C. Heat shocked *C. beijerinckii* spores

(70°C for 3 min) were inoculated in 150ml GCBM broths and grown overnight at 30°C in the anaerobic unit.

Stage 2: Conjugation

The conjugation procedure was carried out in duplicate once the OD_{600} reading of *E. coli* CA474 was in the range of 2 to 4 (~$5x10^{-9}$ cfu/ml) and the OD_{600} reading of *C. beijerinckii* was above 1.2 units. Conjugation was carried out at 30°C. For the heat treatment of recipients, 1ml of *C. beijerinckii* was subjected to a temperature of 50°C for 2 min in a heating block anaerobically, and then left at 30°C for 10 min before conjugation. The recipient to donor ratio was 1:10 (Leat, 1997) with 1ml of *C. beijerinckii* culture mixed with 10ml of the *E. coli* CA474 culture harbouring the respective plasmid construct. The recipient and donor were mixed in a Sterilin™ tube and centrifuged in the Hettich EBA 3S bench-top centrifuge at 3000 xg for 6 min. The supernatant was poured off and the pellet resuspended gently in 100µl GCBM with a plastic inoculation loop. Bacteria were then cultured on GCBM agar plates containing no antibiotics and left at 30°C overnight.

Stage 3: Selection

The following day, 0.8ml of sterile water was placed on each of the conjugation plates and the growth gently scraped off with a glass spreader. The respective conjugation mixes, (~100µl), were cultured on GCBM agar plates containing erythromycim (10µg/ml) and left at 42°C in the anaerobic unit. Overnight anaerobic incubation at 42°C facilitated counter-selection of the donor strain, while the presence of erythromycin facilitated the selection of *C. beijerinckii* transconjugants. Typically, transconjugants took 3-4 days to appear. The conjugation procedure used plasmid pCTC1 and plasmid pMTL30 as the positive and negative controls respectively. For the conjugation frequency, dilutions (10^{-2} to 10^{-8}) were prepared of the respective conjugation mixes. The dilutions were cultured on GCBM agar plates with no antibiotics at 42°C in the anaerobic cabinet and aerobically at 30°C on YT agar plates containing Kan and Ap. The conjugation frequency was expressed as transconjugants per donor.

3.3.6 Electroporation

The electroporation protocol was essentially similar to the one reported by Oultram et.al. (1988a) with the incorporation of some of the modifications recommended by Minton et al. (1993b). Sucrose was substituted with melibiose as the osmotic stabilizer

and 0.3ml to 0.4ml of cells were electroporated in 0.2-cm cuvettes. The electroporation settings for the Bio-Rad Gene Pulser (Richmond, CA) were: 1.25 kV field strength, 25 μF capacitance and 100 to 1000 ohms resistance. Plasmid pCTC1 was used as the positive control since it is known to replicate stably in *C. beijerinckii*. Induction of the Em^r gene was achieved by adding 5μg/ml of erythromycin to the expression mix after 1 h and then incubating for a further 2 h after which it was sub-cultured on GCBM agar plates containing Em. Typically, transformants took 3-4 days to appear.

3.3.7 Construction of suicide vectors based on pMTL vector series

Suicide plasmids were constructed based on vector pMTL30 for the disruption of the *regB* gene. Plasmid pQE0RegB contains *regB* on a *BamH*I/*Pst*I fragment subcloned into the protein expression vector pQE30 (Figure 3.2; see Chapter 4 for further details). The *regB* gene was released from the pQE0RegB plasmid by digestion with *BamH*I and *Pst*I, and subsequently subcloned into plasmid pUC19, which was also previously digested with *BamH*I and *Pst*I. The resulting construct was called pUC19RegB.

A 963 bp fragment encoding a chloramphenicol gene was excised from plasmid pMTL500C (Minton et al., 1990) and then the plasmid was digested with *BamH*I, the restriction sites were filled-in, and then digested with *Stu*I (blunt). Plasmid pUC19RegB was linearised with *Spe*I, which cuts once in the *regB* gene (position 1427 bp) and the restriction sites filled-in. The chloramphenicol gene was then ligated into the blunt *Spe*I site of the pUC19RegB construct. The resultant plasmid was called pUCRegBCm. This construct was then digested with *BamH*I and *Pst*I which released a ~1953 bp fragment containing the *regB* gene interrupted by the chloramphenicol gene. The insert was subcloned into pMTL30 previously digested with *BamH*I and *Pst*I and the resulting construct was called pMTL30RegBCm.

Since *Spe*I and *Xba*I restriction sites are compatible, the internal 583 bp *Xba*I/*Spe*I *regB* fragment of pUC19RegB was subcloned into pMTL30 digested with *Xba*I to form pMTL30RegB.

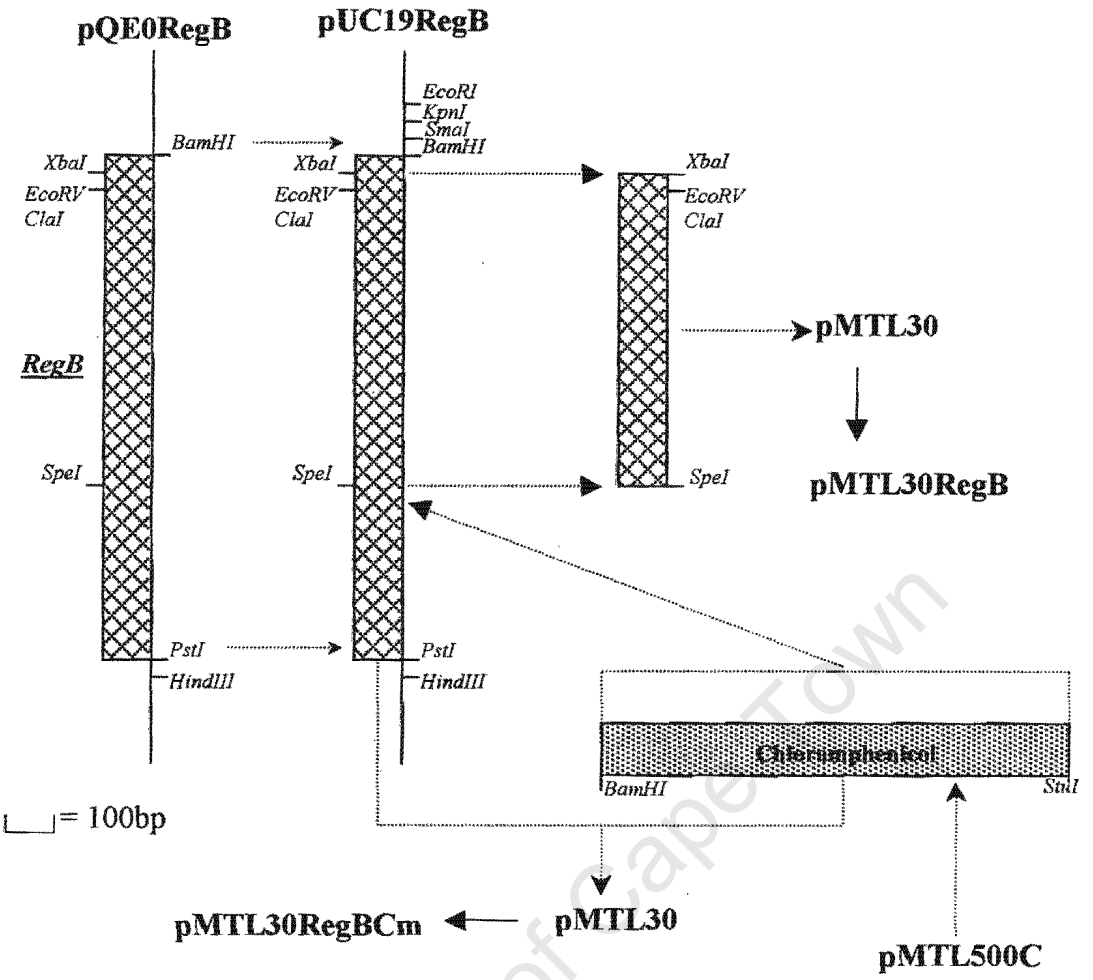

Figure 3.2: Construction of suicide vectors pMTL30RegB and pMTL30RegBCm. The *Bam*HI/*Pst*I fragment containing the *regB* gene from plasmid pQE0RegB was subcloned in pUC19 and the resultant plasmid pUC19RegB was used as the basis for the construction of suicide vectors in pMTL30. The internal *Xba*I/*Spe*I RegB fragment was subcloned into the *Xba*I site of pMTL30 to form pMTL30RegB. For vector pMTL30RegBCm, the chloramphenicol gene from pMTL500C was firstly subcloned into the *Spe*I site of the *regB* gene on pUC19RegB and then the resulting interrupted *regB* gene was subcloned into pMTL30. Refer to the text for further details.

For the construction of vector pCTCRegB refer to Chapter 4.

3.3.8 Construction of suicide vectors based on vector pG$^+$host and development of a system for sco/dco integration into the *C. beijerinckii* genome.

Vector pG$^+$host4 (Biswas et al., 1993) was used to construct a shuttle plasmid for *C. beijerinckii* and *E. coli* to facilitate future work. This vector lacked a gram-negative *ori* and an antibiotic marker for replication, selection and maintenance in *E. coli*. Vector

pG$^+$host4 was digested with *Eco*RV (blunt) and *Eco*RI (in the mcs). Plasmid pBR322 was digested with *Pvu*II (blunt) and *Eco*RI to release a ~2.297 kb band containing the *ori* and ampicillin gene. This insert was subcloned into the previously digested (*Eco*RV/*Eco*RI) pG$^+$host4 vector and selection was based on ampicillin resistant *E. coli* colonies. This new vector was subsequently renamed pGAhost.

Two plasmids containing the *C. beijerinckii regB* gene were constructed from vector pGAhost. The first vector was constructed for a 'sco' recombination event and the second vector for a 'dco' recombination event. The ~380 bp *regB Eco*RV/*Xmn*I fragment which was previously filled-in was ligated into the *Sma*I site of pGAhost and the resulting plasmid was named pGAhost1.

For the second *regB* suicide plasmid, the following constructs were made (Figure 3.3). Plasmid pSKB2 (see Chapter 2.4.2) was digested with *Kpn*I and *Bgl*II to release a ~ 1205 bp band, containing the truncated 5' *regB* gene. The remaining 4518 bp plasmid contained 1555 bp of downstream *regB* region from *Bgl*II to *Cla*I (refer to Figure 2.5). Plasmid pSK2C is a pBluescriptSK$^+$ construct containing 860 bp of the upstream region of the *regB* gene as a result of ExonucleaseIII shortening of pSKB2. Only 75 bp of the 5' region of the *regB* gene remains on this plasmid. This plasmid was digested with *Kpn*I and *BamH*I, and the insert was subsequently subcloned into the previously digested vector pSKB2 to form pSKnull. This construct contains part of the upstream and downstream regions flanking the *regB* gene, and excluding most of the *regB* gene itself except for 75 bp near the 5' end, on a ~ 2.4 kb insert. Vector pGAhost was digested with *Sal*I, the restriction site filled-in and then digested with *Kpn*I. Vector pSKnull was digested with *Kpn*I and *Stu*I (blunt) and the 1822 bp insert subcloned into the previously digested pGAhost to form pGAhost2.

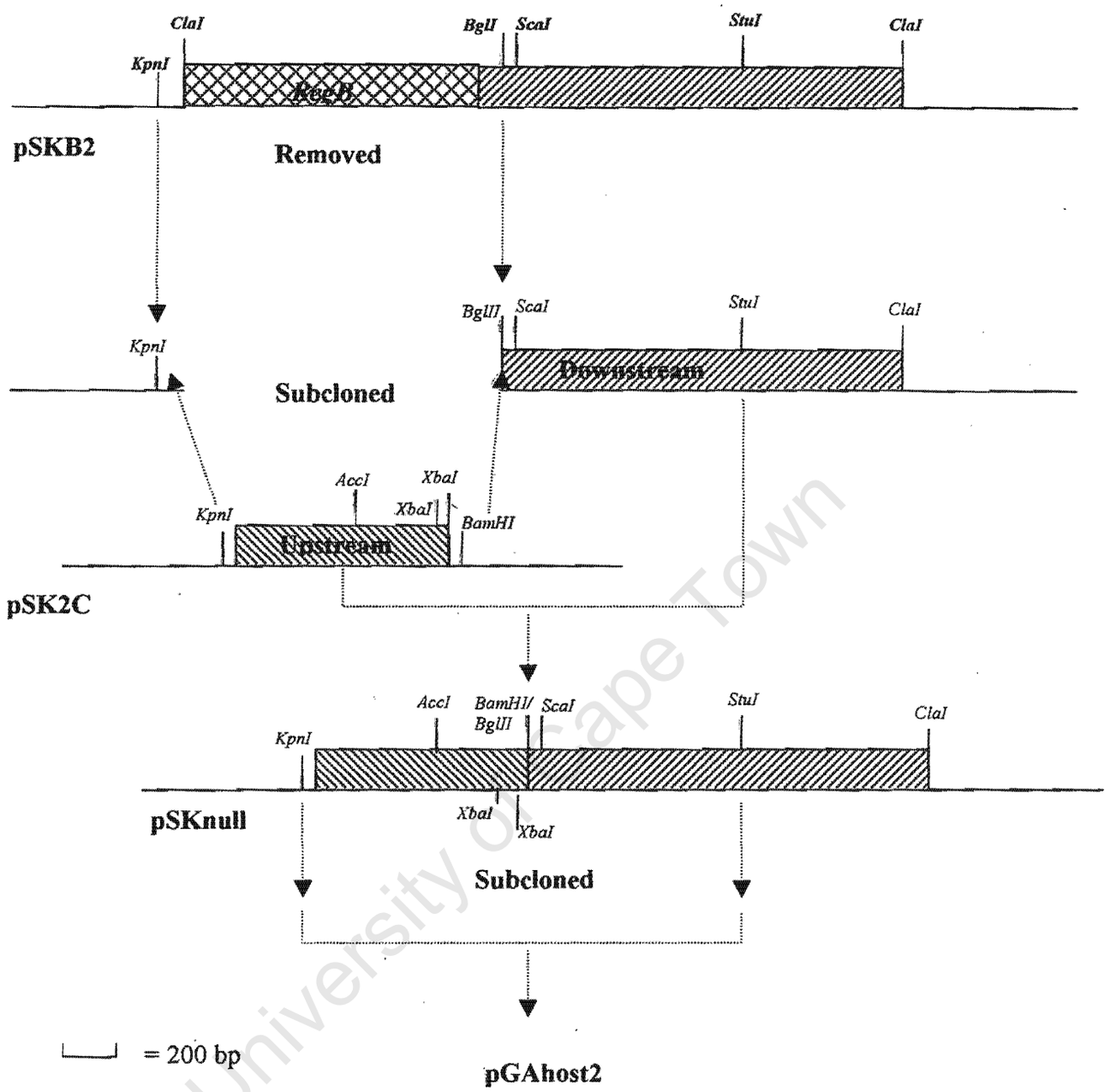

Figure 3.3: Construction of vector pGAhost2. Plasmid pSCB2 was digested with *Kpn*I and *Bgl*II to release a fragment (cross-hatch) which contained the truncated N-terminal *regB* gene and leaving only the downstream region of the *regB* gene. This fragment was replaced with the upstream region of the *regB* gene by digestion of plasmid pSK2C with *Kpn*I and *BamH*I to form pSKnull. Plasmid pSKnull was digested with *Kpn*I and *Stu*I and the resulting fragment subcloned into pGAhost. See text for further details.

In order to obtain sco and dco *C. beijerinckii* recombinants for pGAhost2, we followed a similar protocol to that described by Biswas et al. (1993):

Sco integration into the *C. beijerinckii* chromosome:

C. beijerinckii colonies containing pGAhost2 were grown overnight at 30°C in the presence of erythromycin and then diluted 100-fold in GCBM and grown at 30°C for ~ 2.5 h. Cultures were then shifted to 42°C for 3 hrs in order stop plasmid replication. Samples were diluted and cultured on CBM agar plates at 42°C to select for Em^r integrants.

Dco integration into the *C. beijerinckii* chromosome:

The *C. beijerinckii* colonies obtained at 42°C for sco integration were grown overnight with erythromycin at the same temperature and then diluted ($1:10^5$) in CBM. This inoculum was grown overnight without selection at 30°C to allow stimulation of recombination by plasmid replication. The overnight culture was sub-cultured onto CBM agar plates with and without erythromycin and colonies were screened for Em sensitivity.

3.4 RESULTS AND DISCUSSION

3.4.1 Construction of Vectors

A pG^+host-based shuttle vector, which could be maintained in *E. coli*, was constructed for gene inactivation in *C. beijerinckii* (Figure 3.4). This plasmid with its temperature-sensitive replicon was designated pGAhost, and was subsequently used to construct two plasmids, pGAhost1 and pGAhost2 (Figure 3.5). Plasmid pGAhost1 contains an internal 380 bp *EcoR*V/*Xmn*I fragment from the *regB* gene, and pGAhost2 contains a chimeric fragment consisting of the flanking upstream and downstream regions of the *regB* gene joined together. The internal *regB* fragment on plasmid pGAhost1 allows for a sco recombination event while the regions flanking the *regB* gene on plasmid pGAhost2 allow for a dco recombination event with the *C. beijerinckii* chromosome.

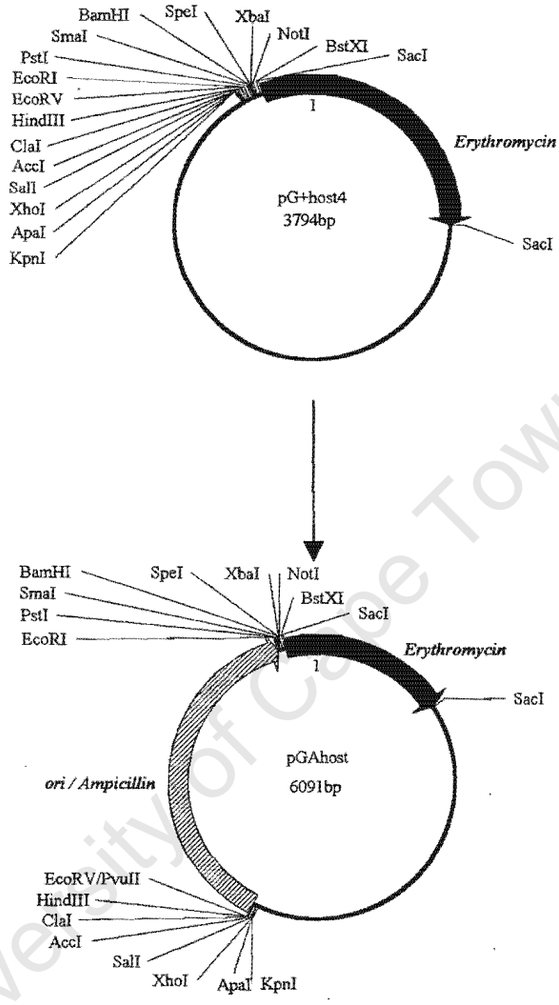

Figure 3.4: Plasmid map of pGAhost. Vector pGAhost was constructed by subcloning a *Pvu*II/*EcoR*I DNA fragment from plasmid pBR322 containing the pBR322 *ori* and ampicillin gene into the *EcoR*V/*EcoR*I site of pG+host4. See section 3.3.8 for further details.

Two suicide vectors based on plasmid pMTL30 were constructed namely pMTL30RegB and pMTL30RegBCm. Plasmid pMTL30RegB contains a 583 bp *Xba*I/*Spe*I fragment internal to the *regB* gene, and pMTL30RegBCm contains a ~1953 fragment consisting of the *regB* gene interrupted by the Cm gene from pMTL500C (Figure 3.5).

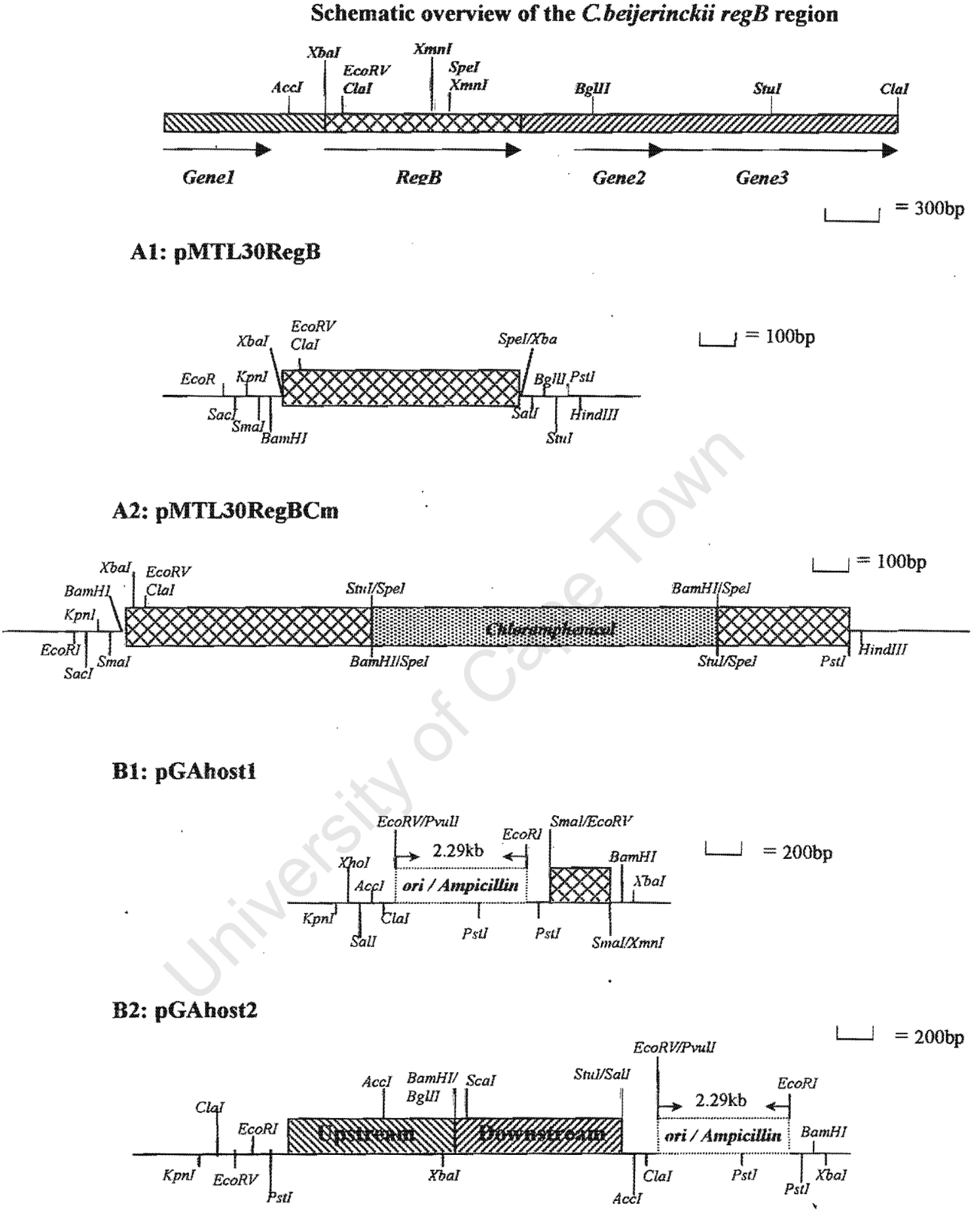

Figure 3.5: Schematic representation of the *C. beijerinckii regB* region including suicide vectors pMTL30RegB, pMTL30RegBCm, pGAhost1 and pGAhost2. A1- Plasmid pMTL30RegB contained an internal *XbaI*/*SpeI* fragment (cross-hatch pattern) of the *regB* gene cloned into the pMTL30 *XbaI* site.

The *Xba*I/*Spe*I sites are compatible but cannot be recut with either enzyme. A2- Plasmid pMTL30RegBCm contained a blunt-ended Cm^r marker from plasmid pMTL500C cloned into the *Spe*I site (filled-in) of the *regB* gene. Restriction sites (*Stu*I/*Spe*I and *BamH*I/*Spe*I) are indicated on both sides of the blunt-ended Cm^r insert since it could have ligated in either orientation in the restriction site. B1- Plasmid pGAhost1 contained a blunt-ended *EcoR*V/*Xmn*I internal fragment of the *regB* gene (cross-hatch pattern) cloned into the *Sma*I site of pGAhost. B2- Plasmid pGAhost2 contained a *Kpn*I/*Stu*I insert from pSKnull cloned into the *Kpn*I and *Sal*I (filled-in) restriction sites of vector pGAhost. This insert contains the upstream and downstream regions flanking the *regB* gene but excludes most of the *regB* gene.

3.4.2 Key factors affecting Conjugation and Electroporation procedures

3.4.2.1 Conjugation and Electroporation frequencies

Exhaustive conjugation experiments using *E. coli* CA474 containing plasmid R702 as the donor and pCTC1 as the mobilizable plasmid, resulted in conjugation frequencies of between 10^{-6} to 10^{-7} transconjugants per donor (Table 3.2). This agrees with frequencies previously reported by Williams et al., (1990b), but was extremely variable even under constant experimental conditions. Plasmid pMTL30 is unable to replicate in *C. beijerinckii*, which is reflected in Table 3.2. Jennert et al. (2000) also observed with *C. cellulolyticum* that the optimized conjugation procedure, which was adapted from Williams et al. (1990a), was not always reproducible even when all the conditions were repeated identically.

Attempts were made to increase the conjugation frequency by using heat treatment in order to concomitantly increase the recombination frequency. Counter-selection of *E. coli* CA474 was normally carried out at 42°C for 2-3 days, so it was known that *C. beijerinckii* was viable at this temperature. A preliminary test indicated that temperatures above 50°C were lethal for *C. beijerinckii*, but that cell viability was not dramatically affected when *C. beijerinckii* was kept at 50°C for 2 min (results not shown). Recipient cultures were therefore heated at 50°C for 2 min, prior to conjugation with the *E. coli* donor. Only a 50-fold increase in the conjugation frequencies of heat-treated cells was observed (Table 3.2).

Table 3.2: The effect of heat treatment on the frequency of plasmid transfer to *C. beijerinckii*

Plasmid	Mobilization	Frequency per donor (30°C)*	Frequency per donor (Heat treatment)*
pCTC1	R702	1.535×10^{-0} [4]	8.352×10^{-5} [4]
pMTL30	R702	$< 3.1 \times 10^{-9}$ [4]$^{\pm}$	$<3.1 \times 10^{-9}$ [4]$^{\pm}$

* The number of determinations is shown in square brackets. Each experiment was done in duplicate while the values represent the average of all the experiments. Numerous conjugation experiments were carried out but only four conjugation frequencies are given since most gave similar values to the average.

$^{\pm}$ Values obtained represents the limit of detection, no transconjugants were obtained.

