

Analysis of a Catabolic Operon for Sucrose Transport and Metabolism in *Clostridium acetobutylicum* ATCC 824

Martin Tangney* and Wilfrid J. Mitchell

Department of Biological Sciences, Heriot-Watt University,
Riccarton, Edinburgh EH14 4AS, UK

Abstract

The utilization of sucrose by *Clostridium acetobutylicum* ATCC 824 was investigated. Sucrose was found to be transported via a phosphoenolpyruvate (PEP)-dependent phosphotransferase system (PTS) and a metabolic pathway identical to that previously identified in *C. beijerinckii*, was established. The genes encoding the proteins of this pathway were identified from the *C. acetobutylicum* genome sequence, in the order *scrAKB* encoding Enzyme II of the sucrose PTS, fructokinase and sucrose 6-phosphate hydrolase respectively. While the pathway for sucrose metabolism is conserved between *C. acetobutylicum* and *C. beijerinckii*, the operons show considerable differences in organization and regulatory elements. The *C. acetobutylicum* *scr* operon contains the elements of an antiterminator-mediated regulation mechanism, typical of the BglG family of regulators. The *scrT* gene, located upstream of *scrA* encodes an antiterminator that is preceded by a transcription terminator, which is overlapped by a classical ribonucleic antiterminator (RAT) sequence. We also propose the existence of a new variant RAT-like sequence which overlaps a terminator between *scrT* and the downstream structural genes.

Introduction

Solventogenic clostridia were traditionally used in the biotechnological production of the organic solvents, acetone, butanol and ethanol, in the ABE fermentation process. Although no longer operated commercially to any significant extent, the attraction of using a renewable resource, coupled with recent advances in both genetic and fermentation technologies have stimulated fresh interest in the ABE fermentation. We have studied carbohydrate transport mechanisms in solventogenic clostridia and in particular in *Clostridium beijerinckii* NCIMB 8052. In a previous publication we described a pathway by which sucrose is taken up and metabolized by this organism and further demonstrated that this process is subject to repression by glucose (Tangney *et al* 1998b). Industrial fermentations commonly used molasses as a growth substrate, the main carbon source in which is sucrose. Regulation of sucrose utilization therefore has

industrial significance, as it represents a potentially important control point in the fermentation.

In *C. beijerinckii*, sucrose is transported into the cell via a sucrose-specific phosphoenolpyruvate (PEP)-dependent phosphotransferase system (PTS). The product, sucrose 6-phosphate, is cleaved by a hydrolase to glucose 6-phosphate and fructose, which is subsequently phosphorylated by fructokinase (Tangney *et al* 1998b). The PTS and enzymes required for metabolism of sucrose have been shown to be encoded by the genes of the *scr* operon (Reid *et al* 1999). The bacterial PTS is composed of two general cytosolic proteins, called Enzyme I and HPr, as well as a substrate-specific enzyme complex called Enzyme II. The EII complex typically contains three distinct functional domains; IIA, IIB and IIC, the architecture of which can vary (Saier Jr. and Reizer, 1992).

In addition to its role in sugar uptake and phosphorylation, the PTS in bacteria is intimately involved in regulation of gene expression. A diverse group of transcriptional regulators has been identified which are characterized by the possession of PTS regulation domains (PRDs), and are subject to regulation by the PTS via phosphorylation (Stülke *et al* 1998). An emerging family of such regulators is exemplified by the BglG protein of *Escherichia coli* (Schnetzer *et al* 1987). The BglG family of proteins are antiterminators with the capacity to interact with mRNA to prevent termination of transcription. Genes subject to regulation are preceded by a p-independent transcription terminator which under non-expressing conditions prevents transcription of the downstream gene(s). Partially overlapping the terminator is a sequence called a ribonucleic antiterminator (RAT), which is highly conserved amongst genes subject to such regulation. RAT sequences have been shown to form a secondary structure which is stabilized by an interaction with the associated antiterminator. Formation of a stable RAT/antiterminator complex precludes formation of the terminator structure and thereby allows transcription of the target operon to proceed (Rutberg, 1997). The antiterminator activity of BglG is antagonistically modulated through phosphorylation by PTS components; either positively by PEP/Enzyme I/HPr, or negatively by the EII product of the operon, BglF (Gorke and Rak, 1999). Multiple PTS-dependent phosphorylation of antiterminators has been proposed as a general mechanism by which it is ensured that operon expression occurs only in the presence of the substrate and in the absence of a more favorable carbon source (Stülke *et al* 1998). Several operons in *B. subtilis*, including the *sacPA* and *sacXY* operons (each of which encodes an EII for sucrose), are subject to this type of control (Rutberg, 1997). The *abg* operon encoding the EII for β -glucoside uptake in the rumen organism *Clostridium longisporum* exhibits similar features (Brown and Thomson, 1998), but as yet no examples have been found in solventogenic clostridia.

The availability of the genome sequence of *C. acetobutylicum* ATCC 824 now allows for a unique insight into the genetics and physiology of this group of bacteria.

Received October 19, 1999; revised November 8, 1999; accepted November 10, 1999. *For correspondence. Email m.tangney@hw.ac.uk; Tel. +44 131 451 3640; Fax. +44 131 451 3009.

The aim of this work was to investigate the mechanism of sucrose transport and metabolism in *C. acetobutylicum* ATCC 824 and to exploit the genome sequence data to conduct a comparative analysis of the sucrose utilization systems in *C. acetobutylicum* and *C. beijerinckii*. We find that although the pathway for sucrose metabolism appears to be identical in the two organisms, the *scr* operons show considerable differences in organization and regulatory elements.