The increase obtained in the conjugation frequency using heat treatment was not regarded as a significant difference since Wilkinson and Young (1994) obtained a frequency per recipient of 1.2×10^{-4} with plasmid pCTC1. In those individual experiments where the heat treatment on *C. beijerinckii* failed or there was a decrease in the conjugation frequency, it was difficult to establish if this was due to the variability within samples, strain degeneration or whether heat had a periodic negative effect on the *C. beijerinckii* physiology and growth phase. Heat treatment was however occasionally employed in addition to the normal procedure in some of the conjugation experiments. Heat treatment of *C. beijerinckii* may improve the conjugation frequency but optimization of the heat treatment is required. This requires the assessment of the effects of heat on the growth phases and the degeneracy state of the culture of *C. beijerinckii*, which was beyond the scope of this study. Overall, conjugation produced more consistent results than the electroporation procedure in terms of the reproducibility of DNA transfer.

The electroporation procedure of Minton et al. (1993b) was employed with minor modifications. The plasmid DNA was however prepared from *E. coli* JM109 or JM110, unlike Minton et al. (1990) who found that plasmid DNA prepared from *B. subtilis* gave higher transformation frequencies. Despite the modifications of Minton et al. (1993b), it was still found that the results of electroporation for *C. beijerinckii* remained inconsistent.

Repetition of the electroporation procedure using the same components and parameters failed to give reproducible results. We obtained no transformants at the reported time

constants of 4.7 ms to 6.9 ms (Oultram et al. 1988a) but obtained transformants at a low frequency at transformation values below those reported (Table 3.3; [B], [C]). In our hands, we observed no consistent correlation between the resistance setting and the time constant. For example, a time constant of 3.3ms and 8.1ms (Table 3.3 [D]) were obtained independently upon electroporation of pGAhost into *C. beijerinckii* using the same conditions and resistance setting (1000Ω). Thus neither time constants, nor any other parameter could be used as an indication of a successful electroporation procedure.

Table 3.3: The electroporation of *C. beijerinckii* with plasmid constructs under different field strengths.

Field Strength (kV)[a]	Time Constants[b] (ms)	Transformants per electroporation
[A] 1.25 (3)	2.7- 4.8	None
[B] 1.5 (3)	3.1-3.5	1-2
[C] 2.0 (4)	2.5-3.8	1-2
[D] 2.5 (5)	3.3-8.1	None

[a]The number of electroporation experiments for each voltage is given in brackets.

[b]In each electroporation experiment, 5 µg of plasmid DNA was mixed with 0.4 ml of *C. beijerinckii* culture (OD_{600} ~ 0.45-0.6) and electroporated using 0.2 cm electroporation cuvettes with resistance in the range of 200Ω – 1000Ω. Only the lowest and the highest time constant values are indicated in order to give the range of values obtained.

In those electroporation experiments that did produce transformants, the transformation frequency was extremely low. The reported electro-transformation frequencies for plasmid pMTL500E prepared from an *E. coli* host were in the range of 10^2 to 10^3 tranformants/µg of DNA (Oultram et al., 1988a). We obtained an average frequency of 10^4 transformants/µg of DNA with plasmid pCTC1, which is a derivative of plasmid pMTL500E. We found that the addition of erythromycin (5ug/ml) to the expression mix, after first incubating for one hour with no selection, increased the number of transformants. This enrichment step helped by conferring a selective advantage to transformants over non-transformed cells. The methylation/restriction endonuclease deficient *E. coli* strain, JM110, was used for the propagation, maintenance and DNA source of the plasmid constructs. Unlike *C. acetobutylicum* ATCC 824, which possesses

a powerful restriction endonuclease system (Mermelstein and Papoutsakis, 1993a), *C. beijerinckii* is reported to lack a restriction endonuclease system (Richards et al., 1988). Thus these factors are unlikely to be the major contributing reasons for the low electro-transformation frequency. However, Minton et al. (1990) did report that electroporation efficiency increased with DNA prepared from *B. subtilis*. The variability and inconsistency in the procedure made it difficult to assess the contribution, if any, of factors inherent in the construction of the pG^+host suicide vectors which could affect the transformation frequency. Although the growth and state of degeneration of *C. beijerinckii* cultures could explain the inconsistency of the electroporation procedure, it cannot explain the poor electro-transformation frequency and there are most likely other unknown factors, which may compromise vector integrity in the host or recipient.

3.4.2.2 Strain degeneration and growth requirements

During the preparation of recipient cultures for both the conjugation and the electroporation procedures, *C. beijerinckii* would periodically undergo a physiological condition during the exponential growth phase in which cell growth appeared stationary, taking many hours to reach the next growth stage, or alternatively cell viability would decrease as indicated by a decrease in the optical density of the culture. When *C. beijerinckii* cultures displaying this growth pattern were left in this condition, growth resumed only after 24 hrs. This physiological condition occurred during the exponential growth phase at OD_{600} readings lower than 1.2 units and conjugation or electroporation procedures using these cultures were unsuccessful.

Strain lysis seems the most likely reason for these growth patterns exhibited by *C. beijerinckii*, since there was a decrease in cell viability during the exponential growth phase and in some cases the condition was reversible. We attempted to circumvent this physiological condition in the conjugation procedure by growing *C. beijerinckii* at a lower temperature (30°C) and initiating the conjugation when the OD_{600} of the culture was above 1.2 units which is near the end of the exponential phase (Kashket and Cao, 1995). The electroporation procedure also used *C. beijerinckii* grown at 30°C but cultures above OD_{600} 0.6 units could not be electroporated successfully. These

alterations to the procedures did not remove strain degeneration entirely but helped reduce the frequency of the condition.

An important part of any gene inactivation procedure is the selection requirement for the resulting mutant. The selection requirement is in turn dependent on the effect that the mutation has on the overall status of the bacterium. CcpA has been shown to affect growth rate and sporulation in gram-positive bacteria due to its global regulatory role in carbon, energy and nitrogen metabolism (Faires et al., 1999, Grundy et al., 1994, Miwa et al, 1994). Inactivation of the *ccpA* gene in *B. subtilis*, *B. megaterium*, *L. lactis* and *Lactobacillus plantarum* lead to a marked decrease in growth rate on minimal medium with various carbon sources (Küster et al., 1999a; Muscariello et al., 2001). The poor growth was a result of the inability of *ccpA* mutants to assimilate ammonia as the nitrogen source, and CcpA was found to control the expression of the *gltAB* operon, which encodes glutamate synthase. The growth defect was relieved by the addition of glutamate or glutamine with as little as 0.5 mM glutamate sufficient to restore growth.

Lastly, we could not rule out the possibility that the creation of a *regB* mutant could also affect nitrogen assimilation and thereby severely impair the growth rate of *C. beijerinckii*. *C. beijerinckii* prefers an organic nitrogen source in the form of casamino acids to inorganic nitrogen in the form of ammonium (Quixley, 1999). Although CBM contains casamino acids as a nitrogen source, we included glutamine (20mM) in the medium to help reduce possible growth defects. We chose glutamine as the organic source since we found that glutamate was inhibitory to *C. beijerinckii* growth when included in the medium (results not shown).

3.4.3 Inactivation of *regB* using the pMTL30 suicide system

Conjugation experiments using *E. coli* CA474 (pR702) as the donor (no heat treatment was employed) produced a single Em^r *C. beijerinckii* transconjugant for construct pMTL30RegB and none for construct pMTLRegBCm. Plasmids pCTCI and pCTCRegB were conjugated into *C. beijerinckii* as controls at the previously reported frequencies (see Table 3.2) in the same experiment.

As expected for the positive control (refer to Figure 3.6), hybridization occurred against plasmid pECORegB (A: lane 3) at ~ 8.76 kb (refer to Table 3.1) while there was no hybridization signal for plasmid pMTL30 (A: lane 2) when both were digested with *Xba*I and probed with the 188 bp *Xmn*I *regB* probe. DNA hybridization analysis of the chromosomal DNA from the putative *C. beijerinckii regB* mutant using the *regB Xmn*I probe revealed an increase in fragment size compared to the wild-type which is indicative of a recombination event between the suicide vector pMTL30RegB and the chromosomal *regB* gene (Figure 3.6 A). However DNA from a contaminant organism whose genome happens to contain a regB homologue cannot be ruled out although morphological analysis indicated that the bacterial culture belonged to the clostridia. For *C. beijerinckii* wt or *C. beijerinckii* (pCTCRegB), where the plasmid is maintained in the cytoplasm, chromosomal digestion with *Xba*I/*Spe*I resulted in expected hybridization signals at ~ 0.584 kb corresponding to the intact wild-type *regB* gene (refer to Figure 3.2). Digestion of the putative *C. beijerinckii regB* mutant with the same enzymes resulted in a ~ 16.0 kb hybridization signal with no signal at ~ 0.584 kb. Restriction analysis indicated that there were only two *Xba*I sites (positions 759 and 844) and a single *Spe*I site (position 1428) on the 3623 bp of *C. beijerinckii* chromosomal DNA flanking the *regB* gene (refer to Figure 2.4). Integration of the ~ 5.084 kb pMTL30RegB construct into the corresponding region on the chromosome would result in an increase in fragment size detected by the *regB* gene probe. A single copy of the plasmid integrated into the *regB* gene would be predicted to yield a fragment of 5.084 kb. However, although a band of 0.584 kb corresponding to the wild-type *regB* gene was missing in the putative mutant, the only band detected by hybridization was ~ 16 kb. It is possible that pMTL30RegB construct was established in the *C. beijerinckii* chromosome in multiple copies either by multiple integration of the construct or by integration of a HMW plasmid construct.

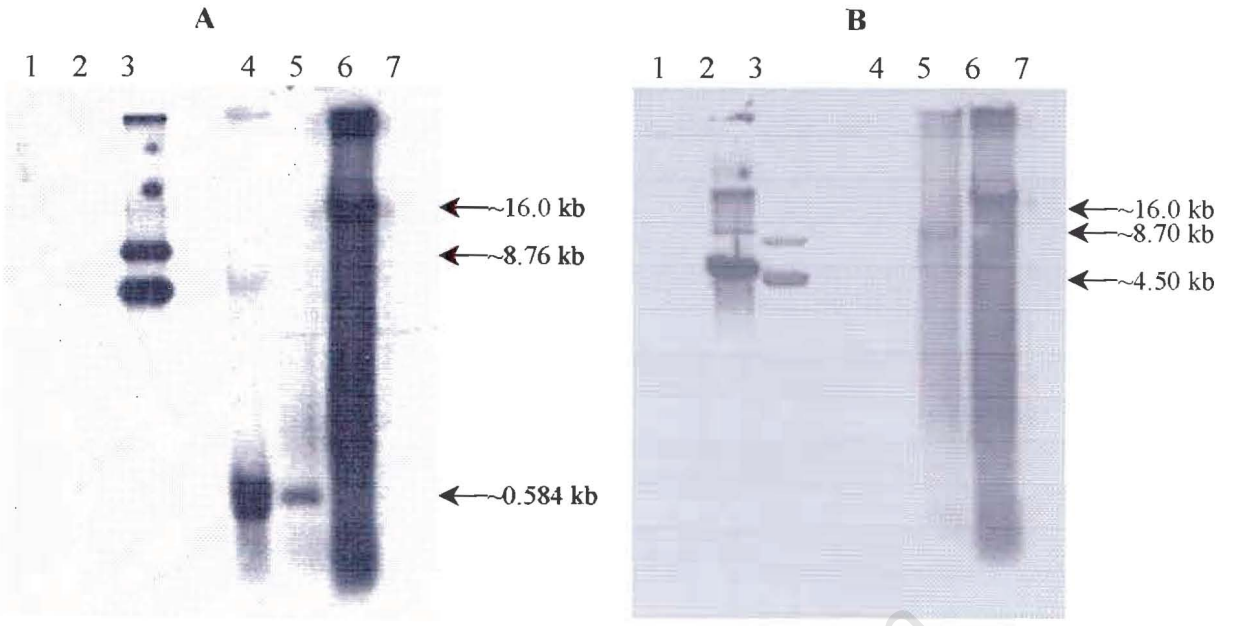

Figure 3.6: Southern hybridization analysis of wild-type *C. beijerinckii*, *C. beijerinckii* (pCTCRegB) and a putative *C. beijerinckii regB* mutant. Chromosomal DNA (50μg) from *C. beijerinckii* wt, *C. beijerinckii* (pCTCRegB) and putative *C. beijerinckii regB* mutant were digested with *Xba*I/*Spe*I. All DNA samples were hybridized respectively with the 188 bp *Xmn*I *regB* probe (A) and the pAMβ1 *erm* gene (B).

Lane 2: Plasmid pMTL30 digested with *Xba*I (positive control for B)

Lane 3: Plasmid pECORegB digested with *Xba*I (postive control for A).

Lane 4: *C. beijerinckii* wt (positive control for A, negative control for B).

Lane 5: *C. beijerinckii* (pCTCRegB).

Lane 6: Putative *C. beijerinckii regB* mutant.

Lanes 1 and 7: Lambda marker (λ DNA digested with *Pst*I).

A single ~ 16 kb band was obtained in the putative mutant suggesting that 3 copies of the plasmid had been integrated into the *C. beijerinckii* chromosome. Although the 0.584 kb *Xba*I/*Spe*I *regB* insert was ligated into the *Xba*I site of pMTL30 resulting in construct pMTL30RegB, it could be recleaved by *Xba*I at the *Xba*I restriction site but by neither *Xba*I nor *Spe*I at the hybrid *Xba*I/*Spe*I site (see Appendix A). If we assume that multiple integration of pMTL30RegB (~5.084 kb) had occurred and multiple copies became established, then digestion with *Xba*I and *Spe*I should yield a ~5.084 kb band (plasmids released from the chromosome) and a 0.584 kb band (right junction fragment), (see Figure 3.7). However, the internal *Xba*I site does not appear to be recognized by the restriction enzyme, resulting in a 16 kb fragment.

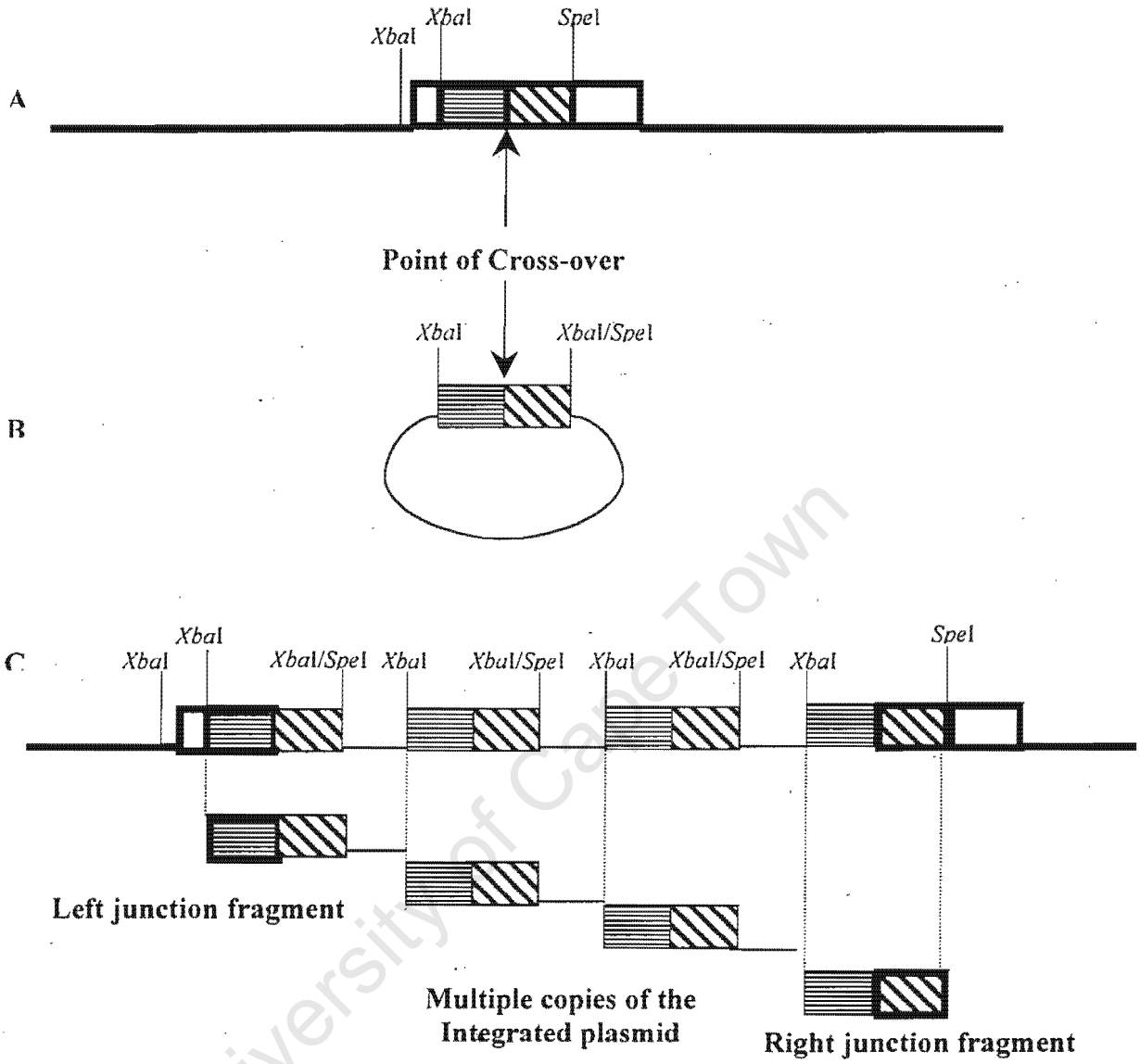

Figure 3.7: Schematic representation of pMTL30RegB integrated into the *C. beijerinckii* genome. Thick lines represent the *C. beijerinckii* chromosome and thin lines represent pMTL30 DNA. The 'target' DNA carried by pMTL30 and its equivalent on the *C. beijerinckii* genome are represented by squares with a designated pattern. [A], The *regB* gene is shown with the DNA region to be disrupted and key restriction enzyme sites. [B], Construct pMTL30RegB. The point of cross-over for the Campbell-like recombination event is indicated by an arrow. [C], The result of multiple insertions into the genome of *C. beijerinckii*.

Wilkinson and Young (1994) reported establishment of single integrated plasmids and amplification of integrated plasmids, which would lead to different results upon

Southern hybridization analysis. Leat (1997) found upon disruption of the *C. beijerinckii scrB* and *scrR* genes, the presence of multiple tandem copies of plasmid pMTL30 integrated into the respective genes.

In order to verify that plasmid pMTL30RegB had integrated into the *C. beijerinckii* genome and thus confirm that we had created a *C. beijerinckii regB* mutant, we re-probed the Southern blot with the pAMβ1 *erm* gene. The *erm* gene probe hybridized to a ~ 16.0 kb DNA fragment from the putative *C. beijerinckii regB* mutant corresponding to the fragment detected by the *regB* probe (Figure 3.6). This indicated that the plasmid had integrated into the chromosome in the same region as the inactivated *regB* gene. Wild-type *C. beijerinckii* gave no signal and *C. beijerinckii* (pCTCRegB) gave the expected hybridization signal at ~ 8.7 kb, corresponding to the plasmid construct (B: lane 4 & 5). Similarly, an expected hybridization signal was obtained at ~ 4.5 kb for plasmid pMTL30 when probed with the pAMβ1 *erm* gene (B: lane 2). The hybridization signal observed in lane 3 (B) when probed with the *erm* gene is most likely due to DNA entanglement or probe contamination with plasmid DNA.

It cannot be readily explained why a single hybridization signal for the integration event was obtained. It is possible that upon integration of the suicide vector into the chromosome, the *Xba*I site was damaged or altered. Alternatively the suicide vector could have assumed a multimeric form in which the *Xba*I site had been modified before the integration event. If this had indeed occurred, then it could explain the ~ 16 kb hybridization signal since only the *Xba*I site upstream of the *regB* gene and the *Spe*I site at the right junction fragment (refer to Figure 3.7) would be recleavable, releasing a single fragment containing multiple copies of pMTL30RegB plasmid. It was verified that the *Xba*I site of the original pMTL30RegB plasmid maintained in *E. coli* was recleavable (Appendix A) and therefore any changes to pMTL30RegB or the *Xba*I site must have occurred during or after the conjugation and integration events.

The above Southern hybridization analysis suggested that construct pMTL30RegB had integrated into the *regB* gene, thereby creating a *C. beijerinckii regB* mutant and that the construct was possibly established in multiple copies on the chromosome. The *C.*

beijerinckii regB mutant was then subjected to sporulation in order to make spore stocks for further characterization. However, upon germination of the spores, it was found that the *C. beijerinckii regB* mutant had reverted to wild-type. PCR analysis was done on chromosomal DNA extracted from the germinated spore stock using primers fRegB31 and rRegB (see 4.3.8), which flanked the *regB* gene. The same *regB* fragment size (~ 990 bp) for the *C. beijerinckii regB* mutant as wild-type *C. beijerinckii* was obtained, and not the increase in size of the DNA fragment expected for a disrupted *regB* gene (results not shown). Numerous attempts at repetition of the conjugation procedure using the same pMTL30 *regB* construct and exhaustive PCR analysis of the spore stock failed to re-produce the *C. beijerinckii regB* mutant.

It has been reported previously that inactivation of the *ccpA* gene in *E. faecalis*, *B. subtilis* and *B. megaterium* led to a distinct reduction in growth rate (Leboeuf et al., 2000; Faires et al., 1999; Miwa et al., 1994; Hueck et al., 1995). This growth defect was less pronounced in complex medium and it has been shown that *ccpA* mutants require the presence of intermediates of the tricarboxylic acid cycle and a nitrogen source for growth on minimal medium (Wray et al., 1994). Furthermore, there is evidence that growth control and *cre*-binding of the CcpA protein are independent properties (Küster et al., 1999a). This has led to the suggestion that the CcpA protein is a central regulator of carbon, energy and nitrogen metabolism in gram-positive bacteria with multiple functions in cellular physiology (Stülke and Hillen, 2000). In *C. beijerinckii*, the growth rate is linked to acidification, the switch from acidogenesis to solventogenesis and subsequent sporulation. Because of the proposed global role of RegB as a CcpA homologue, disruption of *regB* may trigger strain degeneration or compromise cell viability by directly or indirectly affecting the growth rate, growth phase, energy requirements or sporulation. The selective pressure to revert to wild-type would therefore be considerable.

Lastly, Campbell-like integration events generate a duplication of target sequences and a reversal of the integration event is possible. It was found that in the absence of erythromycin selection, erythromycin-sensitive recombinants segregated at a frequency of between 3.7×10^{-4} to 1.3×10^{-3} per generation (Wilkinson and Young, 1994).

It was possible that factors within the conjugation procedure were responsible for the failure to obtain *C. beijerinckii regB* mutants. However using the pMTL30 based construct pCBMU1, the inactivation of the sucrose hydrolase gene, *scrB*, was again repeated. The *scrB* gene was previously inactivated by Leat (1997) who obtained an establishment frequency per recipient of $1.2x10^{-10}$. Repetition of the conjugation procedure as described previously (without heat treatment) with vector pCBMU1 resulted in 7 transconjugants being obtained representing an establishment frequency per recipient of $3.846x10^{-8}$. Thus factors within the conjugation procedure were not responsible for our failure to obtain a *C. beijerinckii regB* mutant. Both Leat (1997) and our non-replicative plasmid establishment frequencies were lower than the reported frequencies (10^{-6} to 10^{-7}) found by Wilkinson and Young (1994). Leat (1997) suggested that this could be due to the fact that conjugation was conducted on CBM agar plates as opposed to being conducted on cellulose filters on CBM agar plates. Though the technique we employed may be more convenient it may lead to an overall decrease in conjugation efficiency. This result further suggested that the failure to obtain *C. beijerinckii regB* mutants could have been because inactivation of the *regB* gene contributed negatively to cell viability or was a lethal event for *C. beijerinckii*. The selective pressure to revert to wild-type by a reversal of the integration event would therefore be extremely high.

There is supporting evidence for both decreased cell viability and instability of single cross-over recombination events. Muscariello et al. (2001) found that *ccpA* mutants of *L. plantarum*, derived from one-step chromosomal integration of suicide vectors carrying sequence duplications, were unstable and they had to use a two-step homologous recombination process in order to obtain a stable chromosomal disruption of *ccpA* gene. Furthermore, they found growth defects in the *ccpA* mutant with a doubling time of the mutant higher than that of the wild type. Based on the observations that CcpA mutants of certain gram-positive organisms have growth impairments, it is possible that the introduction of a mutation in the *C. beijerinckii regB* gene could also result in a negative effect on the growth and metabolism of a viable mutant.

3.4.4 Attempts at inactivation of *regB* using the pGAhost system

The second strategy to obtain a *regB* mutant was based on the temperature-sensitive pGAhost constructs. Electroporation of *C. beijerinckii* with pGAhost, pGAhost1 and pGAhost2 gave a single pGAhost colony, no pGAhost1 colonies and four pGAhost2 colonies after the enrichment step, and the latter were termed "int" strains. The inactivation of *regB* using pGAhost2 would require a 'dco' event between the flanking regions on the suicide vector and the equivalent regions on the *C. beijerinckii* chromosome resulting in deletion of the *regB* gene (Figure 3.5).

In order to obtain 'sco' and 'dco' *C. beijerinckii* recombinants a similar protocol to Biswas et al. (1993) was followed. The system for 'sco' integration should generate *C. beijerinckii* colonies, which are resistant to erythromycin whereas 'dco' recombination should generate erythromycin sensitive colonies because they have undergone a second recombination event which would eliminate the plasmid DNA and the *erm* gene. Erythromycin resistant integrants were obtained after growth at the non-permissive temperature, which should force 'sco' integration and they were designated 'intI'. However, after growth without selection at 30°C, which should force a second recombination event and eliminate the plasmid, there was failure to obtain Em^s strains which would represent 'dco' integration. Attempts at increasing the time period at which the culture was left without selection and repeated sub-culturing without selection failed to produce *C. beijerinckii* 'dco' integrants that were sensitive to erythromycin, and gave only Em^r colonies designated 'intII'. Preliminary Southern hybridization of five 'intI' and five 'intII' Em^r *C. beijerinckii* colonies using the 380 bp *EcoR*V/*Xmn*I *regB* probe indicated that all putative 'intI' recombinants gave two bands, while all the putative 'intII' recombinants gave a single band (results not shown). For further characterization, a single *C. beijerinckii* 'intI' and 'intII' recombinant colony was taken and grown for a few generations at 30°C and 42°C respectively after which further Southern hybridization analysis was performed using the *EcoR*V/*Xmn*I *regB* and 1.180 kb pGAhost *erm* probes (Figure 3.8).

When the chromosomal DNA from the respective samples was digested with *Cla*I and the *EcoR*V/*Xmn*I *regB* probe was used [A], an expected hybridization signal was

obtained at ~ 2.7 kb for wild-type *C. beijerinckii* [lane A2], *C. beijerinckii* (pGAhost) [lane A3] and *C. beijerinckii* (pGAhost2) [lane A4]. A second hybridization signal at ~ 2.0-2.1 kb was also obtained for *C. beijerinckii* (pGAhost2). The second hybridization signal for *C. beijerinckii* (pGAhost2) was not expected since the pGAhost2 construct lacks the 380 bp *EcoR*V/*Xmn*I insert. The possibility of probe contamination however could not be ruled out since the ~ 1.82 kb insert that was subcloned into pGAhost contained 75 bp of the *regB* gene including 1 bp of the *EcoR*V site (position 861 bp, see Figure 2.5; Figure 3.9) and this may have caused the insert to hybridize to the *regB* probe. However even if probe contamination were responsible for the signal then it could still not account for the size of the fragment (~2.0-2.1 kb). All this suggested that pGAhost2 had undergone a rearrangement in *C. beijerinckii* prior to the temperature shift.

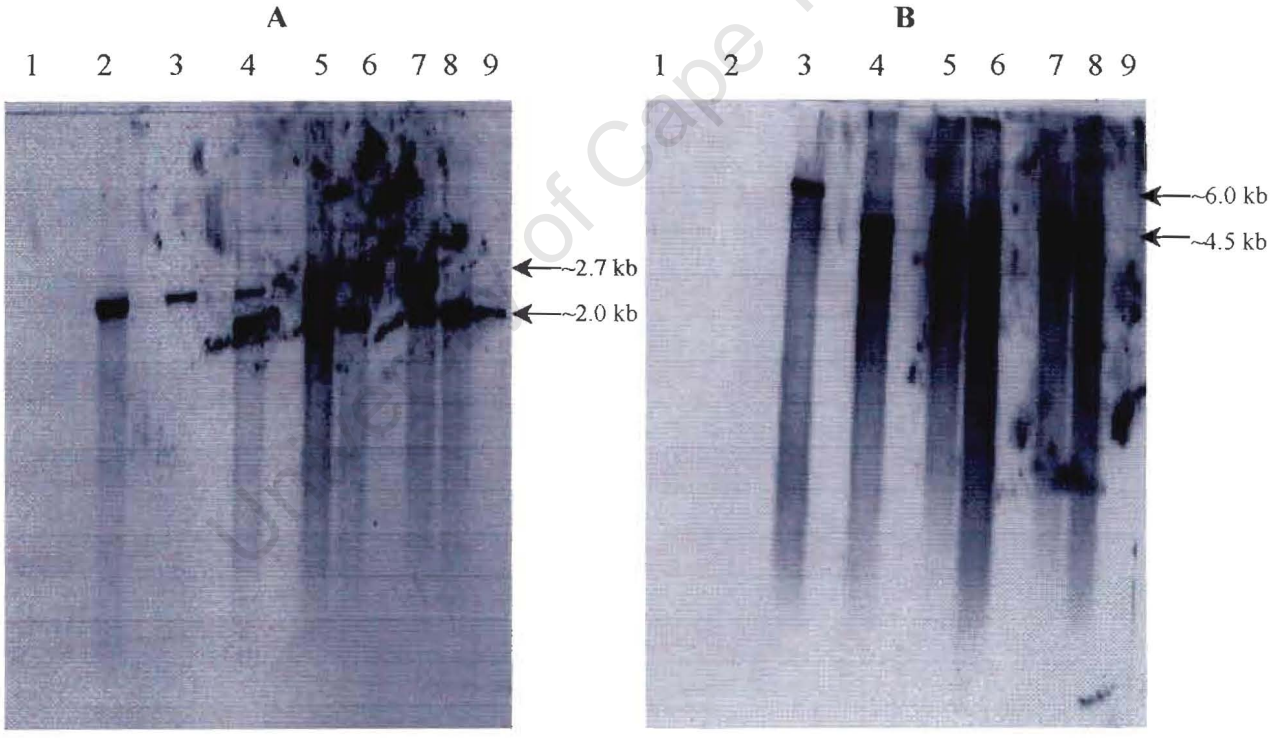

Figure 3.8: Southern hybridization analysis of *C. beijerinckii* wt, *C. beijerinckii* (pGAhost), *C. beijerinckii* (pGAhost2) and putative *C. beijerinckii* 'intI' and 'intII' integrants. Chromosomal DNA (50μg) from the respective *C. beijerinckii* culture was digested with *Cla*I and run on duplicate gels. All DNA samples were hybridized respectively with, the 380 bp *EcoR*V/*Xmn*I *regB* probe (A) and the 1.180 kb *Sac*I *erm* probe from pGAhost (B). The DNA samples in the agarose gel (A) ran unevenly, adjustments with respect to the calculation of the molecular weight of DNA fragments were taken into consideration when determining size using DNAfrag (v.3.03, Canada).

Lane 2: *C. beijerinckii* wt (positive control).

Lane 3: *C. beijerinckii* (pGAhost).

Lane 4: *C. beijerinckii* (pGAhost2).

Lane 5: Putative 'intI' *C. beijerinckii* mutant grown at 30°C.

Lane 6: Putative 'intII' *C. beijerinckii* mutant grown at 30°C

Lane 7: Putative 'intI' *C. beijerinckii* mutant grown at 42°C.

Lane 8: Putative 'intII' *C. beijerinckii* mutant grown at 42°C.

Lanes 1 and 9: Lambda marker (λ DNA digested with *PstI*). The signal in Lane 9 (A) was as a result of sample overflow from Lane 8 (A) during sample loading.

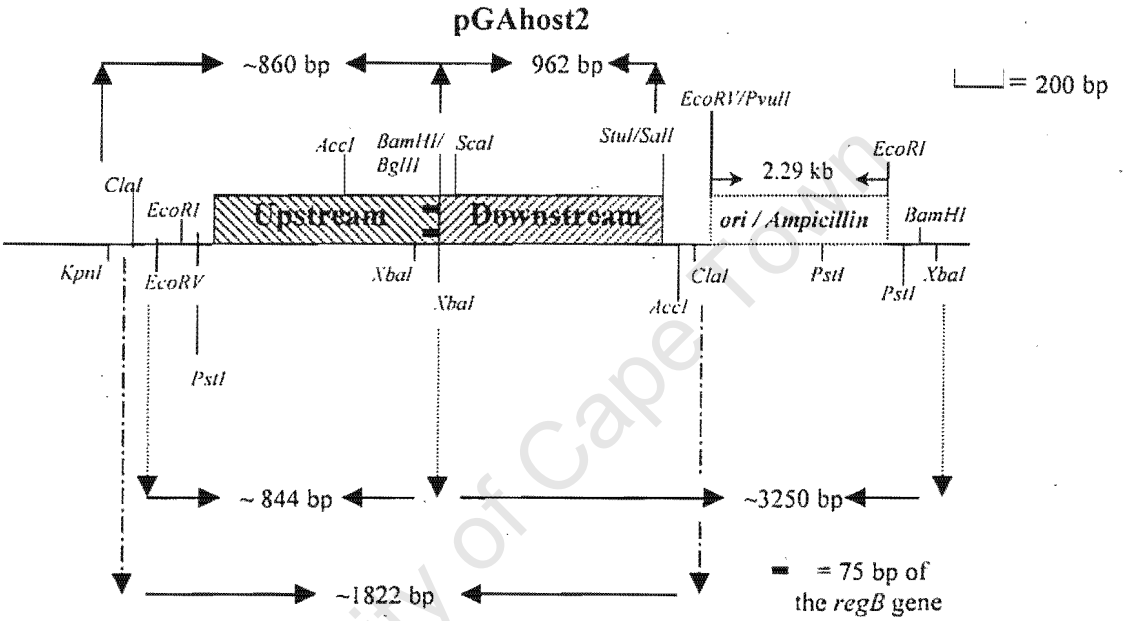

Figure 3.9: A schematic representation of construct pGAhost2 with relevant restriction enzyme sites and fragment sizes. Note that dashed arrows indicate restriction enzyme sites used for hybridization analyses and a thick line on the upstream region indicates the 75 bp of *regB* gene.

The putative *C. beijerinckii* 'intI' integrant (lanes 5 and 7) showed the same hybridization banding pattern as *C. beijerinckii* (pGAhost2) (lane 4). The two bands of ~ 2.0 kb and ~ 2.7 kb, identified in preliminary Southern blots were evident in both cells grown at 30°C and at the non-permissive temperature of 42°C. The ~ 2.7 kb DNA fragment corresponding to the wt *regB* gene and downstream regions therefore has not been disrupted in the 'intI' strains. These strains could therefore represent 'sco' integrants where recombination has occurred in the upstream flanking region (see Figure 3.9), or could be plasmids which have been retained in the cytoplasm as they are identical to pGAhost2 (lane 4). The fact that the intensity of the ~ 2.7 kb *regB* band in

the 'intI' strains is equivalent to the intensity of the ~ 2.0 kb plasmid band would support the latter case. DNA preparations from 'intI' strains were therefore transformed into *E. coli* JM109 in an attempt at plasmid rescue. No transformants were obtained indicating either that plasmids had integrated into the *C. beijerinckii* chromosome or that the plasmid had rearranged and had lost the pBR322 replication functions or selective ampicillin gene.

The putative 'intII' integrant had a single hybridization signal at ~ 2.0 kb, and the ~ 2.7 kb *Cla*I fragment containing the *regB* gene and the downstream regions was not evident. This could represent either a 'dco' event, where a second recombination has occurred in the downstream region, disrupting the ~ 2.7 kb fragment, or a 'sco' event in the downstream region.

To gain further clarification a duplicate blot was probed with the ~ 1.180 kb *erm* gene from pGAhost. This probe gave an expected hybridization signal at ~ 6 kb for *C. beijerinckii* (pGAhost) [lane B3] and no signal for the wild-type *C. beijerinckii*. *C. beijerinckii* (pGAhost2) and the putative *C. beijerinckii* 'intI' and 'intII' integrants gave a hybridization signal at ~ 4.5 kb. All the samples should have given the same signal (~ 6.091 kb) since digestion with *Cla*I releases the ~ 1.82 kb insert from pGAhost2 (see Figure 3.9). This suggested that DNA re-arrangement had occurred in the *C. beijerinckii* (pGAhost2) construct in order to account for the reduction in fragment size. The presence of the *erm* gene in the putative *C. beijerinckii* 'intII' mutant would also explain the erythromycin resistance of the 'intII' integrant, and would suggest that it is not a 'dco' integrant as the plasmid would have been excised in this case.

It has been noted previously that the ss DNA family of plasmids are easily perturbed by insertion or deletion of DNA during vector construction, which leads to structural and segregational instability (Minton et al., 1990). The pG^{+}host system has an active rolling-circle replicon and it has been reported that insertion of these replicons into the chromosome stimulates homologous recombination between flanking repeated sequences 20 to 450 times (Leehouts et al., 1991; Biswas et al., 1993). As discussed previously, 'sco' integrants when subjected to the permissive temperature should

strongly stimulate a second recombination event resulting in high-frequency excision of the replicon giving rise to either the parental or 'dco' chromosomal structure (Biswas et al., 1993).

To confirm the rearrangement of the vector pGAhost2, chromosomal DNA from wild-type *C. beijerinckii*, *C. beijerinckii* (pGAhost) and the putative *C. beijerinckii* 'intI' and 'intII' integrants was digested with *Eco*RV and *Xba*I (Figure 3.10). A hybridization signal at ~ 5.7 kb for wild-type *C. beijerinckii*, *C. beijerinckii* (pGAhost) and the putative 'intI' integrant [Figure 3.10: Lanes 3, 4 and 6), corresponding to the wt *regB* region was obtained. Previous restriction analysis results of the pGAhost2 vector indicated that digestion with *Eco*RV/*Xba*I would release a ~ 3.2 –3.4 kb and 0.84 kb insert (Figure 3.9, Appendix A). These fragments should not hybridize to the 380 bp *regB* probe, however if hybridization should occur due to probe contamination it would give a 0.84 kb or 3.2-3.4 kb signal. However a hybridization signal at ~ 8.6 – 9.0 kb for the putative 'intI' and 'intII' integrants was also obtained. This indicated that probe contamination was not responsible for the hybridization signals obtained for the *C. beijerinckii* 'int' strains and that the wild-type 5.7 kb fragment had been enlarged by the insertion of the plasmid.

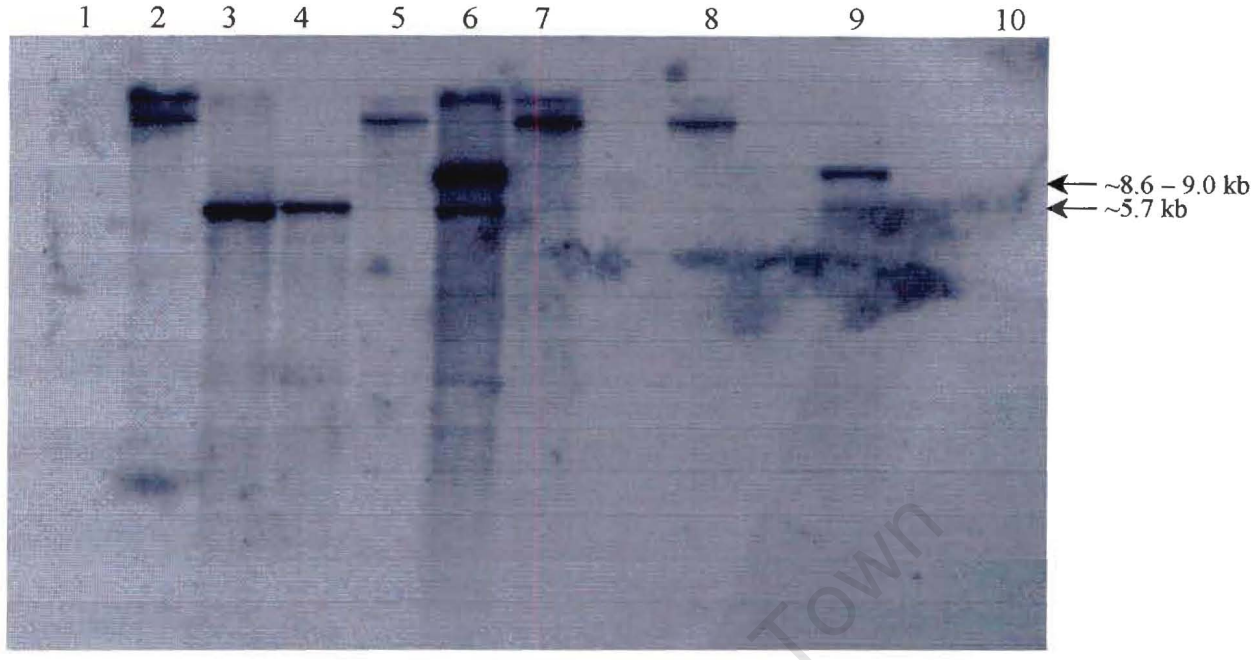

Figure 3.10: Southern hybridization analysis of wild-type *C. beijerinckii*, *C. beijerinckii* (pGAhost) and the putative *C. beijerinckii* 'intI' and 'intII' integrants. Chromosomal DNA (50μg) from the respective *C. beijerinckii* samples were digested with *EcoR*V/*Xba*I where appropriate. All DNA samples were hybridized, with the 380 bp *EcoR*V/*Xmn*I *regB* probe.

Lane 2: Wild-type *C. beijerinckii*, uncut DNA.

Lane 3: Wild-type *C. beijerinckii* digested with *EcoR*V/*Xba*I.

Lane 4: *C. beijerinckii* (pGAhost) digested with *EcoR*V/*Xba*I.

Lane 5: *C. beijerinckii* (pGAhost), uncut DNA.

Lane 6: Putative *C. beijerinckii* 'intI' digested with *EcoR*V/*Xba*I.

Lane 7: Putative *C. beijerinckii* 'intI', uncut DNA.

Lane 8: Putative *C. beijerinckii* 'intII', uncut DNA.

Lane 9: Putative *C. beijerinckii* 'intII' digested with *EcoR*V/*Xba*I.

Lanes 1 and 10: Lambda marker (λ DNA digested with *Pst*I).

In order to confirm whether the *regB* gene from the *C. beijerinckii* 'intI' and 'intII' integrants had been disrupted by the recombination event, the chromosomal DNA from *C. beijerinckii* wild-type and both integrants was digested with *Xmn*I and probed with the 188 bp *Xmn*I probe internal to the *regB* gene. Digestion of chromosomal DNA from wild-type *C. beijerinckii* with *Xmn*I should release a 188 bp internal *Xmn*I *regB* gene fragment and if the *regB* gene was deleted or rearranged in the 'intI' and 'intII' integrants, the hybridization signal should be absent or have increased in size respectively. Previous restriction analysis indicated that the pGAhost2 construct

contained the correct 1.82 kb insert, and that the 188 bp region between the *Xmn*I sites in *regB* (Figure 2.4) was not present. Southern hybridization analysis revealed (results not shown) that no hybridization signal with the *Xmn*I *regB* probe was obtained for plasmid pGAhost2 DNA confirming that the *regB* gene was not present on the plasmid. The predicted 188 bp hybridization signal was detected for wild-type *C. beijerinckii* and also for the 'intI' and 'intII' integrants indicating that the *regB* gene was not disrupted in any of these strains.

Biswas et al. (1993) noted that the thermosensitive vectors replicate via the rolling circle mechanism and could accumulate HMW molecules, which would alter the plasmid conformation and in turn have an affect on the recombination frequency. Gruss and Ehrlich (1988) observed that recombinant plasmids from the ss DNA family do not necessarily retain features of the parental plasmid and that the insertion of foreign sequences may decrease copy number, cause segregational instability or a shift in plasmid distribution from monomeric to multimeric form. Since no *C. beijerinckii regB* mutant was obtained using the pG$^+$host system and because of the difficulties experienced with regard to electroporation and the re-arrangement of the vector constructs, it was decided to abandon attempts to obtain *regB* mutants.

3.5 CONCLUSIONS

Attempts at inactivation of the *regB* gene based on the pMTL30 system (Wilkinson and Young, 1994) resulted in the possible production of a *C. beijerinckii regB* mutant which reverted during growth and sporulation. Gene inactivation using the pG$^+$host system (Biswas et al., 1993) resulted in DNA rearrangements which were difficult to analyse.

Techniques that rely on recombination, such as gene disruption, can only be successful if the plasmid delivery systems are improved. It was found the results from the genetic techniques of electroporation and conjugation were extremely variable and inconsistent in our hands confirming previous reports. Heat treatment of *C. beijerinckii* cultures during the conjugation procedure resulted in only a 50-fold increase in the conjugation frequency. This was not considered a significant difference. Further studies on the effect of heat on the growth phases of *C. beijerinckii* are required in order to optimize

conjugation procedures involving heat treatment. *C. beijerinckii* strain degeneration was a problem that affected the reproducibility of both systems and slowing the growth rate by reducing the growth temperature and initiating the conjugation procedure near the end of the exponential phase helped to reduce the frequency of the condition.

The pMTL30 suicide system produced a single possible *C. beijerinckii regB* mutant that was lost upon subsequent sporulation and germination. Southern hybridization analysis confirmed the disruption of the *regB* gene and suggested that the pMTL30RegB construct was most probably established on the *C. beijerinckii* chromosome in multiple copies. Based on previous growth studies on *ccpA* mutants, it is proposed that the putative *regB* mutant reverted due to a reduction in the growth rate and to the negative effects of the mutation on sporulation. The pMTL30 suicide system produced more consistent results for the creation of *C. beijerinckii regB* mutants. The vector did not appear to be subject to DNA re-arrangement in the host. The major stumbling block in the creation of a *C. beijerinckii regB* mutant using this system appears to be the nature of the mutation itself, the stability of the integration event and the maintenance of the *regB* mutant upon selection.

The nature of the pGAhost2 construct may have been the cause of the failure to obtain a successful recombination and disruption event of the *regB* gene. The *C. beijerinckii* 'intII' strains did demonstrate plasmid integration into the chromosome via 'sco' recombination. If *C. beijerinckii* transformants had been obtained for pGAhost1, which contained only an internal fragment of the *regB* gene, a recombination event leading to disruption may have been more likely even if plasmid re-arrangement should have occurred in the host. Thus the pGAhost vector remains a suitable system to obtain mutants via 'sco' recombination.

There are a number of anomalies obtained with the 'intI' and 'intII' integrants that are difficult to explain. Firstly, the sizes of the DNA fragments detected by either the *regB* probe or the *erm* probe were not as predicted. It would appear that the plasmid pGAhost2 had rearranged on entering *C. beijerinckii*, such that part of the plasmid was

lost. Secondly, plasmid pGAhost2 was detected by a probe designed to detect regions of the *regB* gene which should have been absent from that plasmid.

The failure to obtain a mutant in *C. beijerinckii* resulted in the adoption of alternative strategies to demonstrate (or provide evidence for) a relationship between RegB (CcpA homologue) and the sucrose operon. This was attempted by studying the effects of over-expression of the *regB* gene in *C. beijerinckii* on sucrase activity.

CHAPTER FOUR

CATABOLITE REPRESSION OF THE SUCROSE OPERON OF *C. BEIJERINCKII*

CONTENTS

CHAPTER FOUR

4.1 SUMMARY

Specific glucose repression of the *scrARBK* operon can be demonstrated at the transcriptional level in *C. beijerinckii* by measuring *scrA* expression in a *C. beijerinckii scrR* mutant strain. The ScrR protein was found not to be the only element responsible for the negative regulation of the operon. A difference in glucose repression was obtained when a wild-type *C. beijerinckii* and a *C. beijerinckii scrR* mutant were grown in glucose or glucose and sucrose, indicating that the *scrARBK* operon is subject to glucose repression in addition to negative regulation by ScrR.

Nucleotide and primer extension analysis of the *scrARBK* operon identified a single transcriptional start at an adenine nucleotide, 44 bp upstream of the translational start of the operon. The promoter region, translational start and a *cre* (*cre*I) element had been previously identified (Leat, 1997), and a second, imperfect palindromic sequence (*cre*II), $[A_{12}T_{11}T_{10}A_9T_8A_7G_6T_5T_4T_3T_2C_1/G'_1A'_2A'_3A'_4A'_5T'_6A'_7T'_8A'_9A'_{10}A'_{11}T'_{12}]$ was identified further upstream, 139 bp before the transcriptional start. Analysis of the position of the *scrARBK cre*I element relative to the transcriptional start site showed that it was similar to the *B. subtilis amyE cre* element, which was found to overlap the promoter region and was not subject to Mfd modulation (Zalieckas et al., 1998b). Comparison of the *scrARBK cre*I element with other *cre* elements showed that it had the highest G+C content, predominantly in its flanking sequences, similar to the *B. subtilis acsA cre* site. This suggested that there was a possible mechanism for optimal sucrose induction and glucose repression of the *scrARBK* operon in which total repression by either ScrR or RegB is prevented because of the G+C content in the *scrARBK cre*I flanking regions and because of the suboptimal spacing between the proposed ribosome-binding site and initiation codon in *scrR* which would lower the translational yield of ScrR (Reid et al., 1999).

The RegB protein was successfully expressed in *E. coli* and purified using metal ion chelation chromatography. Using a negative transdominace strategy, it was demonstrated that overexpression of RegB using plasmid pCTC1 in *C. beijerinckii*

resulted in a decrease in the CCR efficiency with respect to sucrose hydrolase activity, in the early and mid exponential *C. beijerinckii* growth phase. This indicated a physiological link between RegB activity and CCR regulation of the sucrose operon and is consistent with RegB as a CcpA homologue.

4.2 INTRODUCTION

The *scrARBK* operon of *C. beijerinckii* has been cloned and sequenced (Reid et al., 1999). Sequence analysis indicated that it consisted of four genes, which are proposed to encode an EIIBC^{Scr} PTS protein (ScrA), a GalR-LacI like transcriptional regulatory protein (ScrR), a sucrose-6-phosphate hydrolase of the glucosyl hydrolase family 32 (ScrB) and an ATP-dependent fructokinase (ScrK) of the ribokinase/*pfkB* family. The distinct homology of the ScrA gene product with EIIBC^{Scr} proteins suggested a role in PTS-dependent sucrose uptake. Sucrose utilization in *C. beijerinckii* therefore involves PTS-dependent sucrose transport, resulting in the accumulation of intracellular sucrose-6-phosphate, which would be hydrolysed by sucrose-6-phosphate hydrolase, releasing glucose-6-phosphate and fructose. The fructose in turn would be phosphorylated by the fructokinase and would be incorporated into the glycolytic pathway. It was shown that the sucrose utilization system was induced by sucrose and not by any other carbohydrates at both the RNA and the enzyme level. Disruption of the *scrR* gene in *C. beijerinckii* resulted in constitutive expression of the *scrARBK* operon under non-inducing conditions indicating that ScrR mediates *scrARBK* regulation and is negatively autoregulated.

In *C. beijerinckii* cultures with glucose and sucrose as substrates, sucrose was utilized more slowly and the initial hydrolase and fructokinase activities were less than half the values observed for cultures grown on sucrose alone (Reid et al., 1999). The *C. beijerinckii* EIIBC^{Scr} protein was thought to interact with a shared EIIA domain, probably that of the glucose PTS (Reid et al., 1999, Tangney et al., 1998). This mechanism would be similar to the phosphorylation of EIIB^{Scr} by EIIA^{Glc} in *B. subtilis* (Sutrina et al., 1990) and the components of a glucose PTS including Enzyme I, HPr and Enzyme II proteins have been identified by complementation in *C. beijerinckii*

(Mitchell et al., 1991). These observations indicated that a mechanism might exist in *C. beijerinckii*, which facilitates glucose-mediated repression of the *scrARBK* operon.

The *C. beijerinckii* and *Staphylococcus xylosus* ScrR proteins shared several conserved domains, corresponding to the inducer binding regions of the LacI protein (Weickert and Adhya, 1992). This suggested that ScrR proteins respond to the same inducer, probably sucrose-6-phosphate (Gering and Brückner, 1996). An imperfect palindromic operator sequence was identified between the *scrA* promoter and initiation codon. The ScrR protein belongs to the GalR-LacI family of transcriptional regulators, which typically act at imperfect palindromic operator sequences centered at a ($A_4A_3N_2C_1/G_1N_2T_3T_4$) consensus (Weickert and Adhya, 1992). It was proposed that ScrR from *C. beijerinckii* negatively regulates *scrARBK* operon transcription by binding to its promoter region, possibly at the palindromic operator sequence, in the absence of sucrose (Reid et al., 1999). However, CcpA homologues, which are a subfamily of the GalR-LacI family, also target palindromic operator sites and could therefore possibly act within the *scrARBK* gene cluster.

CcpA is known to mediate CCR by interacting with *cre* elements in a carbon catabolic operon (Stülke and Hillen, 1999). The *cre* element is a 14 bp core sequence with dyad symmetry, which can be located in the promoter region but has also been reported in the translated region of genes (Hueck and Hillen, 1995). Mutations in the transcription-repair coupling factor (*mfd*) partially relieve CCR at downstream *cre* sites (*hut* and gnt_{down}), but not at *cre* sites located in promoter regions of the *bglPH*, gnt_{up} and *amyE* genes (Zalieckas et al., 1998a). The downstream *cre* element is thought to act as a transcriptional roadblock, stalling the RNA polymerase. Mfd enhances CCR at downstream *cre* sites by displacing the stalled RNA polymerase at CcpA-*cre* complexes.

However, CCR of *acsA* expression, which required a downstream *cre* site (+44.5 bp), was not modulated by Mfd activity (Grundy et al., 1994). Generally, most *cre* sites are flanked by A+T rich sequences, the exception being the *acsA cre* of *B. subtilis*, which contains G or C nucleotides at three or four positions directly adjacent to the 14 bp core

cre sequence (Zalieckas et al., 1998b). Zalieckas et al. (1998b) moved the *acsA cre* element further downstream (to +161.5 bp), and found that Mfd modulated CCR of gene expression demonstrating that the location of the *acsA cre* element relative to the transcriptional start site was the factor that determined Mfd participation. It was hypothesized that RNA polymerase stalled at this position may have a conformation that prevents Mfd-dependent displacement. Alternatively, CcpA may inhibit the initiation of transcription by looping between the downstream *cre* site and an identified upstream *cre* site in the promoter region, which implies that the promoter regions contain unrecognized CcpA binding sites that participate in cooperative binding. The latter is more unlikely since with downstream *xyl cre* sites, HPr-Ser-P is involved in non-cooperative CcpA binding while glucose-6-P triggers cooperative binding of CcpA (Gösseringer et al., 1997). The CCR of *acsA* was shown to be dependent on HPr-Ser-P and not glucose kinase indicating that CcpA binds non-coperatively to the *acsA cre* and acts as a transcriptional roadblock (Galinier et al., 1997).

Zalieckas et al. (1998b) demonstrated that higher levels of CCR were observed at *cre* sites flanked by A+T rich sequences than by those flanked by G+C nucleotides. Therefore the activity of *cre* elements appears dependent on the flanking nucleotides, which suggests that they participate in the formation of the CcpA catabolite repression complex. Kim and Chambliss (1997) demonstrated *in vitro* that CcpA contacts three phosphate groups at the end of the *amyE cre* site. It was hypothesized that replacement of the A+T rich sequences with G and C nucleotides caused subtle DNA conformational changes that diminished and reduced the affinity of CcpA for the *cre* site. This is corroborated by analysis of the ideal 14 bp *lac* operator, which demonstrated that both *in vivo* repression and *in vitro* binding of the *lac* repressor were affected by flanking nucleotides (Lehming et al., 1987).

The proposed mechanism of CCR in most gram-positive species is due to the interaction of three components, namely, a CcpA homologue, HPr-Ser46-P and a *cre* element (Stülke and Hillen, 1999). In the absence of a *C. beijerinckii regB* mutant, this study required evidence for glucose repression of the *scrARBK* operon and confirmation of the presence of cis-acting elements upstream of the *scrARBK* operon. Glucose repression of

the *scrARBK* operon would require physiological analysis at the transcriptional level and the evidence for cis-acting elements would involve genetic analyses employing primer extension and sequencing. Analyses of both of these results would contribute to a proposed mechanism of regulation involving the *scrARBK* operon and RegB (CcpA homologue).

The failure to obtain a *regB* mutant made it difficult to demonstrate direct involvement of RegB in catabolite repression of the *scrARBK* operon. In order to demonstrate the involvement of RegB a strategy used by Heuck et al. (1995) and Kraus et al. (1998) was employed. These authors used negative transdominance in order to investigate the effect of overexpression of CcpA or mutated forms of CcpA on CCR in *B. megaterium*. In the first strategy, Heuck et al. (1995) constructed two plasmids (pWH2014 and pWH2015), both of which contained a XylR repressor and a transcriptional fusion of *ccpA* to the *xylA* promoter region and in addition, construct pWH2015 had an extra *cre* element. The CcpA protein was expressed upon induction with xylose and the amount of CcpA expressed by plasmid pWH2014 was not affected by the addition of glucose or fructose, while there was a 3- to 5-fold reduction in expression from construct pWH2015. These constructs were then transformed into *B. megaterium* WH331, which contained a single chromosomal copy of a *xylA-lacZ* fusion. Plasmid pWH1520, which only contained the XylR, was also transformed into the strain to serve as the control and catabolite repression exerted by glucose on _-galactosidase expression was measured. They found that overproduction of CcpA by strains containing either pWH2014 or pWH2015, decreased the catabolite repression of *xylA* expression by approximately twofold. Therefore, overproduction of CcpA was counterproductive for CCR and was phenotypically similar to a *ccpA* mutant. They proposed that overexpression of CcpA could titrate out co-factors such as phophorylated HPr in such a way that only a part of the CcpA became activated by cofactor binding. This would not then be sufficient to mediate CCR and would effectively result in reduced CCR. These results were in contrast to those of Miwa et al. (1994) who found that overexpression of CcpA in *B. subtilis* did not affect CCR of the *gnt* operon. They however found that overexpression of CcpA was harmful to *B. subtilis* cells because of the induction of gluconate kinase and there was also a negative effect on the growth rate.

Kraus et al. (1998) employed a strategy in which they transformed a *B. megaterium ccpA* deletion mutant (strain WH331) with plasmids encoding mutated CcpA proteins and then tested the *ccpA* mutant proteins for their negative transdominance over wild type. Plasmids such as these, containing the pE194(ts) *ori* would be expected to replicate to about 10 copies per cell (Krause et al., 1998). The gene dosage effect led to an excess of mutant *ccpA* subunits over the wild type, and a decrease in the CCR of the chromosomal *xylA-lacZ* fusion was detected for nearly all the mutants. This indicated that there was Heterodimer formation with the inactive CcpA mutant proteins or titration of the cofactor.

The transcriptional start site and a possible *cre* element (*cre*I) have been identified in the *scrARBK* operon, and in order to demonstrate direct interaction between RegB and the cis-active sites in the *scrARBK* promoter, the RegB protein has been over-expressed in *E. coli* for future gel-shift studies. However confirmation that the *scrARBK* operon is subject to catabolite repression in the absence of a functional *scrR* gene is needed. Also this study sought to examine whether overexpression of *regB* in *C. beijerinckii* would result in a decrease in CCR efficiency of the *scrARBK* operon. Plasmids are available which are maintained in *C. beijerinckii* at copy numbers suitable for gene expression studies. Plasmid pCTC1 is a high-copy number plasmid based on plasmid pMTL500E, which has a copy number of 150-300 per chromosome equivalent (Swinfield et al., 1990). This would, in lieu of a *regB* mutant, provide a link between RegB and CCR regulation of the sucrose operon and be consistent with RegB's role as a CcpA homologue.

4.3 MATERIALS AND METHODS

4.3.1 Bacterial strains, plasmids and culture conditions

All bacterial strains and plasmids are listed in Table 4.1.

Cultures of *E. coli* were routinely maintained on Yeast Tryptone (YT) or Luria-Bertani (LB) medium (Sambrook *et al.*, 1989) at 37°C. YT and LB media were solidified with agar (1.5% w/v). Ampicillin (100ug/ml) and kanamycin (50ug/ml) were included in the media where appropriate.

The laboratory stock of *C. beijerinckii* NCIMB 8052 was maintained, cultured and germinated as previously described (See Chapter 3.3.1).

4.3.2 General DNA manipulation and extraction

See Chapter 2.3.2

4.3.3 Genomic DNA extraction from *C. beijerinckii*

See Chapter 2.3.3

4.3.4 Preparation of RNA from *C. beijerinckii*

C. beijerinckii RNA was prepared from cells, which were grown respectively in CBM containing 1% glucose, 1% sucrose and 1% glucose with 1% sucrose.

See Chapter 2.3.7 for further details.

Table 4.1: Bacterial strains and plasmids used in this study

Name	Relevant Characteristics	Reference
Bacterial strains		
C. beijerinckii		
NCIMB 8052	Wild-type	NCIMB
CBSCRR	*Scr::erm*	Reid et al. (1999)
E. coli		
JM105	*_(lac-proAB) lac^q _(lacZ)M15*	Yanisch-Perron et al. (1985)
JM109	*_(lac-proAB) lac^q _(lacZ)M15 recA1*	Yanisch-Perron et al. (1985)
CA474	*F' thi-1, _cI857, hsdS20(r^-_B, m^-_B), supE44 rpsL20(Sm^r) recA13*	Wilkinson and Young (1994)
M15 (pREP4)	*RecA$^+$, uvr$^+$, mtl^r, lac$^-$, Nal^s, Str^s*	Qiagen GmbH, Germany
Plasmids		
pBluescriptSK$^+$	Ap^r, T3 and T7 polymerase promoters	Stratagene, La Jolla, California
pECORegB	~5.5 kb DNA fragment containing *regB* in pECOR251	See Chapter 2
pCTC1	Amp^r Em^r; Tra$^-$ Mob$^+$	Williams et al. (1990)
pSKRegB	*Acc*I pECORegB fragment subcloned into pBluescriptSK$^+$	This study

pCTCRegB	*Xho*I/*Sac*I insert of pSKRegB subcloned into pCTC1*Xho*I/*Sac*I sites.	This study
pUCRegBCm		See Chapter 3
pCTCRegBCm	*Pst*I/*Bgl*II insert from pUCRegBCm subcloned into pCTC1	This study
pGDR11	Modified pQE31 protein expression vector	Ehrmann M. (University of Konstanz, Germany, personal communication)
pQEG1RegB	*regB Bam*HI/*Pst*I insert subcloned into pGDR11	This study
pCBS3	Amp^r Em^r; 3.2 kb *Hind*III/*Nde*I fragment recovered from CBSCRB	Reid et al., 1999
pCBS3a	*Eco*RI/*Hind*III derivative of pCBS3 subcloned into pBluescriptSK$^+$	Leat N.G. (personal communication)
pCBS4	Amp^r Em^r; 1.2 kb isolated from *C.beijerinckii* by PCR	Reid et al., 1999
pCBS4a	ClaI/XbaI derivative of pCBS4 subcloned into pBluescriptSK$^+$	This study

4.3.5. RNA Slot blotting

Plasmid pCBS3 was a pMTL30 plasmid rescued from the chromosome of *C. beijerinckii* (CBSCRB) mutant by digestion with restriction endonuclease, *Hind*III and *Nde*I (Leat, 1997). The plasmid contained a truncated *scrA* (encodes EIIBC^{Scr} PTS protein), the *scrB* gene and the complete *scrR* gene. Subsequently an *EcoR*I/*Hind*III restriction digest of pCBS3 allowed the insert to be subcloned into pSKBluescript$^+$ and renamed pCBS3a (Leat, personal communication). We constructed a *scrA* DIG probe by digesting pCBS3a with restriction endonucleases *Hind*III and *Sty*I. The digestion released an internal 373 bp *scrA* insert, which was DIG labelled. Non- radioactive DIG-labelled probes were prepared using the Roche Diagnostics DIG random primer labelling kit (catalog number: 1175033).

C. beijerinckii RNA (20µg) from each sample (50µl sample volume) was loaded in duplicate into the wells of the Slot blotting apparatus (Hoefer Scientific Instruments, San Francisco, CA, USA). The RNA was transferred via vaccum blotting onto Hybond+

nylon membranes and UV crosslinked. Hybridization and detection of the *scrA* DIG-labelled probe was conducted according to the protocols of the DIG Easy hybridization system supplied by Roche Diagnostics.

4.3.6 Primer Extension of the *scrARBK* operon

RNA was extracted from *C. beijerinckii* cultures grown in CBM with sucrose (1%) as the carbon source. Primer extension experiments were carried out according to Ausubel et al. (1989), with the addition of actinomycin D (0.225mg/ml). The primers used for primer extension of *scrARBK* were internal to the *scrA* and *scrK* genes: primer a (*scrA* gene: 5'-TTCTTTTGCTACTATTTGTTCC-3') and primer k (*scrK* gene: 5'-TTATCTACACAAAGCATATC-3').

4.3.7 Nucleotide sequencing and sequence analysis

Plasmid pCBS4 (Reid et al., 1999) was digested with restriction endonucleases *Cla*I and *Xba*I to yield a 1.4 kb fragment which contained the *scrA* upstream region including a 140bp region which was previously sequenced. This 1.4 kb *Cla*I/*Xba*I fragment was subcloned into pSKBluescriptSK$^+$ which was also digested with *Cla*I/*Xba*I to yield pCBS4a.

This construct was sequenced using the Sequitherm^{TM} kit (Epicenter Technologies, Madison, USA) with CY-5^{TM} labeled M13 primers and the OmniGene thermocycler (Hybaid). Nucleotide sequence was resolved using the ALFexpress automated DNA sequencer (Pharmacia). Sequence analysis was performed using programs in the DNAMAN V4.0 (Lynnon Biosoft, Canada) and Genetics Computer Group (Devereux et al., 1984) sequence analysis packages.

4.3.8 PCR amplification of the *regB* gene

The *regB* gene contained on plasmid pECORegB was amplified using two PCR primers.

The forward primer, fRegb31, (5'-GCGGATCCTACTTCTATTAAGG-3') contained a *Bam*HI restriction site at the N-terminus (underlined).

The reverse primer, rRegB, (5'-CACTGCAGCTGTACTTTCTCTTTC-3') contained a *Pst*I restriction site at the N-terminus (underlined).

The following optimized PCR regime was used for 30 cycles: Step 1, template denaturation at 95°C for 15 sec; Step 2, primer annealing at 27°C for 30 sec and Step 3, strand synthesis at 72°C for 60 sec. PCR reactions were conducted using an OmniGene thermocycler (Hybaid) and the Expand™ High Fidelity PCR System (Roche Diagnostics). After PCR amplification of the *regB* gene, the PCR product was digested with *Bam*HI and *Pst*I and subcloned into the pQE30/31/32 vector series and pGDR11 (see 4.3.9).

4.3.9 *In vitro* protein expression and purification of RegB

The RegB protein was expressed using the pQE expression system, which is designed to express foreign proteins in *E. coli* (Qiagen GmbH, Germany). This system employs 3 protein expression vectors (pQE30/31/32), in which the genes are fused to six histidine sequences for affinity purification. The histidine tag was at the N-terminus of the expressed protein.

It was subsequently found after subcloning into pQE30/31/32 (resultant vectors pQE0/1/2RegB) that the pQE vectors were unstable and vector pGDR11 was used (Ehrmann M., personal communication). Vector pGDR11 is a pQE31 vector containing an additional *lacI* element, which allowed for plasmid stability. Vector pGDR11, which contained the *regB* gene was named pQEG1RegB and was used for protein expression of the *regB* gene in *E. coli*.

RegB protein was expressed and purified according to the method of Küster et al. (1996) with minor modifications. The vectors pGDR11 and pQEG1RegB were transformed respectively into *E. coli* M15 (pREP4). The *E. coli* cultures (50ml) were grown at 30°C and overexpression was induced by the addition of 2mM IPTG when the culture reached an OD_{600} of 0.5. Growth was allowed to proceed for another 4 h. The optical densities of the cultures were measured and standardized in order to obtain similar protein concentrations for all the samples and the cultures were then harvested (5000 rpm for 5 min at 4°C). The supernatant was discarded and the cell pellet was kept

at –20°C overnight. The frozen pellets were thawed on ice and resuspended in 1 ml of SBT buffer (50mM Tris-HCl, pH 7.5; 200mM NaCl; 5mM 2-mercaptoethanol and 0.1mM PMSF). Cell disruption was achieved by sonication (Virtis Sonic); where 30 s pulses interspersed with 30 s intervals were applied 5 times. The lysate was centrifuged at 8000 xg for 30 min at 4°C and the supernatant was used immediately or kept at –20°C overnight.

The crude extract (4ml) was mixed with 2 ml Ni-NTA resin (Qiagen, Chatsworth, CA), which was pre-equilibrated with Buffer A (20mM Tris-HCl, pH 8.5; 100mM KCl, 5mM 2-mercaptoethanol, 10% glycerol and 20mM imidazole). This slurry was then loaded onto a column (5ml plastic syringe with 5 mm layer of glasswool), which had a pre-determined flow rate of 0.5ml/min using BufferA. The column was washed with 10 vol of Buffer A, followed by 2 vol of Buffer B (20mM Tris-HCl, pH 8.5; 1M KCl, 5mM 2-mercaptoethanol and 10% glycerol) and another 2 vol wash with Buffer A. The bound proteins were eluted with Buffer C (20mM Tris-HCl, pH 8.5; 100mM KCl, 5mM 2-mercaptoethanol, 10% glycerol and 100mM imidazole), collected in 0.5ml fractions and later pooled.

Protein samples were separated by SDS-PAGE using the discontinuousbuffer system of Laemmli (1970). Acrylamide concentrations for SDS-PAGE were 3.5 and 12% (w/v) in the stacking and resolving gels, respectively. Low molecular weight markers (Amersham Pharmacia Biotech) were used as molecular weight standards. A 25µl sample of an elution pool was added to 25µl protein tracking dye (Amersham) and 20µl loaded onto the electrophoresis gel. Samples were heated at 37°C for 10 min before loading the gel and run using a Hoefer SE600 vertical slab electrophoresis unit (Hoefer Scientific Instruments, San Francisco, CA, USA). Gels were fixed and stained with Coomassie blue.

4.3.10 Overexpression of RegB protein in *C. beijerinckii*

Construct pECORegB was digested with *Acc*I which releases a ~1483 bp fragment containing the entire *regB* gene including the transcriptional and translational start. The restriction sites of the insert were filled-in and subcloned into vector pSKBluescript^{+},

previously digested with *Eco*RV. The resultant construct was called pSKRegB. This construct was digested with *Xho*I and *Sac*I to release the 1483 bp insert, which was then subcloned into vector pCTC1 previously digested with *Xho*I and *Sac*I. The resultant construct was called pCTCRegB where the *regB* promoter drove expression of the *regB* gene.

Construct pUCRegBCm (refer to 3.3.7) was digested with *Pst*I, the restriction end filled-in and then digested with *Bgl*II which released a ~1953 bp fragment containing the *regB* gene interrupted by the chloramphenicol resistance gene. The insert was subcloned into pCTC1 previously digested with *Bam*HI and *Stu*I (blunt) and the resultant construct was called pCTCRegBCm. Both constructs were transferred to *C. beijerinckii* via conjugation as previously described in Chapter3.

4.3.11 Enzyme and protein assays

Sucrose hydrolase activity was assayed using an adaption of the dinitrosalicylic acid method, as described by Blatch and Woods (1993). *C. beijerinckii* containing the relevant plasmid constructs were grown in CBM containing respectively glucose (1%), sucrose (1%) and glucose (1%) with sucrose (1%). Duplicate cultures (50ml) were centrifuged at 3000 xg for 10 min at 4°C. The cell pellet was washed in 5ml citrate-phosphate buffer (0.0179M citric acid, 0.0642M Na_2HPO_4, pH 6.0) and then centrifuged at 6000rpm for 5 min. The pellets from the early to mid log phase cells (OD_{600nm} < 0.7) and the late log phase cells (OD_{600nm} > 1.0) were washed in 1 ml and 2 ml citrate-phosphate buffer respectively. The cells were then frozen overnight at –20°C or immediately sonicated. The sonication (Virtis Sonic) consisted of three 30 sec bursts with 30 sec intervals in between and cells were kept on ice at all times. The cell-free extracts (CFEs) were spun at 10000rpm for 30 min at 4°C. The CFEs were standardized to the same protein concentration and the assays done in triplicate. 15µl of sucrose (0.88M) was added to 35µl of CFE and incubated at 30°C for 15 min. The reaction was terminated with the addition of 150µl DNS and boiling for 5 min. After the reaction mixture was cooled on ice, 800µl of distilled water was added and absorbance readings taken at OD_{510nm}. Sucrose hydrolase activities were expressed as mmoles reducing sugar

produced per min per mg protein. Protein concentrations were determined using the method of Bradford (1976), with bovine serum albumin as the standard.

4.4 RESULTS AND DISCUSSION

4.4.1 Catabolite repression of the *C. beijerinckii scrARBK* operon

There was physiological evidence of glucose repression of the *scrARBK* operon when *C. beijerinckii* was grown in glucose and sucrose (Leat, 1997). This however did not indicate that there was any direct link between the sucrose operon and the effects of glucose metabolism. The repressive effect of glucose on sucrose utilization in *C. beijerinckii* may have been due to secondary metabolic effects higher up in the glycolytic pathway.

This study therefore sought to demonstrate specific glucose repression of the *scrARBK* operon at the transcriptional level. Leat (1997) inactivated the *scrR* gene of this operon by insertional inactivation, creating a mutant strain which was unable to produce a functional repressor protein, and resulting in constitutive expression of the *scrA* gene upstream of *scrR*. This would allow the study of any glucose effects on the *scrARBK* operon, without interference by the ScrR protein. RNA from *C. beijerinckii* wild-type and CBSCRR (*C. beijerinckii scrR* mutant), which were grown in glucose and glucose with sucrose respectively, was hybridized with a *scrA* probe after slot blotting and the specific signals were measured.

From the results (Table 4.2), it can be seen that *scrA* expression in wild-type *C. beijerinckii* is induced in the presence of sucrose while virtually no transcription is observed when glucose is the sole carbon source. The presence of both carbon sources in the wild-type leads to a 83% reduction in transcription of *scrA*. In glucose-grown cells of the *C. beijerinckii scrR* mutant, *scrA* transcription occurs at nearly half the value of that observed in sucrose-induced cells of the wild-type These results and their and their limitations are considered more fully in the report. Furthermore, glucose repression still occurred in the *scrR* mutant, because there was a 1.4-fold reduction in *scrA* transcription when grown in glucose plus sucrose as compared to glucose alone. This difference in transcription levels of *scrA* in CBSCRR is consistent with the

physiological evidence that the *scrARBK* operon is subject to glucose repression in addition to negative regulation by ScrR.

Table 4.2: Densitometer measurements of *scrA* transcription in *C. beijerinckii* wild-type and CBSCRR

Carbon Source†	Densitometer measurements					
	C. beijerinckii wild-type			*C. beijerinckii* CBSCRR		
	Sample 1	Sample2	Average	Sample 1	Sample 2	Average
Sucrose*	0.50	0.56	0.53	–	–	–
Glucose	0	0	0	0.22	0.21	0.215
Glucose + Sucrose	0.08	0.10	0.09	0.15	0.16	0.155

†CBM containing 1% glucose and 1% sucrose respectively or both carbon sources

**C. beijerinckii* CBSCRR cannot grow with sucrose as the sole carbon source (Leat, 1997)

Vibrio Cholerae RNA served as the negative control.

4.4.2 Primer extension analysis of the upstream region of the *C. beijerinckii scrARBK* operon

Primer extension of the *scrARBK* operon using oligonucleotide primer a revealed a single transcriptional start at an adenine nucleotide, 44 bp upstream of the translational start of the operon (Figure 4.1). This confirms the putative –10 and –35 promoter regions which were previously identified at 51 bp and 80 bp before the translational start (Reid et al., 1999). A putative *cre* element ($G_6A_5A_4A_3A_2C_1/G'_1G'_2T'_3T'_4T'_5C'_6$), *cre*I, was also identified by Leat (1997), between the –10 region and the ATG start codon (Figure 4.2). A second imperfect palindromic sequence ($A_{12}T_{11}T_{10}A_9T_8A_7G_6T_5T_4T_3T_2C_1/G'_1A'_2A'_3A'_4A'_5T'_6A'_7T'_8A'_9A'_{10}A'_{11}T'_{12}$) 139 bp before the transcriptional start, *cre*II (Figure 4.2) was identified.

Primer extension analysis using oligonucleotide primer k did not yield any primer extension reaction product (results not shown). Previous northern blot analysis of wild-type *C. beijerinckii* indicated that all four *scrARBK* genes were co-transcribed as a 5.0 kb operon, however an additional 1.1 kb transcript was detected with a *scrK*-specific probe (Reid et al., 1999). It was thought that the transcript represented an additional *scrK* transcript expressed from a putative promoter identified within the coding region of *scrB*. However primer extension analysis failed to confirm this as a functional

promoter and the 1.1 kb transcript was thought to be a mRNA from an independent gene with homology to *scrK*.

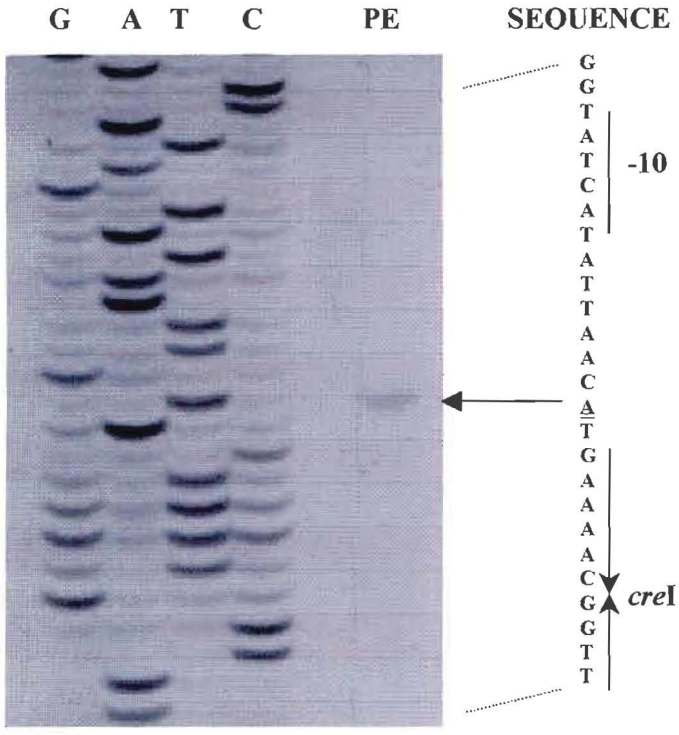

Figure 4.1: Identification of the promoter and transcription initiation site of the *C. beijerinckii scrARBK* operon. A DNA sequence ladder of the promoter region is shown on the left. PE = primer extension reaction product. An oligonucleotide primer a was used as the primer for both the DNA sequence and primer extension reactions. The transcription initiation site at an A (underlined) residue is indicated, and the position of the palindromic sequence is indicated by arrows.

```
  1   CCCTTAAATG AAGAACTTGG AAAAAATAGT CAAGATAATT TATGTTCAAA
 51   AAAATTACAT TTTAAAAATA CAGTATAGTT AATAAGTGTA TTCCTTAAAC
101   TACAAAAGGA GTACACTTTT TTTATGTTCT AAGAACTAAT TGTTTTATAT
151   ATAAATTTAA TTCTTAAAAT TAGCAATAAT AGTGGGAACA GTACCACATA
201   AGATGAATTA TAGTTTTC♦GA AAATATAAAT CATATGAATA GGGAGGCGTA
251   ATGGATAATA TATTAGAAAC TATGCATTGT ACATGGGGGA GAAAGGAAAT
                                                  -35
301   TTATATAATT TATTATTTAG GTAAAAAACG TATTGACACA TTATATATGG
                -10
351   AACTGGTATC ATATTAACAT GAAAAC♦GGTT TCCGAGGTTA ATTGAATTAG

401   GAGGGGACAT TAATGAAGGA ACAAATAGTA GCAAAAGAAA
      SD          ScrA ──►
```

Figure 4.2: Nucleotide sequence of the upstream region of the *scrARBK* operon of *C. beijerinckii*. Two palindromic sequences with a CG (♦) center are shown (double underlined), the first, *cre*I, starting at

base pair position 371 was identified previously (Leat, 1997). The second, *cre*II, was identified starting at base pair position 207.

↓ = Transcriptional start; SD = Shine Dalgarno

If the *cre*I element of the *scrARBK* operon is compared to other *cre* elements it is seen that it has the highest G+C content, especially in its flanking regions, even more than that of the *acsA* cre_{down} element (Table 4.3).

Table 4.3: Sequence alignment and positions of *cre* elements

Gene or operon[a]	DNA sequence	Position (bp)[c]	G+C[d] content
Consensus *cre*[b]	WWWWTGWAARCGYTWNCWWWWW		
B. subtlis amyE	AAATTGTAAGCGTTAACAAAAT	+4.5	0 (5)
**B. subtlis xylA*	ATTTTGGAAGCGTAAACAAAGT	+140.5	1 (7)
B. subtlis $acsA_{down}$	AACTTGAAAGCGTTACCAGCAA	+44.5	3 (9)
B. subtlis hutP	GAATTGAAACCGCTTCCAAAAA	+209.5	0 (8)
*C. beijerinckii scrARBK cre*I	AACATGAAAACGGTTTCCGAGG	+7.5	4 (10)
*C. beijerinckii scrARBK cre*II	TTATAGTTTTCGAAAATATAAA	-150.5	0 (3)
S. xylosus scrB[e]	ATTATGGAACCGGTACCATTTA	+11.5	0 (8)
B. subtlis acu	TTGTTGAAAACGCTTTATAATT	-26.5	1 (5)
B. subtlis lev	ACAATGAAAACGCTTAACACAA	-45.5	2 (7)

[a] References for all the *cre* sequences except (e)were obtained from Zalieckas et al (1998b)

[b] Symbols for ambiguous nucleotides in the consensus *cre* sequence are as follows: W represents A or T; R represents A or G; Y represents C or T; and N represents A, C, G or T.

[c] The location of each *cre* is given as a center position relative to the transcriptional start

[d] The G+C content refers to the 4 bp (in bold) on either side of the main 14 bp consensus (underlined) and the total G+C content (in brackets).

[e] *Staphylococcus xylosus* palindromic sequence (O_B) in the promoter region of the *scrB* gene (Gering and Brückner, 1996).

*cre elements that have been confirmed to have bound to CcpA.

The *scrARBK cre*I site has two noteworthy features; firstly the position of the *scrARBK cre* element relative to the transcriptional start site is similar to that of the *amyE cre*

element (see Table 4.3). The *amyE cre* element overlaps the promoter region and is not subject to Mfd modulation of CCR. Secondly, the *scrARBK cre* element sequence context is similar to the *acsA cre* site with its flanking regions rich in G and C nucleotides. Based on these observations, it suggests that the *scrARBK cre* site is not subject to modulation by a transcription-repair factor and in this regard is similar to the *amyE cre* site. However, the CCR of the *scrARBK* operon may be influenced in a manner similar to the CCR of *acsA* by the G and C nucleotides surrounding their respective *cre* sites resulting in possible suboptimal binding of regulatory proteins at the *scrARBK* operon. Reid et al., (1999) also noted that only 5 bp separated the proposed ribosome-binding site and initiation codon in *scrR* compared to the conventional 7-9 bp spacing. Vellanoweth and Rabinowitz (1992) demonstrated in *B. subtilis* that the translational yield was substantially reduced if the ribosome-binding sites and initiation codons were separated by less than 6 bp. This suboptimal spacing may limit translation of *scrR*, in order to allow for optimal expression of the *scrARBK* operon and prevent large amounts of ScrR from being produced and conceivably repressing *scrARBK* transcription even if the inducer is present. Reid et al. (1999) demonstrated that the *scrARBK* operon is subject to tight regulation, being only induced by sucrose. These observations together with ours suggest that the *scrARBK* operon could allow optimal sucrose induction and glucose repression by preventing total repression by either ScrR or RegB via the G+C content in the *scrARBK cre*I flanking regions as well as by lowering the translational yield of ScrR. Both mechanisms, in theory, would allow for low-level expression of the *scrARBK* operon and the build-up of necessary gene products while the inducer is absent.

It is interesting to note that in the *C. acetobutylicum* ATCC824 sucrose operon, an element similar to the *cre* consensus was identified in the promoter region of *scrT* before an attenuator sequence (Tangney and Mitchell, 2000). This *cre* site differed from the consensus *cre* in having a TG palindromic center rather than the more common CG palindromic center. However, their sucrose operon has a very different mechanism of regulation because it contains elements of an antiterminator-mediated mechanism (ScrT and RAT sequences), typical of the BglG family of regulators.

The *S. xylosus* genes for sucrose utilization are very homologous to the *C. beijerinckii scrARBK* genes, however the *S. xylosus* sucrose genes are not clustered in the genome (Wagner et al., 1993). In the *scrA* and *scrB* promoter regions of this bacterium, a palindromic sequence (O_A and O_B) has been identified respectively (Gering and Brückner, 1996, see Table 4.3). The O_B sequence was identical to the *cre* consensus sequence except for a deviation in one nucleotide (see Table 4.3). The ScrR protein had about 30% identity to RegA of *C. acetobutylicum* NCP262 and to the CcpA proteins of *B. subtilis* and *B. megaterium*. The authors demonstrated *in vitro* that ScrR binds to *scrA* O_A and *scrB* O_B, inactivation of ScrR led to constitutive expression of *scrB* and that upon deletion of 4 bp of the right-arm of O_B, sucrose specific regulation was lost. Thus ScrR is the transcriptional repressor that controls the sucrose-specific regulon and O_A and O_B serve as the cis-active sequences mediating this specific regulation. The *S. xylosus* ScrR, being a member of the GalR-LacI protein family was thought to bind O_B as a dimer since there was no further possible binding region for ScrR. However they have not yet performed similar experiments using CcpA proteins, and these sites could also function as *cre* elements.

The identified imperfect palindromic sequence, *cre*II, 139 bp from the *scrARBK* transcriptional start resembles a typical GalR-LacI operator sequence centered around a ($A_4A_3N_2C_1/G'_1N'_2T'_3T'_4$) consensus (Weickert and Adhya, 1992). The dyad symmetry would allow members of the GalR-LacI protein family to bind as dimers to its target DNA. In this case, one monomer of the regulator would interact with one half-site of the palindromic sequence. The DNA binding domain of PurR, a GalR-LacI protein member, was shown to bind into the minor groove of DNA and bend it open by intercalation of a leucine residue into the central CG of the operator (Schumacher et al., 1994). Tetramerization is also observed within this family allowing them to bind more than one operator simultaneously (Weickert and Adhya, 1992). The *scrARBK cre*I site differed by only 2 bp from the concensus *cre* element while *scrARBK cre*II differed by 7 bp (see Table 4.3, Zalieckas et al., 1998b). Both *scrARBK cre* sites are likely binding sites for RegB or ScrR, and binding of either regulator to *cre*I would block the transcriptional start site of the *scrA* gene, causing repression of the operon.

Alternatively, the imperfect palindrome could facilitate DNA looping involving ScrR or RegB and possible tetramization of either regulator.

If binding of RegB and ScrR proteins occurred at specific *scrARBK cre* sites, independent of each other, then the mechanism of regulation of the *scrARBK* operon would be very similar to that of the *B. subtilis xyl* operon (Dahl and Hillen, 1995; Schmiedel and Hillen 1996). Xylose utilization in *B. subtilis* is encoded by *xylA* (xylose isomerase) and *xylB* (xylulose kinase) and negatively regulated at the level of transcription by XylR (Kraus et al., 1994; Dahl and Hillen, 1995; Dahl et al., 1995). In the absence of the inducer xylose, XylR represses *xyl* operon transcription by interaction with a tandem *xyl* operator, O_L and O_R (Figure 4.3). Catabolite repression occurs via CcpA/HPr-Ser-P interaction with a *cre* site in the *xylA* gene. Similarly, in the absence of sucrose, *C. beijerinckii* ScrR could bind to a specific *scrARBK cre* site, turning down transcription of the *scrARBK* operon. In the presence of both carbon sources, RegB modulated by phosphorylated HPr (or other co-factors), could maintain this repression by binding to the other *scrARBK cre* site until the glycolytic status falls, relieving binding of RegB/HPr to that *cre* site and allowing transcription of the *scrARBK* operon.

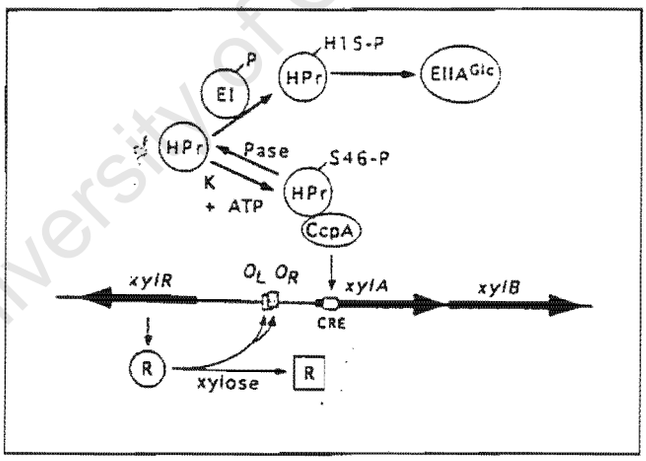

Figure 4.3: Genetic organization of the *xyl* operon from *B. subtilis* composed of *xylA* and *xylB* and the divergently orientated gene *xylR* (encoding Xyl repressor). Under uninduced conditions, the Xyl repressor (R) exhibits a conformation (circle) able to bind the *xyl* operators O_L and O_R. Xylose interacts with XylR to induce the expression of the *xyl* operon. Xylose bound XylR (square) is not able to interact with *xylO*. A *cre* element is located within the coding sequence of *xylA*. Serine 46 (S46) phosphorylated HPr interacts with CcpA leading to catabolite repression by binding at the *cre* element. The phosphorylation

of HPr at S46 is catalyzed by a protein kinase (K) in the presence of ATP, dephosphorylation is catalyzed by a phosphatase (Pase). HPr is also phosphorylated at histidine 15 (H15) by the P-EI. This phosphate is transferred to enzymeIIA^{Glc}. Adapted from Dahl and Hillen (1995).

4.4.3 Overexpression of RegB in *C. beijerinckii*

The *regB* gene was subcloned into plasmid pCTC1 to create construct pCTCRegB to study the effect of over-production of the RegB protein on sucrose hydrolase activity in *C. beijerinckii*. In order to ascertain whether the vector copy number could influence sucrose hydrolase activity, a *regB* gene interrupted by the chloramphenicol resistance gene was subcloned into the pCTC1 vector to serve as the control. Both vectors were transferred to *C. beijerinckii* via the conjugation procedure as previously described (see Chapter 3). The sucrose hydrolase activities were measured from *C. beijerinckii* cell-free extracts containing the respective plasmid constructs at three different growth phases (Figure 4.5).

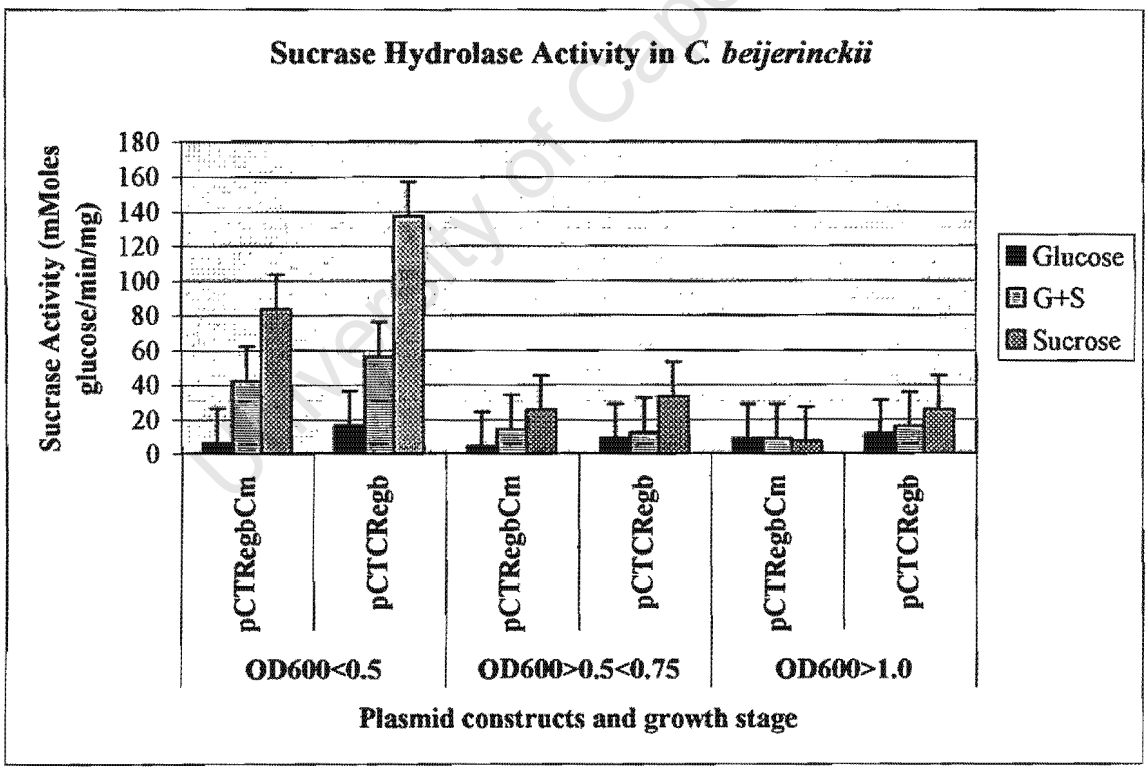

Figure 4.5: Bar graph indicating sucrase hydrolase activities in *C. beijerinckii* containing either plasmid pCTRegBCm or pCTCRegB and grown respectively in 1% glucose, 1% sucrose and 1% glucose with 1% sucrose. Results are the average of at least three experiments, and standard deviations are given as bars.

Early exponentially growing cultures of *C. beijerinckii* (OD_{600} < 0.5), which overexpressed the RegB protein showed a greater glucose repression ratio (2.442) than the control (1.996) (Table 4.4). The glucose repression ratio increased in parallel with *C. beijerinckii* (pCTCRegB) growth phases, with the degree of catabolite repression being the greatest in the late growth phase. The sucrose induction ratio for *C. beijerinckii* (pCTCRegB) in the early- and mid- growth phases was lower (8.331 and 3.780 respectively) when compared to the control (13.469 and 6.192 respectively). The sucrose induction ratio for *C. beijerinckii* (pCTCRegB) in the late growth phase (OD_{600}>1.0) was higher (2.25) when compared to the control (0.834). However the error bars for the mid- and late- growth phase data overlap each other and are thus not statistically reliable and interpretable.

Since there was a large standard deviation in the measurements only the early growth phase data can be interpreted. According to the data, there was greater repression by glucose in the *C. beijerinckii* strain over-expressing RegB with a concomitant decrease in the sucrose induction ratio when compared to the control. This would indicate that something, presumably RegB, is interfering with full induction of the sucrose operon.

Table 4.4: Sucrose hydrolase activities of *C. beijerinckii* containing pCTC1-based plasmid constructs

Growth Phase[f]	***C.beijerinckii* with the relevant plasmid construct**	**Sucrose hydrolase Activities (mmoles glucose/min/mg) Carbohydrate source[a]**			
		Glucose	Sucrose	Glucose +Sucrose	
Early	pCTCRegBCm	6.21	83.63	41.90	
	pCTCRegB	16.47	137.22	56.19	
Mid	pCTCRegBCm	4.06	25.17	14.01	
	pCTCRegB	8.73	33.01	12.16	
Late	pCTCRegBCm	8.56	7.22	8.58	
	pCTCRegB	11.24	25.27	15.65	
		Repression by glucose ratio[b]	**Efficiency of Glucose repression ratio[d]**	**Sucrase induction ratio[c]**	**Efficiency of Sucrase induction ratio[e]**
Early	pCTCRegBCm	1.996	1.233	13.469	1.6
	pCTCRegB	2.442		8.331	
Mid	pCTCRegBCm	1.794	1.513	6.192	1.6
	pCTCRegB	2.715		3.780	
Late	pCTCRegBCm	0.841	1.919	0.843	0.37
	pCTCRegB	1.614		2.25	

[a] Carbohydrate sources were 1%w/v

[b] Sucrase activity in the presence of sucrose divided by activity in the presence of sucrose and glucose (values obtained using the same method of Dahl and Hillen, 1995).

[c] Sucrase activity in the presence of sucrose [induced] divided by activity in the presence of glucose [repressed] (values obtained using the same method of Kraus and Hillen, 1997).

[d] Values obtained for repression of glucose for pCTCRegB divided by the same values obtained for pCTCRegBCm.

[e] Values obtained for repression of sucrase induction for pCTCRegBCm minus the same values obtained for pCTCRegB.

[f]Growth phases: early (OD_{600} < 0.5), mid (OD_{600} >0.5<0.75) and late (OD_{600} > 1.0).

Thus it was observed that overexpression of RegB on a plasmid construct in the early exponential growth phase enhances *scrA* activity thereby causing a modest increase in the efficiency of catabolite repression.

These differences in glucose repression were not due to the effects of the plasmid vector since the *C. beijerinckii* control contained the same vector with a disrupted *regB* gene. We assume that the increase in the number of RegB molecules led to the enhancement and increase in the regulatory control of *scrA* at the transcriptional level. Although this assumption would still fit in with the proposed model of CcpA function in CCR (Figure 1.5), these results are in contrast to those of Hueck et al., 1995 where it was found that overproduction of CcpA leads to decreased CCR efficiency probably due to titration of co-factors. An overexpressed CcpA could titrate out these co-factors to the extent that only one subunit in the CcpA dimer would be activated by binding the cofactor and would be insufficient for mediation of CCR. The interference of RegB in *C. beijerinckii* was more pronounced in the early exponential phase and may not be significant in the mid or late-exponential phase. The early exponential phase of *C. beijerinckii* is characterized by high glycolytic activity with glucose being the major energy source allowing for full glucose repression. The late exponential phase is characterized by low glycolytic activity due to depletion of energy sources. The effects of overproduction of RegB would therefore more likely occur in a state of high glycolytic activity with decreasing activity as growth time increases and glucose depletes in the environment.

Overproduction of RegB was dependent on two factors, firstly, the copy number of the plasmid vector and secondly, expression from the *scrARBK* promoter. Negative

transdominance studies in *B. subtlis* (Kraus and Hillen, 1997) used plasmid vectors containing multiple copies of *ccpA* encoded mutants and an inducible *xylA-lazZ* fusion which was subject to strong glucose repression. In this study, the amount of RegB protein should be directly correlated to the copy number of plasmid pCTC1 since each plasmid encoded a single *regB* gene. Previous results indicated that the *scrARBK* operon is subject to tight transcriptional control by glucose (Table 4.2). It is therefore possible that a further increase in CCR efficiency could be elicited by stronger expression of RegB on a plasmid construct. Alternatively, greater expression could result in decreased CCR efficiency as obtained by Heuck et al., 1995.

4.4.4 Overexpression of RegB protein in *E. coli*

Initial attempts to express the *regB* gene in pQE30, 31 and 32 expression vectors did not succeed as no protein overexpression was observed on PAGE gels although the correct DNA sequence was confirmed. The *regB* gene was therefore subcloned into plasmid pGDR11, a His-tag expression vector, and DNA sequencing confirmed the correct reading frame. The RegB protein was successfully expressed in *E. coli* and was purified using a histidine affinity column (see Figure 4.4). The purified protein bands ran just below the 43 kD MW marker.

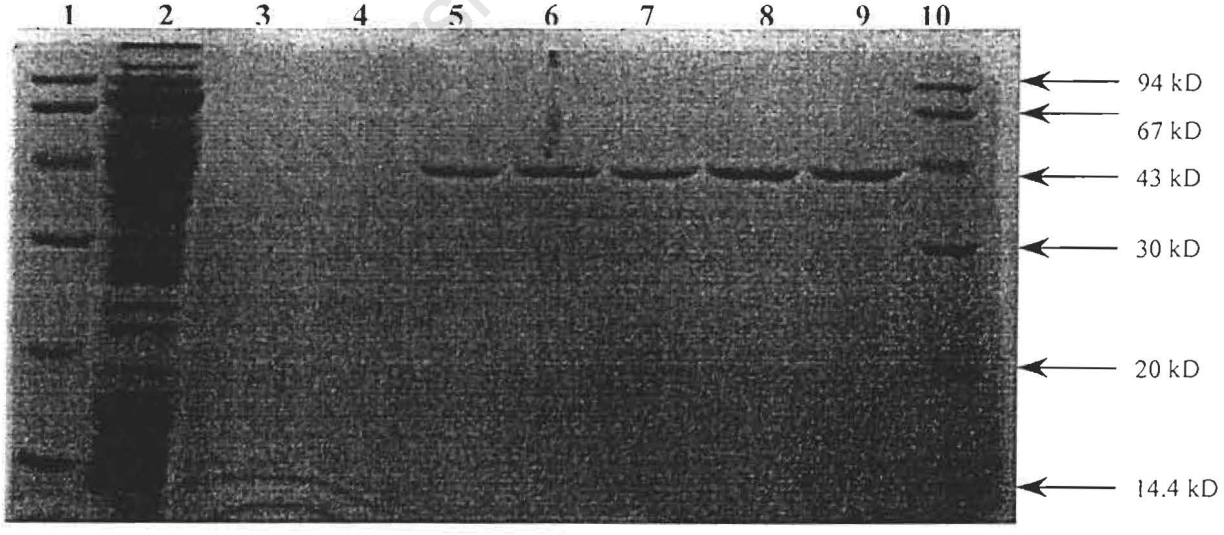

Figure 4.4: SDS-PAGE of cell-free extracts from *E. coli* containing plasmid pQEG1RegB. Lane 2 is the crude protein extract before loading the Ni column. Lanes 3 and 4 are wash buffer A and B respectively which indicate that all the histidine tagged proteins were bound to the column. Lanes 5 to 6 represent protein pools 7, 8, 9, 10 and 11 respectively which were eluted with bufferC. Lanes 1 and 10 are the protein molecular weight markers.

This result confirmed that plasmid pQEG1RegB containing the *regB* gene produces a protein product with a MW of ~ 40 kD corresponding to the deduced RegB protein MW of 39 840 Da (see Chapter 2). It will therefore be possible to purify RegB protein for future gel-binding studies.

4.5 CONCLUSION

There was physiological evidence of glucose repression of the *scrARBK* operon when *C. beijerinckii* was grown in glucose and sucrose (Leat, 1997). This study therefore sought to demonstrate specific glucose repression of the *scrARBK* operon at the transcriptional level in the absence of a functional *scrR* gene. It was observed that there was a difference at the transcriptional level in the expression of the upstream gene, *scrA*, in wild-type *C. beijerinckii* as compared with CBSCRR, the *C. beijerinckii scrR* mutant, when grown in glucose and glucose with sucrose. This suggested that the *scrARBK* operon was subject to glucose repression in addition to negative regulation by ScrR.

Primer extension and nucleotide sequence analysis of the *scrARBK* operon revealed a single transcriptional start at an adenine nucleotide, 44 bp upstream of the translational start of the operon. The putative *cre* element ($G_6A_5A_4A_3A_2C_1/G'_1G'_2T'_3T'_4T'_5C'_6$), *cre*I, downstream of the −10 region was previously identified in the *scrARBK* operon (Leat, 1997). A second imperfect palindromic sequence of 22 bp, situated 139 bp before the transcriptional start was identified. The imperfect palindromic sequence, *cre*II, resembled a typical GalR-LacI operator sequence centered around a ($A_4A_3N_2C_1/G'_1N'_2T'_3T'_4$) consensus (Weickert and Adhya, 1992). This dyad symmetry would allow members of the GalR-LacI protein family to bind as dimers to its target DNA.

The *scrARBK cre*I element had the highest G+C content, especially in the flanking sequences, when compared to other *cre* elements and the sequence position of the *scrARBK cre* was similar to that of the *B. subtilis amyE cre* site. It was previously demonstrated in *B. subtilis* that *cre* sites situated in the promoter region are not subject to CCR enhancement by Mfd and that higher levels of CCR were observed by *cre* sites

flanked by A+T rich sequences rather than by G+C nucleotides. (Zalieckas et al., 1998a, 1998b). Reid et al. (1999) also demonstrated that the *scrARBK* operon is only induced by sucrose and noted that there was suboptimal spacing between the proposed ribosome-binding site and initiation codon in *scrR*, which would limit translation yield of *scrR*, in order to prevent large amounts of ScrR from being produced and conceivably repressing *scrARBK* transcription even if the inducer is present. It is possible based on the above observations that the *scrARBK cre*I site is not subject to modulation by a transcription-repair factor and the CCR of the *scrARBK* operon may be influenced by the flanking G and C nucleotides resulting in possible inefficient binding of regulatory proteins at the *scrARBK* operon. This would in theory allow optimal sucrose induction and CCR of the *scrARBK* operon by preventing total repression by either ScrR or RegB via both the lowered translational yield of ScrR and the G+C content in the *scrARBK cre*I flanking regions.

Since both *scrARBK cre* sites resemble members of the GalR-LacI operator sequence, these cis sequences could act as binding sites for either ScrR or RegB. It is possible that the mechanism of regulation for the *scrARBK operon* in the absence of sucrose could occur in a number of ways. Firstly, ScrR could bind to one of the *scrARBK cre* elements reducing transcription of the *scrARBK* operon. In the presence of sucrose. In the presence of both glucose and sucrose, RegB modulated by phosphorylated HPr (or other co-factors), could bind to the other *cre* element and thereby maintain this repression until the glycolytic status falls to levels which relieve binding of RegB/HPr sites and allow transcription of the *scrARBK* operon. Secondly, a specific *scrARBK cre* site could act as a target for both ScrR and RegB independently depending on the glycolytic status of the cell environment. When no sucrose is present, ScrR could act as the main repressor of transcription whereas when both glucose and sucrose are present, RegB could act at the same *cre* site to maintain this repression till the glycolytic status favours transcription. Lastly, both ScrR and RegB could bind together to one or both of the *scrARBK cre* sites in a mechanism involving DNA looping resulting in regulation of the operon.

Over-expression of RegB in *C. beijerinckii* using the pCTC1 vectors demonstrated a physiological link between RegB activity and CCR regulation of the sucrose operon. This was indicated by an increase in the efficiency of glucose repression of the sucrose hydrolase activity and this was most pronounced in the early exponential growth phase, possibly because this growth phase is characterized by high glycolytic activity when glucose is the main energy source. The increase in glucose repression of the *scrARBK* operon is consistent with the role of RegB as a CcpA homologue.

CHAPTER FIVE

GENERAL CONCLUSIONS

The sucrose operon of *C. beijerinckii*, *scrARBK* was previously isolated and characterized (Reid et al., 1999). The *scrARBK* operon was transcriptionally induced by sucrose and negatively auto-regulated by ScrR, a member of the GalR-LacI family of transcriptional regulators. Physiological studies indicated that the *scrARBK* operon was subject to catabolite repression by glucose but the mechanism was not established (Leat, 1997). It was proposed that ScrR negatively regulated the transcription of the operon by binding to the promoter region at an imperfect palindromic sequence. This palindromic sequence resembled the *cre* elements identified in the operons of other Gram-positive bacteria and it is possible that other members of the GalR-LacI repressor family, specifically CcpA, could also regulate *scrARBK* gene expression since this regulator acts at *cre* elements. CcpA had been previously identified as the global transcriptional regulator of CCR for catabolite genes in *B. subtilis* and other Gram-positive species (Stülke and Hillen, 2000). In order to determine whether ScrR was the only regulator responsible for regulation of the *scrARBK* operon, we compared expression of *scrA* at the transcriptional level in wild-type *C. beijerinckii* and a *C. beijerinckii scrR* mutant after growth in glucose and glucose plus sucrose. A difference in the *scrA* expression level was observed in the *scrR* mutant, indicating that the *scrARBK* operon was subject to glucose repression in addition to regulation by ScrR.

Therefore, to gain a better understanding of carbon catabolite repression of the sucrose operon in *C. beijerinckii*, a gene, *regB*, was isolated and characterized. The protein encoded by *regB* was shown phylogenetically to belong to the CcpA sub-family of the GalR-LacI family of transcriptional regulators. A protein encoded by the *ccpA* homologue from *C. acetobutylicum* NCP262, *regA*, had previously been shown to complement a *B. subtilis ccpA* mutant (Davison et al., 1995). Amino acid sequence analysis revealed that both RegA and RegB conserved key structural and functional amino acid residues when compared to the CcpA sub-family (Kraus et al., 1998; Chapter 2), verifying that RegA and RegB are homologues of CcpA and suggesting that

these CcpA homologues are global regulators of CCR in *C. acetobutylicum* NCP262 and *C. beijerinckii* respectively.

An isoleucyl tRNA-synthetase gene was identified upstream the *regB* gene in *C.beijerinckii*. The isoleucyl tRNA-synthetase gene was also identified upstream of the *ccpA* homologues for *C. acetobutylicum* NCP262 and *C.acetobutylicum* ATCC824 (www.ncbi.nlm.nih.gov/Microb_blast/unfinishedgenome.html). Mahr et al. (2000) found that lactic acid bacteria and bacilli possess a characteristic *ccpA* gene order with the former group having the [*pepQ*]-[*ccpA*]-[variable] gene order while the latter group possess the [*aroA*]-[*ccpA*]-[variable]-[*acuC*] gene order. This study proposed that *C. beijerinckii*, *C. acetobutylicum* NCP262 and *C. acetobutylicum* ATCC824 possess a characteristic *ccpA* genetic structure, namely the [isoleucyl tRNA-synthetase]-[*ccpA*]-[variable] gene order. It is very likely that taxonomic groups or sub-groups within the clostridial species may also share a characteristic gene order for their respective *ccpA* regions. Characteristic *ccpA* genetic contexts would allow for easier search and identification of *ccpA* homologues in other clostridial species. However, further sequence analysis is required of the relevant regions in other clostridial species to confirm the prevalence of unique *ccpA* genetic arrangements.

A gene encoded a putative transcriptional regulator belonging to the MarR family of transcriptional regulators was identified immediately downstream of the *regB* gene. The MarR family of regulators are involved in a number of functions such as antibiotic resistance, sporulation, protease production and anaerobic aromatic compound degradation (Sulavik et al., 1997; Koide et al, 1999; Egland and Harwood, 1999). Downstream of the gene encoding the putative transcriptional regulator was a truncated gene, which encoded a putative protein with identity to an uncharacterized membrane protein family (UPF0013) whose function is currently unknown. No promoter region could be identified for this gene and its translational start appears to fall within the gene encoding the putative transcriptional regulator. This suggests that the genes that encode the Mar-like regulator and the membrane protein form part of an operon, but confirmation of this will require further downstream sequencing and northern blot analysis. A *cre* element could not be identified in the sequence of these downstream

genes, which suggests that these genes or operon are not subject to regulation by RegB. However, *cre* elements have been found to occur in various positions within operons (Zalieckas et al., 1998b) and complete sequence characterization of downstream genes is necessary in order to determine if RegB is involved in regulation of this genetic region.

Attempts were made to inactivate the *regB* gene in order to demonstrate functional identity of RegB at the physiological level and to determine if there was a physiological link with the *scrARBK* operon. Inactivation of *regB* using the pMTL30 suicide system (Wilkinson and Young, 1994) resulted in possible disruption of the *regB* gene as suggested by Southern hybridization analysis. The plasmid was established on the *C. beijerinckii* chromosome in tandem copies by multiple integration of the construct. However, the *C. beijerinckii regB* mutant was lost upon subsequent sporulation and germination. The loss was most likely caused by negative growth effects of the mutant as observed in *B. subtilis* (Muscariello et al., 2001) which favoured a reversal of the integration event. In this regard, this was similar to findings in *L. plantarum* where *ccpA* mutants generated by 'sco' recombination events were found to be unstable and *ccpA* mutants generated by 'dco' recombination had a reduced growth rate (Muscariello et al., 2001). A suicide vector for 'dco' recombination events, pMTL30RegBCm (refer to Chapter 3), was constructed based on chloramphenicol selection but no transconjugant for this construct was obtained. Chloramphenicol has proven problematic for use in selection of *C. beijerinckii* recombinants where antibiotic resistance is a frequent problem due to enzymatic inactivation of the antibiotic by either reduction of the aryl group or O-acetylation (Minton et al., 1993a; Staudenbauer and Dubbert, 1993). It therefore should be possible to obtain *C. beijerinckii regB* mutants via 'dco' recombination using the pMTL30 system, either by employing alternative selective markers or by re-construction of the suicide constructs, if steps are taken to minimize negative growth effects upon selection.

Despite evidence for recombination and excision events when the pG$^+$host system was used in *C. beijerinckii*, there was no inactivation of the *regB* gene. This could have been due to the nature of the *C. beijerinckii* DNA insert on the pG$^+$host vector, factors

associated with plasmid stability or sub-optimal thermosensitivity control of the suicide system in *C. beijerinckii*. The pG$^+$host system therefore requires that the problems associated with electroporation variability, thermosensitivity, vector construction and plasmid stability be addressed before it can be used as a useful alternative suicide system for *C. beijerinckii*. Dunny et al. (1991) developed an electroporation protocol for *E. faecalis* through a systematic examination of various parameters including growth conditions, composition of the electroporation solution, field strength, characteristics of plasmid DNA and conditions for selection of the transformants. Szostkova et al. (1999) also observed that for *Salmonella typhimurium* the survival of the bacteria after a high-voltage electrical pulse was influenced by the growth phase. A similar study may be required in order to optimize conditions and remove the variability associated with the *C. beijerinckii* electroporation procedure.

An experiment based on the negative transdominance strategy (Kraus et al, 1998; Heuck et al., 1995) was employed in order to demonstrate a physiological link between RegB expression and the *scrARBK* operon. It was demonstrated previously using this system that overproduction of CcpA leads to a decrease in CCR efficiency probably as a result of titration of co-factors (Heuck et al., 1995). Overexpression of RegB in *C. beijerinckii* using vector pCTC1 resulted in an increase in the CCR efficiency of sucrose hydrolase activity in the early exponential *C. beijerinckii* growth phases, thereby suggesting a link between RegB activities to CCR regulation of the sucrose operon. This result would be consistent with the role of RegB as a CcpA homologue. Future work could concentrate on confirming and defining key RegB amino acid residues by site-directed mutagenesis and specific deletions within the *regB* gene and monitoring the physiological consequences upon overexpression in *C. beijerinckii*.

Primer extension and sequence analysis of the *scrARBK* upstream region revealed a single transcriptional start site and a 22 bp imperfect palindromic sequence, *cre*II, 139 bp before the transcriptional start. This imperfect palindrome resembled a typical GalR-LacI operator sequence with dyad symmetry for binding by members of the GalR-LacI protein family. The *cre* site identified by Leat (1997), re-named *cre*I, which was situated in the *scrARBK* promoter region differed by only 2 bp from the consensus *cre* element

(Weickert and Chambliss, 1990) which is specifically recognized by CcpA (Ramseier et al., 1995). This *cre*I element had the highest G+C content when compared to other *cre* sequences, whereas *scrARBK cre*II site differed by 7 bp from the consensus *cre* element and had very low G+C content.

It is possible that the binding of a regulator to the *scrARBK cre* site may be sub-optimal due to the flanking G+C nucleotides, as seen with the *Bacillus subtilis acsA cre* site (Zalieckas et al., 1998b). Reid et al (1999) also noted that there was sub-optimal spacing between the proposed *scrARBK* ribosome binding site and the initiation codon in *scrR*, which would limit the translational yield of ScrR and prevent undue repression of the operon.

There are several hypothetical mechanisms whereby CCR of the *scrARBK* operon in *C. beijerinckii* may occur (Figure 5.1). The presence of sucrose as the sole carbon source could cause the glycolytic status of *C. beijerinckii* not to favour the interaction of RegB with HPr, and binding to the *scrARBK cre* sites (I or II) would therefore not take place. Furthermore, if the *scrARBK cre*I site was the target for RegB, the flanking G+C nucleotides of the *scrARBK cre*I site could help prevent RegB from binding to this *cre* site tightly and inhibiting or interfering with transcription of the operon when sucrose is present. Also, the putative inducer (sucrose-6-P) generated by sucrose metabolism could interact with ScrR resulting in relief of repression and transcription of the operon (Figure 5.1). The sub-optimal ribosome spacing prevents overproduction of ScrR, which could conceivably repress *scrARBK* transcription even when the inducer is present.

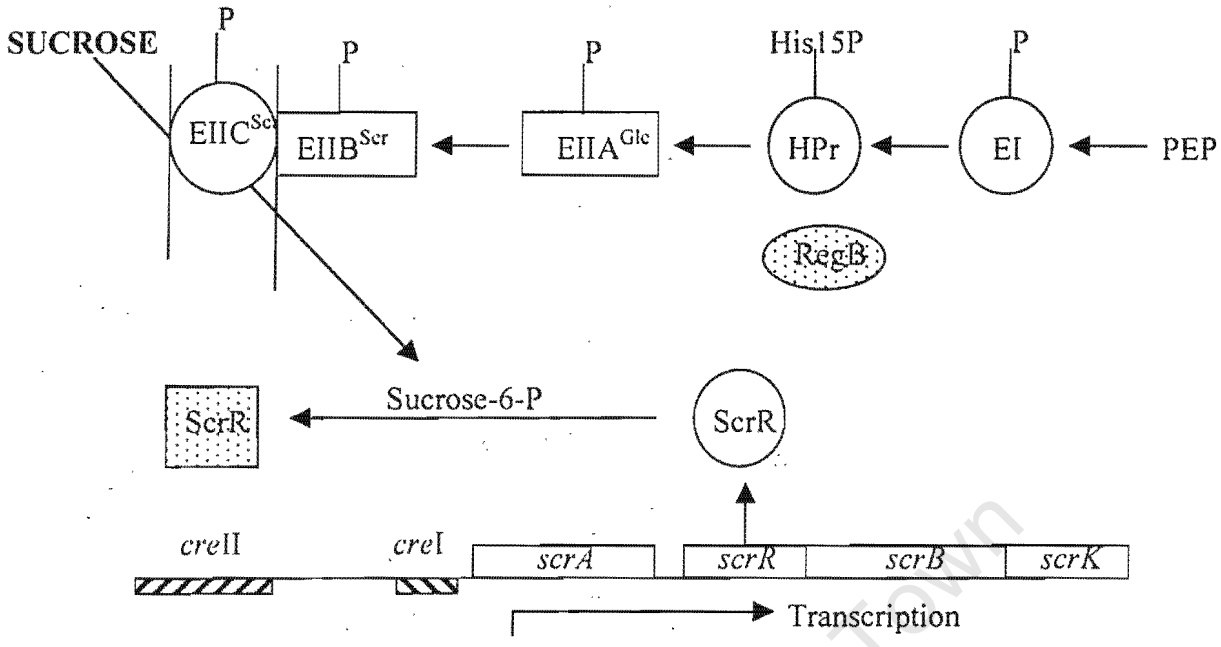

Figure 5.1: Schematic representation of a hypothetical mechanism for relief of CCR of the *scrARBK* operon of *C. beijerinckii* when sucrose is sole carbon source. The organization of the *C. beijerinckii* sucrose specific EII^{Scr} proteins of the PTS is similar to that found in *B. subtilis* where the EIIB and EIIC domains are fused to form an EIIBC^{Scr} protein (Leat, 1997). The *C. beijerinckii* EIIB^{Scr} domain is thought to be phosphorylated by EIIA^{Glc} (Reid et al., 1999; Tangney et al., 1998). When sucrose is present the putative inducer, sucrose-6-phosphate (Leat, 1997), interacts with ScrR repressor preventing ScrR from binding to the putative *scrR* operator sequence. The induction of the sucrose PTS favours the phosphorylation of HPr at His15, which does not interact with RegB and cannot therefore form a functional complex capable of binding to the *scrARBK cre* element. The flanking G+C nucleotides around the *cre* element may prevent RegB from binding independently to the *scrARBK cre* site. The end result is optimal transcription of the sucrose operon.

When there is an absence of sucrose with glucose as the sole carbon source, negative regulation of the *scrARBK* operon could occur by ScrR binding to one of the *cre* sites (Figure 5.2- A). Glycolysis would activate HPr kinase in the presence of ATP and result in the formation of HPr-Ser46-P (or other co-factors), which would modulate RegB to bind to the other *scrARBK cre* site resulting in complete repression of the operon (Figure 5.2- A). This regulatory mechanism is similar to CCR of the *B. subtlis* xylose operon where in the absence of xylose and in the presence of glucose the XylR repressor binds to upstream operator sequences and the CcpA/HPr-Ser46-P complex binds to a *cre* site situated located within the coding region of *xylA* (Dahl and Hillen,

1995). Alternatively, ScrR could bind to a specific *cre* site in the presence of both carbon sources and the level of the inducer would determine the affinity of binding (Figure 5.2- B). The presence of the inducer could then relieve ScrR binding to *cre* once a certain inducer level is reached but the presence of glucose allows RegB/HPr-Ser46-P to maintain repression by binding to the same specific *cre* site.

Irrespective of which *scrARBK cre* elements are the binding target sites for ScrR or RegB, the RegB/HPr-Ser46-P complex could still maintain glucose repression since HPr-Ser46-P has been implicated in inducer expulsion due to activation of sugar phosphatases (Saier et al., 1996). Another level of maintaining CCR would be inducer exclusion as a result of competition for $EIIA^{Glc}$, which would preferentially facilitate glucose uptake over sucrose uptake. Once the glycolytic status falls below a certain level, the presence of PTS sugars will favour dephosphorylation of HPr-Ser46-P and phosphorylation of the HPr histidyl residue. This would dissociate the RegB binding complex from the *cre* site(s) and result in transcription of the *scrARBK* operon.

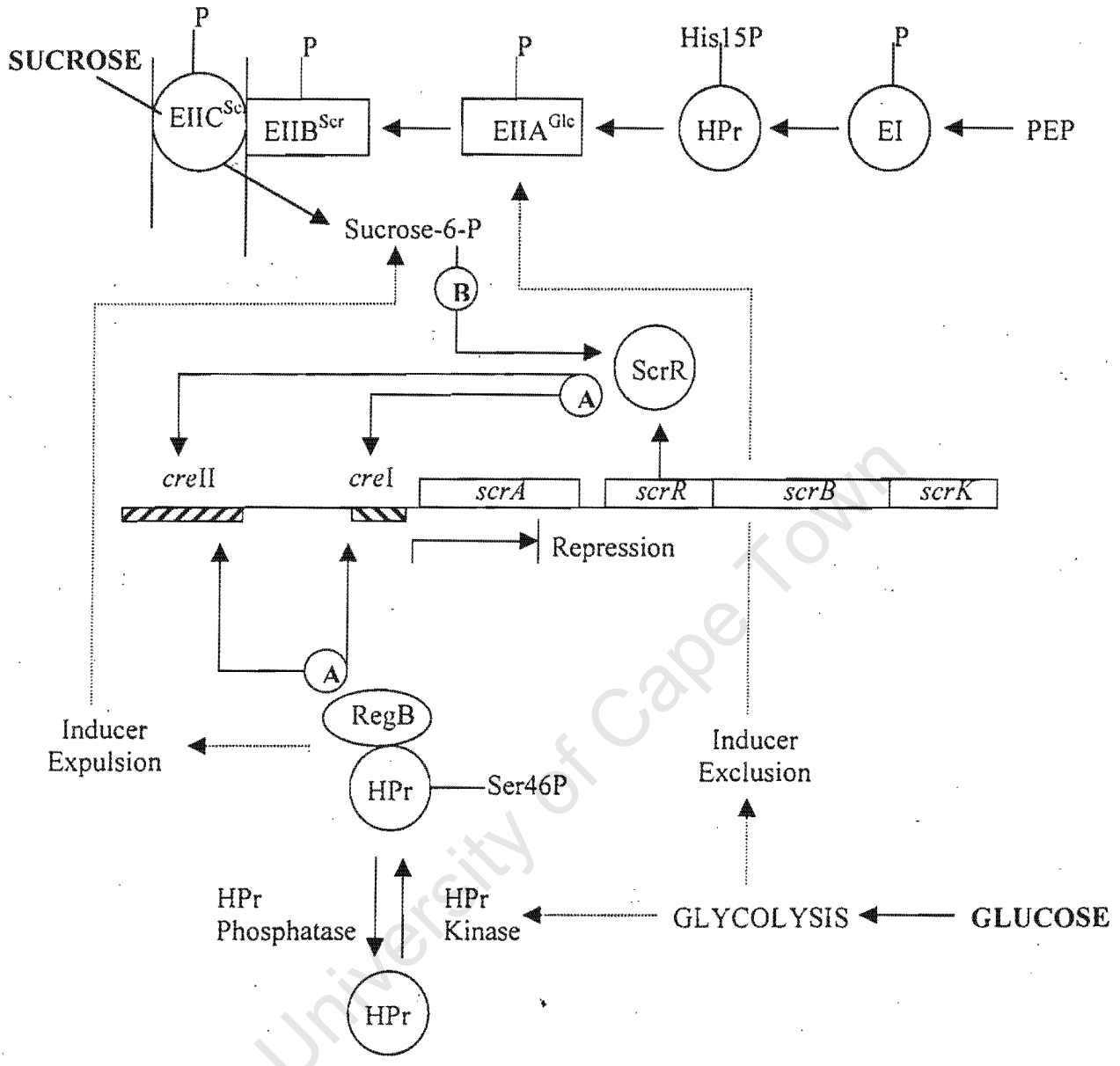

Figure 5.2: Schematic representation of a possible hypothetical mechanism for CCR of the *scrARBK* operon of *C. beijerinckii* when glucose and sucrose are present as carbon sources. (A) In glucose medium, the absence of an inducer allows negative regulation of the *scrARBK* operon by ScrR binding to one of the two *scrARBK cre* sites (I or II). Repression is maintained when glucose is present as glycolysis activates HPr kinase (in the presence of ATP) to phosphorylate HPr at Ser46, which interacts with RegB and this complex binds to either *scrARBK cre* site. (B) When both glucose and sucrose are present, ScrR binding to the *cre* site will depend on the level of inducer, and if sufficient inducer molecules relieve ScrR *cre* binding, the level of glucose will still allow RegB/HPr-Ser46-P to maintain repression by binding to either *scrARBK cre* site. The glycolytic status could be maintained when sucrose is present as HPr-Ser46P could activate sugar phosphatases as a prelude to inducer (sucrose-6-P) expulsion and $EIIA^{Glc\text{-}P}$ would preferentially phosphorylate glucose transport over sucrose transport (inducer exclusion).

It is also possible that both ScrR and RegB could bind at a single *scrARBK cre* site or at both *cre* sites at the same time in a manner that may involve specific binding domains and DNA looping. Lastly it is possible that both *scrARBK cre* sequences may function independently from each other in CCR in order to recognize different glycolytic intermediates as found for the two *cre* elements of the *B. subtilis gnt* operon (Miwa et al., 1997). In the *gnt* operon, a *cre* element was identified in the *gntR* gene, (cre_{down}) and in the promoter region (cre_{up}). The experimental evidence implied that CCR exerted by cre_{up} was probably independent of cre_{down} and involved CcpA. The study found that CCR exerted by cre_{up} was independent of HPr-Ser-P and was not recognized by the CcpA/HPr-Ser-P complex, in contrast to cre_{down}, which was specifically recognized by the complex and was partially independent of HPr-Ser-P. Interestingly, CcpA in complex with glucose-6-P specifically recognized both *B. subtilis gnt cre* sites. In a similar manner, RegB could, by forming complexes with various glycolytic intermediates during metabolism, convey the glycolytic status of the cell environment to the *scrARBK* operon via the *scrARBK cre* sites.

In order to determine the mechanism of CCR for the *scrARBK* operon, gel mobility shift experiments involving ScrR, RegB and the *scrARBK cre*-like sequences must first be done to distinguish and identify the respective target sites for these regulators. Once this has been confirmed, the phosphorylated derivatives of sucrose, glucose and fructose must be tested for their ability to prevent ScrR interacting with the respective palindromic site in order to identify the inducer molecule. The use of HPr mutants in complementation studies would also confirm co-factors for RegB. Deletion or mutation of the *scrARBK cre*I and *cre*II sites would also help clarify the specific role of this element in negative regulation and CCR of the *scrARBK* operon. Finally, renewed attempts must be made to obtain a RegB mutant in order to identify other catabolite genes subject to RegB regulation.

Finally, conserved key amino acid residues were previously identified by Kraus et al. (1998) in RegA, and a *B. subtilis ccpA* mutant was successfully complemented by the protein encoded by the *C. acetobutylicum regA* gene (Davison et al., 1995). We present new evidence in this thesis namely:

(i) The presence of the conserved aa residues of the CcpA sub-family in RegB.

(ii) The possible negative growth effects of a *regB* mutation.

(iii) Possible catabolite repression of *scrARBK* operon in the absence of the *scrR* gene.

(iv) Identification of *cre* elements in the *scrARBK* operon.

(v) Overexpression of the *regB* gene affects the expression of the *scrARBK* operon *in vivo*.

This study therefore proposes that the *regA* and *regB* genes are indeed homologues of the *ccpA* gene found in many Gram-positive bacteria and that the nomenclature of *regA* and *regB* is changed to *ccpA*.

APPENDIX A

RESTRICTION ANALYSIS

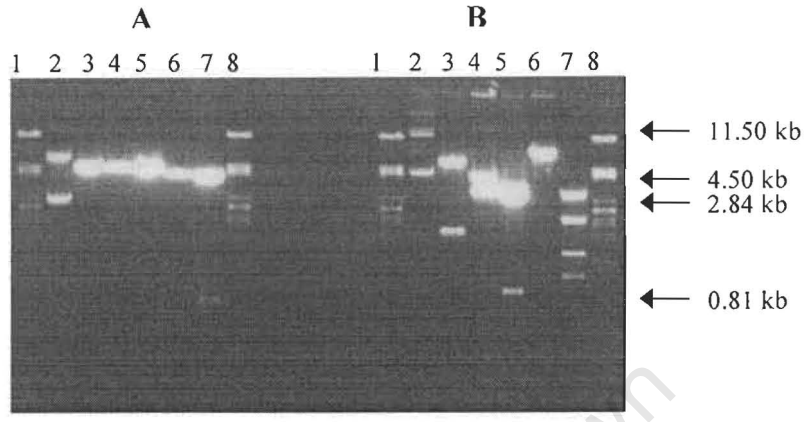

Figure A.1: DNA Restriction analysis of plasmids pMTL30RegB [A] and pGAhost2 [B].

[A]- pMTL30RegB

Lane 2: Uncut pMTL30RegB DNA

Lane 3: Digested with *Xba*I

Lane 4: Digested with *Xba*I/*EcoR*I

Lane 5: Digested with *EcoR*I

Lane 6: Digested with *EcoR*I/*Pst*I (A faint ~0.5 kb band, was observed in this lane)

Lane 7: Digested with *Xba*I/*Pst*I

[B]- pGAhost2

Lane 2: Uncut pGAhost2 DNA

Lane 3: Digested with *Cla*I

Lane 4: Digested with *Xba*I

Lane 5: Digested with *EcoR*V/*Xba*I

Lane 6: Digested with *EcoR*V

Lane 7: Digested with *Xmn*I

Lanes 1 and 8 in both [A] and [B] is Lambda marker (λ DNA digested with *Pst*I)

APPENDIX B

CLONING VECTORS

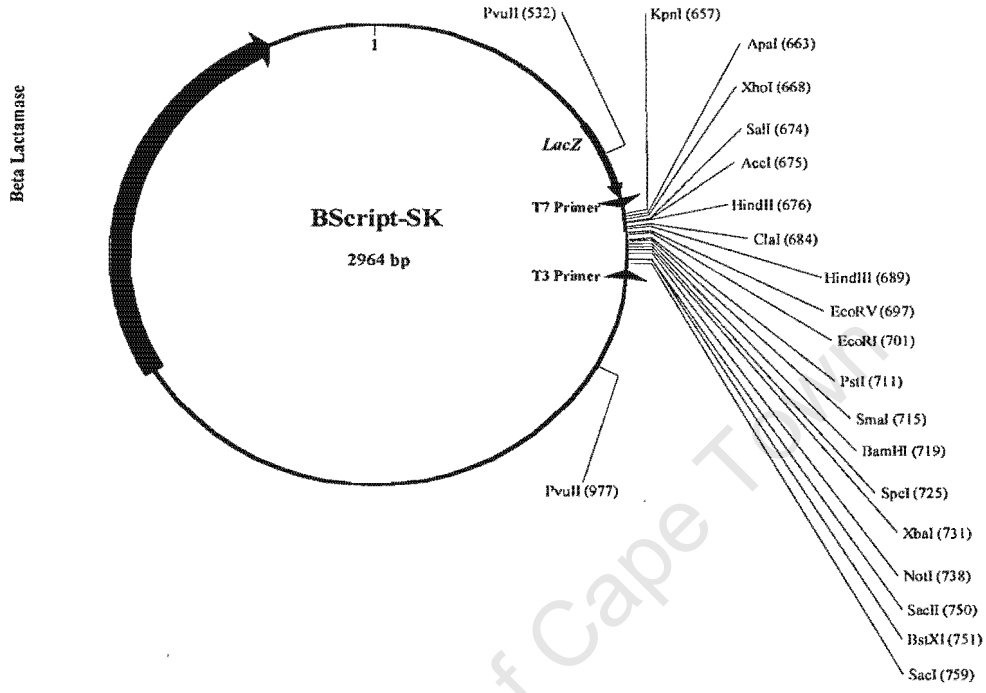

Figure B.1: Restriction map of pBluescriptSK$^+$ (Stratagene, La Jolla, California)

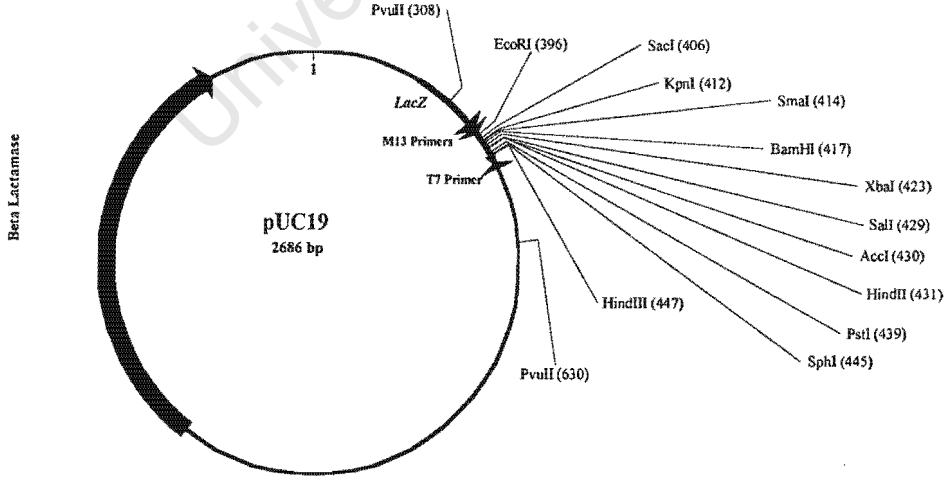

Figure B.2: Restriction map of pUC19 (Roche Diagnostics).

LITERATURE CITED

Achen, M.G., B.E. Davidson and A.J. Hillier. 1986. Construction of plasmid vectors for the detection of streptococcal promoters. Gene. **45**: 45-49.

Adler, H.I. and W. Crow. 1987. A technique for predicting the solvent-producing ability of *Clostridium acetobutylicum*.Appl. Environ. Microbiol. **53**: 2496-2499.

Alekshun, M. N., and S. B. Levy. 1997. Regulation of chromosomally mediated multiple antibiotic resistance: the *mar* regulon. Antimicrob. Agents Chemother. **10:** 2067-2075.

Alekshun, M.N. and S.B. Levy. 1999. Characterization of MarR Superrepressor mutants. J. Bacteriol. **181**(10): 3303-3306.

Annous, B.A. and H.P. Blaschek. 1991. Isolation and chracterization of *Clostridium acetobutylicum* mutants with enhanced amylolytic activity. Appl. Environ. Microbiol. **57**:2544-2548.

Archer, G. L., and D. M. Niemeyer. 1994. Origin and evolution of DNA associated with resistance to methicillin in staphylococci. Trends Microbiol. **2**:343-347.

Armitage, P., R. Walden and J. Draper. 1988. Vectors for the transformation of plant cells using agrobacterium. In: *Plant genetic transformation and gene expression, a laboratory manual*. J. Draper, R. Scott, P. Armitage and W. Walden (eds). Blackwell Scientific Publications ltd. Oxford.

Ausubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith and K. Struhl. 1989. *Current Protocols in Molecular Biology*, vol. 1. New York: Wiley.

Bahl, H. 1993. Heat shock response and onset of solvent formation in *Clostridium acetobutylicum*. In: *The Clostridia and Biotechnology*. D.R. Woods (ed.). **pp.** 247-258. Butterworth-Heinemann. London.

Biswas, I., A. Gruss, S.D. Ehrlich and E. Maguin. 1993. High-efficiency gene inactivation and replacement system for gram-positive bacteria. J. Bacteriol. **175**(11): 3628-3635.

Blaschek, H.P. and B.A. White. 1995. Genetic systems development in the clostridia. FEMS Micro. Rev. **17**: 349-356.

Blatch, G.L., R.R. Scholle and D.R. Woods. 1990. Nucleotide sequence and analysis of the *Vibrio alginolyticus* sucrose uptake-encoding region. Gene. **95**: 17-23.

Blatch, G.L. and D.R. Woods. 1993. Molecular characterization of a fructanase produced by *Bacteroides fragilis* BF-1. J.Bacteriol. **175**: 3058-3066.

Bolivar, F., R.L. Rodrigeuz, P.J. Greene, M.C. Betlach, H.L. Heyneker, H. Boyer, J.H. Crosa and S. Falkow. 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene 2: 95-113.

Bowring, S.N. and J.G. Morris. 1985. Mutagenesis of *Clostridium acetobutylicum*. J. Appl. Bacteriol. **58**:577-584.

Bradford, M.M. 1976. A rapid and sensitive method for quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**: 248-254.

Brehm, J.K., G.P.C. Salmond and N.P. Minton. 1987. Sequence of the adenine methylase gene of *Streptococcus faecalis* plasmid pAM_1. Nucl. Acids Res. **15**: 3177.

Brendal, V. and E.N. Trifonov. 1984. A computer algorithm for testing potential prokaryote terminators. Nucleic Acid Res. **12**:4411-4427.

Bruand, C., S.D. Ehrlich and L. Jannière. 1991. Unidirectional theta replication of the structurally stable *Enterococcus faecalis* plasmid pAM_1. EMBO. **10**(8): 2171-2177.

Cases, I. and V. de Lorenzo. 1998. Expression systems and physiological control of promoter activity in bacteria. Current Opinion in Microbiol. **1**(3): 303-310.

Chambers, S.P., S.E. Prior, D.A. Barstow and N.P. Minton. 1988. The pMTL*nic*˙ cloning vectors. Improved pUC polylinker regions to facilitate the use of sonicated DNA for nucleotide sequencing. Gene. **68**: 139-149.

Chambliss, G.H. 1993. Carbon source-mediated catabolite repression. In: *Bacillus subtilis and Other Gram-Positive Bacteria: Biochemistry, Physiology and Molecular Genetics*. A.L. Sonenshein, J.A. Hoch and R. Losick (eds.). pp. 213-219. Washington DC: American Society for Microbiology.

Chauvaux, S., I.T. Paulsen and M.H. Saier. 1998. CcpB, a novel transcription factor implicated in catabolite repression in *Bacillus subtilis*. J. Bacteriol. **180**(3): 491-497.

Claassen, P.A., M.A. Budde and A.M. Lopez-Contreras. 2000. Acetone, butanol and ethanol production from domestic organic waste by solventogenic clostridia. J. Mol. Microbiol. Biotechnol. **2**(1): 39-44.

Compere, A. and W.L. Griffith. 1979. Evaluation of substrates for butanol production. Dev. Ind. Microbial. **20**:509-517.

Cummins, C.S. and J.L. Johnson. 1971. Taxonomy of the clostridia: wall composition and DNA homologies in *Clostridium butyricum* and other butyric acid-producing clostridia. J.Gen.Microbiol. **67**: 33-46.

Dahl, M.K. and W. Hillen. 1995. Contributions of XylR, CcpA and HPr to catabolite repression of the *xyl* operon in *Bacillus subtilis*. FEMS Micro. Lett. **132**: 79-83.

Dahl, M.K., D. Schmiedel and W. Hillen. 1995. Glucose and Glucose-6-Phosphate interaction with Xyl Repressor proteins from *Bacillus* spp. may contribute to regulation of xylose utilization. J. Bacteriol. **177**(19): 5467-5472.

Davison, S.P., J.D. Santangelo, S.J. Reid and D.R. Woods. 1995. A *Clostridium acetobutylicum* regulator gene (*regA*) affecting amylase production in *Bacillus subtilis*. Microbiology. **141**(4): 989-996.

Deutscher, J., J. Reizer, C. Fischer, A. Galinier, M.H. Saier and M. Steinmetz. 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 catabolite genes of Bacillus subtilis. J. Bacteriol. **176**: 3336-3344.

Deutscher, J., E. Kuster, U. Bergstedt, V. Charrier and W. Hillen. 1995. Protein kinase-dependent HPr/CcpA interaction links glycolytic activity to carbon catabolite repression in gram-positive bacteria. Mol. Micro. **15**(6): 1049-1053.

Devereux, J., P. Haberli and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acid Res. **12**: 387-395.

Dunny, G.M., L.N. Lee and D.J. LeBlanc. 1991. Improved electroporation and cloning vector system for gram-positive bacteria. Appl. Environ. Microbiol. **4**: 1194-1201.

Dürre, P. 1998. New insights and novel developments in clostridial acetone/butanol/isopropanol fermentation. Appl. Microbiol. Biotechnol. **49**(6): 639-648.

Dürre, P., R.J. Fischer, A. kuhn, K. Lorenz, W. Schreiber, B. Stürzenhofecker, S Ullmann, K. Winzer and U. Sauer. 1995. Solventogenic enzymes of *Clostridium acetobutylicum*: catalytic properties, genetic organization, and transcriptional regulation. FEMS Micro. Lett. **17**: 251-262.

Egeter, O. and R. Brückner. 1996. Catabolite repression mediated by the catabolite control protein CcpA in *Staphylococcus xylosus*. Mol Microbiol. **21**(4): 739-749.

Egland, P.G. and C.S. Harwood. 1999. BadR, a new MarR family member, regulates anaerobic benzoate degradation by *Rhodopseudomonas palustris* in concert with AadR, an Fnr family member. J. Bacteriol. **181**(7): 2102-2109.

Faires, N., S. Tobisch, S. Bachem, I. Martin-Verstraete, M. Hecker and J. Stülke. 1999. The catabolite control protein CcpA controls ammonium assimilation in *Bacillus subtilis*. J. Mol. Microbiol. Biotechnol. **1**(1): 141-148.

Fisher, S.H. 1991. Control of Carbon and Nitrogen metabolism in *Bacillus subtilis*. Ann. Rev. Microbiol. **45**: 107-135.

Fraser, C.M., S. Casjens, W.M. Huang, G.G. Sutton et al. 1997. Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. Nature. **390**:580-586.

Fujita, Y. and Y. Miwa. 1994. Catabolite repression of the *Bacillus subtilis gnt* operon mediated by the CcpA protein. J. Bacteriol. **176**(2): 511-513.

Fujita, Y., Y. Miwa, A. Galinier and J. Deutscher. 1995. Specific recognition of the *Bacillus subtilis gnt cis*-acting catabolite-responsive element by a protein complex formed between CcpA and seryl-phosphorylated HPr. Mol Microbiol. **17**(5): 953-960.

Galinier, A., J. Haiech, M.C. Kilhoffer, M. Jaquinod, J. Stülke, J. Deutscher and I. Martin-Verstraete. 1997. The *Bacillus subtilis crh* gene encodes a HPr-like protein involved in carbon catabolite repression. Proc. Natl. Acad. Sci. **94**(16): 8439-8444.

Galinier, A., J. Deutscher and I. Martin-Verstraete. 1999. Phosphorylation of either Crh or HPr mediates binding of CcpA to the *Bacillus subtilis xyn* cre and catabolite repression of the *xyn* operon. J. Mol. Biol. 19, **286**(2): 307-314.

Gering M. and R. Brückner. 1996. Transcriptional regulation of the sucrase gene of *Staphylococcus xylosus* by the Repressor ScrR. J. Bacteriol. **178**(2): 462-469.

Gosalbes M.J., V. Monedero, C.A. Alpert and G. Perez-Martinez. 1997. Establishing a model to study the regulation of the lactose operon in *Lactobacillus casei*. FEMS Micro. Lett. **148**(1): 83-89.

Gösseringer, R., E. Kuster, A.Galinier, J. Deutscher and W. Hillen. 1997. Cooperative and non-cooperative DNA binding modes of catabolite control protein CcpA from *Bacillus megaterium* result from sensing two different signals. J. Mol. Biol. 7, **266**(4): 665-676.

Green, E.M. and G.N. Bennett. 1996. Inactivation of an aldehyde/alcohol dehydrogenase gene from *Clostridium acetobutylicum* ATCC824. Appl. Biochem. Biotech. 57: 213-221.

Green, M.E., Z.L. Boynton, L.M. Harris, F.B. Rudolph, E.T. Papoutsakis and G.N. Bennett. 1996. Genetic manipulation of acid formation pathways by gene inactivation in *Clostridium acetobutylicum* ATCC 824. Microbiology. **142**:2079-2086.

Grundy, F.J., D.A. Waters, T.Y. Takova and T.M. Henkin. 1993a. Identification of genes involved in utilization of acetate and acetoin in *Bacillus subtilis*. Mol Microbiol. 10(2): 259-271.

Grundy, F.J., D.A. Waters, S.H. Allen and T.M. Henkin. 1993b. Regulation of the *Bacillus subtilis* acetate kinase gene by CcpA. J. Bacteriol. **175**(22): 7348-7355.

Grundy, F.J, A.J. Turinsky and T.M. Henkin. 1994. Catabolite regulation of *Bacillus subtilis* acetate and acetoin utilization genes by CcpA. J. Bacteriol. **176**(15): 4527-4533.

Gruss, A. and S.D. Ehrlich. 1988. Insertion of foreign DNA into plasmids from gram-positive bacteria induces formation of high-molecular-weight plasmid multimers. J. Bacteriol. **170**: 1183-1190.

Gruss, A. and S.D. Ehrlich. 1989. The family of highly interrelated single-stranded deoxyribonucleic acid plasmids. Microbiol. Rev. **53**: 231-241

Hastings, J.H.J. 1978. Acetone-butyl alcohol fermentation. In: *Economic Microbiology, vol. 2. Primary products of metabolism*. A.H. Rose (ed.). pp. 31-45. Academic Press, Inc., New York.

Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene. **28**:351-359.

Henkin, T.M., F.J. Grundy, W.L. Nicholson and G.H. Chambliss. 1991. Catabolite repression of _-amylase gene expression in *Bacillus subtilis* involves a trans-acting gene product homologous to the *Escherichia coli lacI* and *galR* repressors. Mol. Microbiol. 5(3): 575-84.

Henkin, T.M., 1996. The role of CcpA transcriptional regulator in carbon metabolism in *Bacillus subtilis*. FEMS Microbiol. Lett. 1, 135(1): 9-15.

Hogema, B.M., J.C. Arents, R. Bader, K. Eijkemans, H. Yoshida, H. Takahashi, H. Aiba and P.W. Postma. 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 IIA^{Glc}. Mol. Microbiol. **30**: 487-498.

Holt, R.A., G.M. Stephens and J.G. Morris. 1984. Production of solvents by *Clostridium acetobutylicum* cultures maintained at neutral pH. Appl. Environ. Microbiol. **48**: 1166-1170.

Horaud, T., C. Le Bouguenec, and K. Pepper. 1985. Molecular genetics of resistance to macrolides, lincosamides and streptogramin B (MLS) in streptococci. J. Antimicrob. Chemother. **16**: 111-135.

Hueck, C., A. Kraus and W. Hillen. 1994. Sequences of *ccpA* and two downstream *Bacillus megaterium* genes with homology to the *motAB* operon from *Bacillus subtilis*. Gene. 27, **143**(1): 147-148.

Hueck, C.J. and W. Hillen. 1995. Catabolite repression in *Bacillus subtilis*: a global regulatory mechanism for the gram-positive bacteria? Mol. Microbiol. **15**(3): 395-401.

Hueck, C.J., A. Kraus, D. Schmiedel and W. Hillen. 1995. Cloning, expression and functional analyses of the catabolite control protein CcpA from *Bacillus megaterium*. Mol. Microbiol. **16**(5): 855-64.

Inada, T., K. Kimata and H. Aiba. 1996. Mechanism responsible for glucose-lactose diauxie in *Escherichia coli*: challenge to cAMP model. Genes to Cells. 1:293-301.

Ish-Horowicz, D. and J.F. Burke. 1981. Rapid and efficient cosmid cloning. Nucleic Acid Res. 9(13): 2989-2998.

Jankovic, I., O. Egter and R. Brükner. 2000. Analysis of catabolite control protein A-dependent repression in *Staphylococcus xylosus* by a genomic reporter gene system. J. Bacteriol. **183**(2): 580-586.

Janniére L., C. Bruand and S.D. Ehrlich. 1990. Structurally stable *Bacillus subtilis* cloning vectors. Gene. **87**: 53-61.

Janssen, P.J., D.T. Jones and D.R. Woods. 1990. Studies on *Clostridium acetobutylicum glnA* promoters and antisense RNA. Mol. Microbiol. 4:1575-1583.

Jennert, K.C.B., C. Tardif, D.I. Young and M. Young. 2000. Gene transfer to *Clostridium cellulolyticum* ATCC 35319. Microbiology. **146**: 3071-3080.

Johnson, J.L. and J.S. Chen. 1995. Taxonomic relationships among strains of *Clostridium acetobutylicum* and other phenotypically similar organisms. FEMS Microbiol. Lett. **17**:233-240.

Johnson, J.L., J. Toth, S. Santiwatanakul and J.S. Chen. 1997. Cultures of "*Clostridium acetobutylicum*" from various collections comprise *Clostridium acetobutylicum*, *Clostridium beijerinckii*, and two other distinct types based on DNA-DNA reassociation. Int. J. Syst. Bacteriol. **47**(2): 420-424.

Jones, D.T., A. van der Westhuizen, S. Long, E.R Allcock, S.J. Reid and D.R. Woods. 1982. Solvent production and morphological changes in *Clostridium acetobutylicum*. Appl. Environ. Microbiol. **43**: 1434-1439.

Jones, D.T and D.R. Woods. 1986. Acetone-butanol fermentation revisited. Microbiol. Rev. **50**:484-524

Jones, D.T. 1993. Mutagenesis and its application in biotechnology. In: *The Clostridia and Biotechnology*. D.R. Woods (ed.). **pp.** 77-97. Butterworth-Heinemann. London.

Jones, B.E., V. Dossonnet, E. Kuster, W. Hillen, J. Deutscher and R.E. Klevit. 1997. Binding of the catabolite repressor protein CcpA to its DNA target is regulated by phosphorylation of its corepressor HPr. J. Biol. Chem. 17, **272**(42): 26530-5.

Kashket, E.R. and Z. Cao. 1993. Isolation of a degeneration resistant mutant of *C.acetobutylicum* NCIMB 8052. Appl. Environ. Microbiol. **59**: 4198-4202.

Kashket, E.R. and Z. Cao. 1995. Clostridial strain degneration. FEMS Microbiol. Rev. **17**: 307-315.

Keis, S., C.F. Bennet, V.K. Ward and D.T. Jones. 1995. Taxonomy and phylogeny of industrial solvent-producing clostridia. Int. J. Syst. Bacteriol. **45**: 693-705

Kim, A.Y., A.A. Vertes and H.P. Blaschek. 1990. Isolation of single-stranded plasmid from *Clostridium acetobutylicum* NCIB 6444. Appl. Environ. Microbiol. **56**: 1725-1728.

Kim, A.Y. and H.P. Blaschek. 1993. Construction of a phage-plasmid hybrid (phagemid), pCAK1, containing the replicative form of virus-like particle CAK1 isolated from *Clostridium acetobutylicum* NCIB 6444. J. Bacteriol. **175**(12): 3838-3843.

Kim, J.H. and G.H. Chambliss. 1997. Contacts between *Bacillus subtilis* catabolite regulatory protein CcpA and *amyO* target site. Nucleic Acids Res. **25**(17): 3490-3496.

Kim, J.H., Z.T. Guvener, J.Y. Cho, K.C. Chung and G.H. Chambliss. 1995. Specificity of DNA binding activity of the *Bacillus subtilis* catabolite control protein CcpA. J. Bacteriol. **177**(17): 5129-5134.

Kim, J.H., M.I. Voskuil and G.H. Chambliss. 1998. NADP, corepressor for the *Bacillus* catabolite control protein CcpA. Proc. Natl. Acad. Sci. 4, **95**(16): 9590-9595.

Koide, A., M. Perego and J.A. Hoch. 1999. ScoC regulates peptide transport and sporulation initiation in *Bacillus subtilis*. J. Bacteriol. **181**(13): 4114-4117.

Komeda, H., M. Kobayashi, and S. Shimizu. 1996. Characterization of the gene cluster of high-molecular-mass nitrile hydratase (H-NHase) induced by its reaction products in *Rhodococcus rhodocrous*. Proc. Natl. Acad. Sci. USA. **93**:4267-4272.

Kraus, A., C. Hueck, D. Gärtner and W.Hillen. 1994. Catabolite repression of the *Bacillus subtilis xyl*-operon involves a *cis* element functional in an unrelated sequence context and glucose exerted additional *xylR* dependent repression. J.Bacteriol. 176: 1738-1745.

Kraus, A. and W. Hillen. 1997. Analysis of CcpA mutations defective in carbon catabolite repression in *Bacillus megaterium*. FEMS Microbiol Lett. **153**(1): 221-226.

Kraus, A., E. Kuster, A. Wagner, K. Hoffmann and W. Hillen. 1998. Identification of a co-repressor binding site in catabolite control protein CcpA. Mol. Microbiol. **30**(5): 955-963.

Kruger, S., J. Stülke and M. Hecker. 1993. Catabolite repression of beta-glucanase synthesis in *Bacillus subtilis*. J. Gen. Microbiol. **139**(9): 2047-2054.

Krüger, S. and M. Hecker. 1995. Regulation of the putative *bglPH* operon for aryl-_-glucoside utilization in *Bacillus subtilis*. J.Bacteriol. 177: 5590-5597.

Küster, E., E.J. Luesink, W.M. de Vos and W. Hillen. 1996. Immunological crossreactivity to the catabolite control protein CcpA *Bacillus megaterium* is found in many gram-positive bacteria. FEMS Microbiol Lett. 1, **139**: 109-115.

Küster, E., T. Hilbich, M.K. Dahl and W. Hillen. 1999a. Mutations in catabolite control protein CcpA separating growth effects from catabolite repression. J. Bacteriol. **181**(13): 4125-4128.

Küster, E., A. Wagner, U. Volker and W. Hillen. 1999b. Mutations in catabolite control protein CcpA showing glucose-independent regulation in *Bacillus megaterium*. J. Bacteriol. **181**(24): 7634-7638.

Kyte, J. and R.F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. **157**(1): 105-132.

Laemlli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature. **227**: 680-685.

Langlykke, A.F., J.M. Van Lanen and D.R. Fraser. 1948. Butyl alcohol from xylose saccharification liquors from corncobs. Ind. Eng. Chem. **40**:1716-1719.

Law, J., G. Buist, A. Haandrikman, J. Kok, G. Venema and K. Leenhouts. 1995. A system to generate chromosomal mutations in *Lactococcus lactis* which allows fast analysis of targeted genes. J. Bacteriol. 177(24):7011-7018.

Leat, N. 1997. Studies on the sucrose utilization system of *Clostridium beijerinckii* NCIMB 8052. PhD Thesis. University of Cape Town, South Africa.

Leboeuf, C., Y. Auffray and A. Hartke. 2000. Characterization of the *ccpA* gene of *Enterococcus faecalis*: identification of starvation-inducible proteins regulated by ccpA. J. Bacteriol. **182**(20): 5799-5806.

Lee, S.Y., G.N. Bennett and E.T. Papoutsakis. 1992. Construction of *Escherichia coli-Clostridium acetobutylicum* shuttle vectors and transformation of *Clostridium acetobutylicum* strains. Biotech. Lett. 14(5):427-432.

Lee, S.Y., L.D. Mermelstein and E.T. Papoutsakis. 1993. Determination of plasmid copy number and stability in *Clostridium acetobutylicum* ATCC824. FEMS Microbiol. Lett. **108**: 319-324.

Leehouts, K.J., J. Kok and G. Venema. 1991. Replacement recombination in *Lactococcus lactis*. J. Bacteriol. 173: 4794-4798.

Lehming, N., J. Sartorius, M. Niemöller, G. Geneger, B. Wilcken-Bergemann and B. Müller-Hill. 1987. The interaction of the recognition helix of *lac* repressor with *lac* operator. EMBO J. 6: 3145-3153.

Lengeler, J.W., R.J. Mayer and K. Schmid. 1982. Phosphoenolpyruvate-dependent phosphotransferase system enzyme III and plasmid-encoded sucrose transport in *Escherichia coli* K-12. J. Bacteriol. **151**: 468-471.

Lengeler, J.W., K. Jahreis and U.F. Wehmeier. 1994. Enzymes II of the phosphoenolpyruvate-dependent phosphotransferase systems: their structure and function in carbohydrate transport. Biochim. Biophys. Acta. **1188**: 1-28.

Lokman, B.C, M. Heerikhuisen, R.J. Leer, A. van den Broek, Y. Borsboomand, S. Chaillou, P.W Postma and P.H. Pouwels. 1997. Regulation of expression of the *Lactobacillus pentosus xylAB* operon. J. Bacteriol. 179(17):5391-5397.

Lopez-Contreras, A.M., P.A. Claassen, H. Mooibroek and W.M. De Vos. 2000. Utilisation of saccharides in extruded domestic organic waste by *Clostridium acetobutylicum* ATCC 824 for production of acetone, butanol and ethanol. Appl. Microbiol. Biotechnol. **54**(2): 162-167.

Maguin, E., P. Duwat, T. Hege, D. Ehrlich and A. Gruss. 1992. New thermosensitive plasmid for gram-positive bacteria. J. Bacteriol. 174(17): 5633-5638.

Mahr, K., W. Hillen and F. Titgemeyer. 2000. Carbon catabolite repression in *Lactobacillus pentosus*: analysis of the *ccpA* region. Appl. Environ. Microbiol. **66**(1): 277-283.

Maniatis, T., E.F. Fritsch and J. Sambrook. 1982. *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory. Cold Spring Harbor. NY.

Marchal, R., D. Blanchet D. and J.P. Vandecasteele. 1985. Industrial optimization of acetone-butanol fermentation: a study of the utlization of Jerusalem artichokes. Appl. Microbiol. Biotechnol. **23**:92-98.

Martin, F.H. and I. Jr Tinoco. 1980. DNA-RNA hybrid duplexes containing oligo(dA:rU) sequences are exceptionally unstable and may facilitate termination of transcription. Nucleic Acids Res. **8**: 2295-2299.

Martin-Verstraete, I., J. Deutscher and A. Galinier. 1999. Phosphorylation of HPr and Crh by HprK, early steps in the catabolite repression signalling pathway for the *Bacillus subtilis* levanase operon. J. Bacteriol. **181**(9): 2966-2969.

McCoy, E. and L.S. McClung. 1935. Studies on anaerobic bacteria. V. The serological agglutination of *Clostridium acetobutylicum* and related species. J. Infect. Dis. **56**: 333-346.

Mclaughlin, J.R., C.L. Murray and J.C. Rabinowitz. 1981. Unique feature in the ribosome binding site sequence of the gram-positive *Staphylococcus aureus* _-lactamase gene. J. Biol. Chem. **256**:11283-11291.

Meinecke, B., H.Bahl and G. Gottschalk. 1984. Selection of an asporogenous strain of *Clostridium acetobutylicum* in continuous culture under phosphate limitation. Appl. Environ. Microbiol. **48**: 1064-1067.

Mermelstein, L.D., N.E. Welker, G.N. Bennett and E.T. Papoutsakis. 1992. Expression of cloned homologous genes in *Clostridium acetobutylicum* ATCC 824. Biotechnology. **10**:190-195.

Mermelstein, L.D. and E.T. Papoutsakis. 1993a. *In* vivo methylation in *Escherichia coli* by the *Bacillus subtilis* phage 3T I methyltransferase to protect plasmids from restriction upon transformation of *Clostridium acetobutylicum* ATCC 824. Appl. Environ. Microbiol. **59**:1077-1081.

Mermelstein, L.D. and E.T. Papoutsakis. 1993b. Evaluation of macrolide and lincosamide antibiotics for plasmid maintenance in low pH *Clostridium acetobutylicum* ATCC 824 fermentations. FEMS Microbiol. Lett. **113**:71-76.

Mermelstein, L.D., D.J. Petersen, G.N. Bennett and E.T. Papoutsakis. 1993. Metabolic engineering of *Clostridium acetobutylicum* ATCC 824 for increased solvent formation by enhancement of acetone formation enzyme activities using synthetic acetone operon. Biotech. & Bioeng. **42**: 1053-1060.

Mes-Hartree, M. and J.N. Saddler. 1982. Butanol production of *Clostridium acetobutylicum* grown on sugars found in hemicellulose hydrolysates. Biotechnol. Lett. **4**:247-252.

Miller, P.F. and M.C Sulavik. 1996. Overlaps and parallels in the regulation of intrinsic multiple-antibiotic resistance in *Escherichia coli*. Mol. Microbiol. **21**(3): 441-448.

Minton, N.P., J.K. Brehm, J.D. Oultram, T.J. Swinfield and D.E. Thompson. 1988. Construction of plasmid vectors for gene transfer in *Clostridium acetobutylicum*. In: *Anaerobes Today*. J.M. Hardie & S.P. Borriello (eds.). pp. 125-140. Chichester:Wiley.

Minton, N.P., J.K. Brehm, J.D. Oultram, T.J. Swinfield, S. Whelan, D.E. Thompson, M. Young and W. Staudenbauer. 1990. Development of Genetic Systems for *Clostridium acetobutylicum*. In: *Proceedings of the 6TH International Symposium on Genetics of Industrial Microorganisms (GIM90)*. H. Heslot et al. (eds.). pp. 759-770. Strasbourg, France. Société Française de Microbiologie.

Minton, N.P., J.K. Brehm, T.J. Swinfield, S.M. Whelan, M.L. Mauchline, N. Bodsworth and J.D. Oultram. 1993a. Clostridial cloning vectors. In: *The Clostridia and Biotechnology*. D.R. Woods (ed.). pp. 119-150. Butterworth-Heinemann. London.

Minton, N.P., J.K. Brehm, T.J. Swinfield, S.M. Whelan, M.L. Mauchline, N. Bodsworth and J.D. Oultram. 1993b. Vectors for *Clostridium acetobutylicum*. In: *Genetics and Molecular Biology of Anaerobic Bacteria*. M. Sebald (ed.). pp. 120-140. Springer-Verlag. New York.

Mitchell, W.J., J.E. Shaw and L. Andrews. 1991. Properties of the glucose phosphotransferase system of *Clostridium beijerinckii* NCIB 8052. Appl. Environ. Microbiol. 57: 2534-2539.

Mitchell, W.J., K.A. Albasheri and M. Yazadanian. 1995. Factors affecting utilization of carbohydrates by clostridia. FEMS Microbiol. Rev. **17**: 317-329.

Mitchell, W.J. 1998. Physiology of carbohydrate to solvent conversion by clostridia. Adv. Microbiol. Physiol. **39**: 31-130.

Miwa, Y., M. Saikawa and Y. Fujita. 1994. Possible function and some properties of the CcpA protein of *Bacillus subtilis*. Microbiology. **140** (10): 2567-75.

Miwa, Y., K. Nagura, S. Eguchi, H. Fukuda, J. Deutscher and Y. Fujita. 1997. Catabolite repression of the *Bacillus subtilis gnt* operon exerted by two catabolite-responsive elements. Mol. Microbiol. **23**(6): 1203-1213.

Miwa, Y., A. Nakata, A. Ogiwara, M. Yamamoto and Y. Fujita. 2000. Evaluation and characterization of catabolite-responsive elements (cre) of *Bacillus subtilis*. Nucleic Acids Res. **28**(5): 1206-1210.

Monedero, V., M.J. Gosalbes and G. P-Martínez. 1997. Catabolite repression in *Lactobacillus casei* ATCC 393 is mediated by CcpA. J. Bacteriol. **179**(21):6657-6664.

Moon, M.W., S.Y. Park, H.K. Kim, D.Y. Yum, H.S. Kwak, T.K. Oh and J.K. Lee. 2000. Cloning and expression of the ccpA gene encoding catabolite control protein from Theremoactinomyces sp. E79. Biosci. Biotechnol. Biochem. **64**(10): 2254-2258.

Moreira, A.R. 1983. Acetone-butanol fermentation. In: *Organic chemicals from biomass*. D.L. Wise (ed.). **pp.** 385-406, Benjamin/Cummings Publishing Co. Inc., Menlo Park, Calif.

Moreno, M.S., B.L. Schneider, R.R. Maile, W. Weyler and M.H. Saier. 2001. Catabolite repression mediated by CcpA protein in *Bacillus subtils*: novel modes of regulation revealed by whole-genome analyses. Mol. Microbiol. **39**(5): 1366-1381.

Morris J.G. 1993. History and future potential of the Clostridia in biotechnology. In: *The Clostridia and biotechnology*. D.R. woods (ed.). **pp.** 1-23. Butterworth-Heinemann. London.

Muscariello, L., R. Marasco, M. de Felice and M. Sacco. 2001. The functional *ccpA* is required for carbon catabolite repression in *Lactobacillus plantarum*. Appl. Environ. Microbiol. **67**(7): 2903-2907.

Nakhmanovich, B.M. and N.A. Shcheblykina. 1959. Fermentation of pentoses of corn cob hydrolyzates by *Clostridium acetobutylicum*. Mikrobiologiya. **29**: 67-72.

Nakas, J.P., M. Schaedle, C.M. Parkinson, C.E. Coonley and S.W. Tanenbaum. 1983. System development for linked fermentation production of solvents from algal biomass. Appl. Environ. Microbiol. **46**:1017-1023.

Nakano, M.N., Y.P. Daily, P. Zuber and D.P. Clark. 1997. Characterization of anaerobic fermentative growth of *Bacillus subtilis*: identification of fermentation end products and genes required for growth. J. Bacteriol. **179**: 6749-6755.

Noirot, P., M.A. Petit and S.D. Ehrlich. 1987. Plasmid replication stimulates DNA recombination in *Bacillus subtilis*. J. Mol. Biol. **196**:39-48.

Notley-McRobb, L., A. Death and T. Ferenci. 1997. The relationship between external glucose concentration and cAMP levels inside *Escherichia coli*: implications for models of phosphotransferase-mediated regulation of adenylate cyclase. Microbiology. **143**: 1909-1918.

Nguyen, C.C. and M.H. Saier. 1995. Phylogenetic, structural and functional analyses of the LacI-galR family of bacterial transcription factors. FEBS Lett. **377**:98-102.

O'Brien, R.W. and J.G. Morris. 1971. Oxygen and the growth and metabolism of *Clostridium acetobutylicum*. J. Gen. Microbiol. **68**:307-318.

Oscarsson, J., Y. Mizunoe, B. E. Uhlin, and D. J. Haydon. 1996. Induction of haemolytic activity in *Escherichia coli* by the *slyA* gene product. Mol. Microbiol. **20**:191-199.

Oultram, J.D., A. Davies and M. Young. 1987. Conjugal transfer of a small plasmid from *Bacillus subtilis* to *Clostridium acetobutylicum* by cointegrate formation with plasmid pAM_1. FEMS Microbiol. Lett. **42**:113-119.

Oultram, J.D., M. Loughlin, T.J. Swinfield , J.K. Brehm, D.E. Thompson and N.P. Minton. 1988a. Introduction of plasmids into whole cells of *Clostridium acetobutylicum* by electroporation. FEMS Microbiol. Lett. **56**: 83-88.

Oultram, J.D., H. Peck, J.K. Brehm, D.E. Thompson, T.J. Swinfield and N.P. Minton. 1988b. Introduction of genes for leucine biosynthesis from *Clostridium pasterianum* into *C.acetobutylicum* by cointegrate conjugal transfer. Mol. Gen. Genet. **214**:177-179.

Page, R.D.M. 1996. Treeview: An application to display phylogenetic tree on personal computers. Computer application in the Biosciences. **12**: 357-358.

Papoutsakis, E.T and G.N. Bennett. 1993. Cloning, structure and expression of acid and solvent pathway genes. In: *The Clostridia and Biotechnology*. Woods, D.R. (ed.). pp. 157-195. Butterworth-Heinemann. London.

Poole, K., K. Tetro, Q. Zhao, S. Neshat, D. E. Heinrichs, and N. Bianco. 1996. Expression of the multidrug resistance operon *mexA-mexB-oprM* in *Pseudomonas aeruginosa*: *mexR* encodes a regulator of operon expression. Antimicrob. Agents Chemother. **40**:2021-2028.

Postma, P.W., J.W. Lengeler and G.R. Jacobson. 1993. Phosphoenolpyruvate: carbohydrate phosphotransferase systems of bacteria. Microbiol. Rev. **57**: 543-594.

Powell, B.S., D.L. Court, T. Inada, Y. Nakamura, V. Michotey, X. Cui, A. Reizer, M.H. Saier and J. Reizer. 1995. Novel proteins of the phosphotransferase system encoded within the *rpoN* operon of *Escherichia coli*. Enzyme IIA^{Ntr} affects growth on organic nitrogen and the conditional lethality of an era^{ts} mutant. J.Biol. Chem. **270**: 4822-4839.

Reizer, J., U Bergstedt, A. Galinier, E. Kuster, M.H. Saier, W. Hillen, M. Steinmetz and J. Deutscher. 1996. Catabolite repression resistance of *gnt* operon expression in *Bacillus subtilis* conferred by mutation of His-15, the site of phosphoenolpyruvate-dependent phosphorylation of the phosphocarrier protein HPr. J. Bacteriol. 178(18): 5480-5486.

Pujic P, R. Dervyn, A. Sorokin and S.D. Ehrlich. 1998. The *kdgRKAT* operon of *Bacillus subtilis*: detection of the transcript and regulation by the *kdgR* and *ccpA* genes.
Microbiology. **144**(11): 3111-3118.

Quixeley, K.W.M. 1999. Studies on the regulation of nitrogen metabolism in *Clostridium acetobutylicum* and *Clostridium beijerinckii* NCIMB 8052. PhD Thesis. University of Cape Town, South Africa.

Ramseier, T.M., J. Reizer, E. Kuster, W. Hillen and M.H. Saier. 1995. In vitro binding of the CcpA protein of *Bacillus megaterium* to *cis*-acting catabolite responsive elements (CREs) of gram-positive bacteria. FEMS Microbiol. Lett. **129**:207-213.

Reed, K.C. and D.A. Mann. 1985. Rapid transfer of DNA from agarose gels to nylon membranes. Nucleic Acid Res. **13**: 7207-7221.

Reid, S.J., E.R. Allcock, D.T. Jones and D.R. Woods. 1983. Transformation of *Clostridium acetobutylicum* protoplasts with bacteriophage DNA. Appl. Environ. Microbiol. **45**: 305-307.

Reid, S.J, M.S. Rafudeen and N.G. Leat. 1999. The genes controlling sucrose utilization in *Clostridium beijerinckii* NCIMB 8052 constitute an operon. Microbiology. **145**:1461-1472.

Reizer, J., U. Bergstedt, A. Galinier, E. Kuster, M.H. Saier, W. Hillen, M. Steinmetz and J. Deutscher. 1996. Catabolite repression resistance of gnt operon expression in *Bacillus subtilis* conferred by mutation of His-15, the site of phosphoenolpyruvate-dependent phosphorylation of the phosphocarrier protein HPr. J. Bacteriol. **178**(18): 5480-5486.

Reizer, J., C. Hoischen, F. Titgemeyer, C. Rivolta, R. Rabus, J. Stülke, D. Karamata, M.H. Saier and W. Hillen. 1998. A novel protein kinase that controls carbon catabolite repression in bacteria. Mol. Microbiol. **27**:1157-1169

Richards, D.F., P.E. Linnett, J.D. Oultram and M. Young. 1988. Restriction endonucleases in *Clostridiun pasteurianum* ATCC 6013 and *Clostridium thermohydrosulfuricum* DSM 568. J. Gen. Microbiol. **134**: 3151-3157.

Reynolds, H. and C.H. Werkman. 1934. Fermentation of artichokes. Proc. Iowa Acad. Sci. **41**:75-78.

Robinson, G.C. 1922. A study of acetone and butyl alcohol fermentation of various carbohydrates. J. Biol. Chem. **52**:125-155.

Ross, D. 1961. The acetone-butanol fermentation. Proc. Ind. Microbiol. **3**:73-85.

Rutberg, B. 1997. Antitermination of transcription of catabolic operons. Mol. Microbiol. **23**(3):413-421.

Saddler, J.N., E.K.C. Yu, M. Mes-Hartree, N. Levitin and H.H. Brownell. 1983. Utilization of enzymatically hydrolyzed wood hemicelluloses by microorganisms for production of liquid fuels. Appl. Environ. Microbiol. **45**: 153-160.

Saier, M.H. and J. Reizer. 1992. Proposed uniform nomenclature for the proteins and protein domains of the bacterial phosphoenolpyruvate:sugar phosphotransferase system. J. Bacteriol. **174**: 1433-1438.

Saier M.H., S. Chauvaux, J. Deutscher, J. Reizer and J-J. Ye. 1995. Protein phosphorylation and regulation of carbon metabolism in gram-negative versus gram-positive bacteria. TIBS. **20**:267-271.

Saier, M.H. 1996. Cyclic AMP-independent catabolite repression in bacteria. FEMS Microbiol. Lett. **138**: 97-103.

Saier M.H., S. Chauvaux, G.M. Cook, J. Deutscher, I.T. Paulsen, J. Reizer and J-J. Ye. 1996. Catabolite repression and inducer control in gram-positive bacteria. Microbiology. **142**:217-230.

Sambrook, J., E.F. Fritsch and T. Maniatis. 1989. *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, Cold Spring Harbor Laboratory, NY.

Sanger, F., S. Nicklen and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA. 74:5463-5467.

Schäfer, A., J. Kalinowski, R. Simon, A-H. Seep-Feldhaus and A. Pühler. 1990. High-frequency conjugal plasmid transfer from gram-negative *Escherichia coli* to various gram-positive coryneform bacteria. J. Bacteriol. **172**(3): 1663-1666.

Schick, J.B., B. Weber, J.R. Klein and B. Henrich. 1999. PepR1, a CcpA-like transcription regulator of *Lactobacillus delbrueckii* subsp. *lactis*. Microbiology. **145**(11): 3147-3154.

Schmiedel, D. and W. Hillen. 1996. Contributions of XylR, CcpA and *cre* to diauxic growth of *Bacillus megaterium* and to xylose isomerase expression in the presence of glucose and xylose. Mol. Gen. Genet. **250**(3): 259-266.

Schmiedel, D., M. Kintrup, E. Kuster and W. Hillen. 1997. Regulation of expression, genetic organization and substrate specificity of xylose uptake in *Bacillus megaterium*.
Mol. Microbiol. **23**(5):1053-1062.

Schumacher, M.A., K.Y. Choi, H. Zalkin and R.G. Brennan. 1994. Crystal structure of LacI member, PurR, bound to DNA: Minor groove binding by _ helices. Science. **266**:763-770.

Sebald, M. 1993. Mutations. In: *Genetics and Molecular Biology of Anaerobic Bacteria.* M.Sebald (ed.). pp. 64-97. Springer-Verlag, New York.

Selby, C.A and A. Sancar. 1995. Structure and function of transcription-repair coupling factor II. Catalytic properties. J. Biol. Chem. **270**: 4890-4895.

Shaheen, R., M. Shirley, D.T. Jones. 2000. Comparative fermentation studies of industrial strains belonging to four species of solvent-producing clostridia. J. Mol. Microbiol. Biotechnol. **2**(1): 115-124.

Simpson, C.L. and R.R.B. Russell. 1998. Identification of a homolog of CcpA catabolite repressor protein in *Streptococcus mutans*. Infect. Immun. **66**(5):2085-2092.

Sizemore, C., B. Wieland, F. Götz and W.Hillen. 1992. Regulation of *Staphylococcus xylosus* xylose utilization genes at the molecular level. J.Bacteriol. **174**: 3042-3048.

Sprenger, G.A. and J.W. Lengeler. 1988. Analysis of sucrose catabolism in *Klebsiella pneumoniae* and in Scr^+ derivatives of *Escherichia coli* K-12. J. Gen. Microbiol. 134: 1635-1644.

Staudenbauer W.L. and W. Dubbert. 1993. Genetics and molecular biology of chloramphenicol acetyltransferase of *Clostridium butyricum*. In: *Genetics and Molecular Biology of Anaerobic Bacteria.* M. Sebald (ed.). pp. 174-178. Springer-Verlag. New York

Steinmetz, M. 1993. Carbohydrate catabolism: pathways, enzymes, genetic regulation and evolution. In: *Bacillus subtlis and Other gram-positive bacteria.* A.L. Sonenshein (ed.). pp. 157-170. Washigton DC: American Society for Microbiology.

Strauch, M.A. 1995. AbrB modulates expression and catabolite repression of a *Bacillus subtilis* ribose transport operon. J. Bacteriol. **177**(23): 6727-6731.

Stülke J., M. Arnaud, G. Rapoport and I. M-Verstraete. 1998. PRD- a protein domain involved in PTS-dependent induction and carbon catabolite repression of catabolic operons in bacteria. Mol. Microbiol. **28**(5):865-874.

Stülke J. and W. Hillen. 1999. Carbon catabolite repression in bacteria. Current Opinion in Microbiol. 2:195-201.

Stülke J. and W. Hillen. 2000. Regulation of carbon catabolism in Bacillus species. Ann. Rev. Microbiol. **54**(1):849-880.

Swinfield, T.J., J.D. Oultram, D.E. Thompson, J.K. Brehm and N.P. Minton. 1990. Physical characterization of the replication region of the *Streptococcus faecalis* plasmid pAM_1. Gene. **87**:79-90.

carbon catabolite repression by a non-phosphoenolpyruvate-dependent phosphotransferase system sugar. J. Bacteriol. **182**(21): 5982-5989.

Vellanoweth, R.L. and J.C. Rabinowitz. 1992. The influence of ribosome-binding-site elements on translational efficiency in *Bacillus subtilis* and *Escherichia coli in vivo*. Mol. Micobiol. 6:1105-1114.

Verhasselt P. and J. Vanderleyden. 1993. Cloning and expression of *Clostridium acetobutylicum* genes involved in carbohydrate utilization. In: *Genetics and molecular biology of anaerobic bacteria*. M. Sebald (ed.). pp. 301-316. Springer-Verlag, New York.

Voet, D. and J.G. Voet. 1995. Chapter 30: Transfer RNA and its aminoacylation. In : *Biochemistry*. D. Voet and J.G. Voet (ed.). pp. 967-978. John Wiley & Sons. New York.

Voget, C.E., C.F. Mignone and R.J. Ertola. 1985. Butanol production from apple pomace. Biotechnol. Lett. 7:43-46.

Wagner, E., F. Götz and R. Brückner. 1993. Cloning and characterization of the *scrA* gene encoding the sucrose-specific enzyme II of the phosphotransferase system from *Staphylococcus xylosus*. Mol. Gen. Genet. **241**:33-41.

Wayman, M. and S. Yu. 1985. Acetone-butanol fermentation of xylose and sugar mixtures. Biotechnol. Lett. 7:255-260.

Weickert, M.J. and S. Adhya. 1992. A family of bacterial regulators homologous to Gal and Lac repressors. J. Biol. Chem. 267(22):15869-15874.

Weickert, M.J. and G.H. Chambliss. 1990. Site-directed mutagenesis of a catabolite repression operator sequence in *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA. 87: 6238-6242.

Welsh F.W. and I.A. Veliky. 1984. Production of acetone-butanol from acid whey. Biotechnol. Lett. 6:61-64.

Weyer, E.R. and L.F. Rettger. 1927. A comparative study of six different strains of the organism commonly concerned in large-scale production of butyl alcohol and acetone by the biological process. J. Bacteriol. 14:399-424.

Wiley, A.S. M.J. Johnson, E. McCoy and W.H. Peterson. 1941. The fermentation of waste sulfite liquor. Ind. Eng. Chem. 35:606-610.

Wilkinson, S.R. and M. Young. 1993. Wide diversity of genome size among different strains of *Clostridium acetobutylicum*. J. Gen. Microbiol. **139**: 1069-1076.

Wilkinson, S.R. and M. Young. 1994. Targeted integration of genes into the *Clostridium acetobutylicum* chromosome. Microbiology. **140**: 89-95.

Wilkinson, S.R. and M. Young. 1995. Physical map of the *Clostridium beijerinckii* (formely *Clostridium acetobutylicum*) NCIMB 8052 chromosome. J. Bacteriol. **177**: 439-448.

Wilkinson, S.R., M. Young, R. Goodacre, J.G. Morris, J.A.E. Farrow and M.D. Collins. 1995a. Phenotypic and genotypic differences between certain strains of *Clostridium acetobutylicum*. FEMS Microbiol. Lett. **125**: 199-204.

Wilkinson, S.R., D.I. Young, J.G. Morris and M. Young. 1995b. Molecular genetics and the initiation of solventogenesis in *Clostridium beijerinckii* (formerly *Clostridium acetobutylicum*) NCIMB 8052. FEMS Microbiol. Lett. **17**:275-285.

Willets, N. and B. Wilkins. 1984. Processing of plasmid DNA during bacterial conjugation. Microbial. Rev. **48**:24-41.

Williams, D.R., D.I. Young, J.D. Oultram, N.P Minton and M. Young. 1990a. Development and optimization of conjugative plasmid transfer from *Escherichia coli* to *Clostridium acetobutylicum* NCIB 8052. In: *Clinical and molecular aspects of anaerobes*. S.P. Borriello (ed.). **pp.** 239-246. Wrightson Biomedical Publishing Ltd.

Williams, D.R., D.I. Young and M. Young. 1990b. Conjugative plasmid transfer from *Escherichia coli* to *Clostridium acetobutylicum*. J. Gen. Micro. **136**: 819-826.

Wilson, K. 1987. Preparation of genomic DNA from Bacteria. Current Protocols, Unit 2.4, **pp** 2.4.1-2.4.5.

Wray, L.V., F.K. Pettengill and S.H. Fisher. 1994. Catabolite repression of the *Bacillus subtilis hut* operon requires a *cis*-acting site located downstream of the transcription initiation site. J. Bacteriol. **176**(7): 1894-1902.

Woods, D.R. 1995. The genetic engineering of microbial solvent production. Trends Biotechnol. **13**:259-264

Wouters, J.A., H.H. Kamphuis, J. Hugenholtz, O.P. Kuipers, W.M. de Vos and T. Abee. 2000. Changes in glycolytic activity of *Lactococcus lactis* induced by low temperature. Appl Environ Microbiol. 66(9): 3686-3691.

Wu, S., L. de Herminia and A. Tomasz. 1998.Genetic Organization of the *mecA* Region in Methicillin-Susceptible and Methicillin-Resistant Strains of *Staphylococcus sciuri*. J. Bacteriol. **180**(2): 236-242.

Yanish-Perron, C., J. Viera and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 cloning vectors. Gene. **33**: 103-199.

Young, M., N.P. Minton and W. Staudenbauer. 1989. Recent advances in the genetics of the clostridia. FEMS Microbiol. Rev. **63**:301-326.

Young, M. 1993a. Conjugative gene transfer in Clostridia. In: *Genetics and Molecular Biology of Anaerobic Bacteria*. M. Sebald (ed.). **pp.** 99-110. Springer-Verlag. New York.

Young, M. 1993b. Development and exploitation of conjugative gene transfer in clostridia. In: *The Clostridia and Biotechnology*. D.R. Woods (ed.). **pp.** 99-117. Butterworth-Heinemann. London.

Yu, E.K.C., L. Deschatelets and J.N. Saddler. 1984a. The bioconversion of wood hydrolyzates to butanol and butanediol. Biotechnol. Lett. **6**:327-332.

Yu, E.K.C., N. Levitin and J.N. Saddler. 1984b. Utilization of wood hemicellulose hydrolyzates by microorganisms for the production of liquid fuels and chemicals. Dev. Ind. Microbiol. **25**: 613-620.

Yu, P.L. and L.E. Pearce. 1986. Conjugal transfer of streptococcal antibiotic resistance plasmids into *Clostridium acetobutylicum*. Biotechnol. Lett. **8**: 469-474.

Zalieckas, J.M., L.V. Wray and S.H. Fisher. 1998a. Transcription-repair coupling factor is involved in carbon catabolite repression of the *Bacillus subtilis hut* and *gnt* operons. Mol. Microbial. **27**:1031-1038.

Zalieckas, J.M., L.V. Wray and S.H. Fisher. 1998b. Expression of the *Bacillus subtilis acsA* gene: position and sequence context affect *cre*-mediated carbon catabolite repression. J Bacteriol. **180**(24): 6649-6654.

Zappe, H., D.T. Jones, W.A Jones and D.R. Woods. 1988. Structure of an endo-_-1,4-glucanase gene from *Clostridium acetobutylicum* P262 showing homology with endoglucanase genes from *Bacillus* spp. Appl. Environ. Microbiol. 54: 1289-1292.

Zhou, C., Y. Yang and A.Y. Jong. 1990. Mini-prep in ten minutes. Biotechniques. 8(2): 172-173.

Zuker, M. and P. Stiegler. 1981. Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information. Nucleic Acids Res. 9:133-148