Results

Sucrose Metabolism by *C. acetobutylicum*

Cultures of *C. acetobutylicum* were found to be capable of growth on minimal medium supplemented with sucrose as a sole fermentable carbon source (data not shown). In the related organism *C. beijerinckii* we have previously demonstrated that sucrose metabolism is dependent on the presence of three activities; a PEP-dependent phosphotransferase system, a hydrolase and a fructokinase (Tangney *et al* 1998b). We therefore examined *C. acetobutylicum* for evidence of these activities. Cultures were grown on sucrose as the sole carbon source, from which cell-free extracts were prepared, as described in the Experimental Procedures section. Assays performed with cell-free extracts established the presence of sucrose activity, while the demonstration of ATP-dependent phosphorylation of fructose in such extracts confirmed the presence of fructokinase activity (data not shown). We also assayed for PEP-dependent phosphorylation of sucrose, and, as is shown in Figure 1A, the existence of a sucrose PTS in this organism was established. The pathway for sucrose metabolism in *C. acetobutylicum* therefore appears to be the same as that in *C. beijerinckii*.

Sucrose PTS activity was not found in extracts prepared from cultures of *C. acetobutylicum* grown on glucose as the sole carbon source (Figure 1A). We therefore examined the effect of glucose on sucrose utilization by *C. acetobutylicum* by monitoring cultures growing in the presence of both sugars. As shown in Figure 1B, when a culture was grown initially on sucrose as the sole carbon source (to ensure that the components necessary for sucrose metabolism are present) and subsequently inoculated into medium containing both glucose and sucrose, glucose is utilized preferentially and only when it has been consumed is sucrose taken up by the cells appreciably. Glucose can therefore regulate sucrose metabolism in *C. acetobutylicum*.

Sequence Analysis of the *C. acetobutylicum scr* Structural Genes

Analysis of the *C. acetobutylicum* genome revealed the presence of three clustered open reading frames (ORFs) which by homology are equivalent to the structural genes for sucrose uptake and metabolism in other bacteria; we have accordingly designated the genes *scrA*, *scrB* and *scrK*, encoding Enzyme II of the PTS, sucrose 6-phosphate hydrolase and fructokinase respectively. The genes are in the same orientation and in the same reading frame, with the gene order *scrAKB*. Interestingly, this differs from the gene order of the *C. beijerinckii scr* operon, where *scrB* precedes *scrK*, as depicted in Figure 2. It is also notable that in the *C. beijerinckii scr* operon, the 3'-end of *scrB* overlaps the 5'-end of *scrK* (Reid *et al* 1999). No such

overlap exists between the *C. acetobutylicum* genes. Northern blot analysis confirmed that the three genes are all expressed on a single transcript in cells grown on sucrose as the sole carbon source and therefore constitute an operon (data not shown).

The *scrA* gene is preceded, at a distance of 8 nucleotides, by a putative ribosome binding site (rbs),

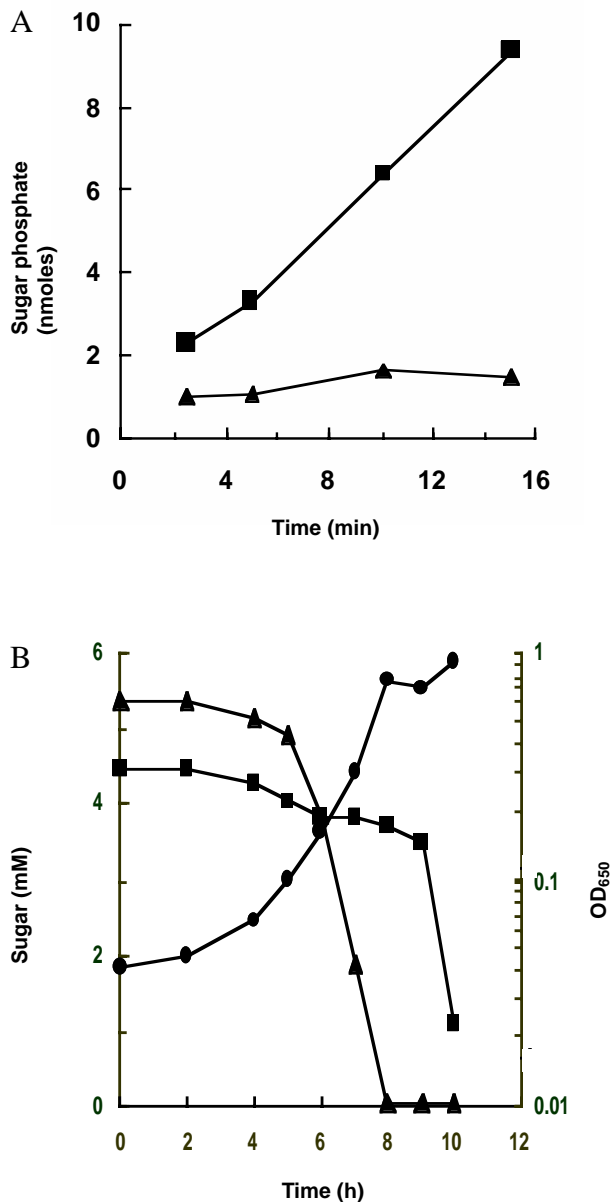


Figure 1. Sugar utilization and transport activity in *C. acetobutylicum*. A. Sucrose phosphorylation in cell free extracts of *C. acetobutylicum*. Extracts were prepared from cultures grown on either sucrose (■) or glucose (▲) as the sole carbon source. Extracts were assayed for sucrose PTS activity in phosphorylation assays. B. Culture growth and carbohydrate utilization by *C. acetobutylicum*. A culture was grown in a defined medium containing glucose (▲) and sucrose (■). The OD₆₅₀ (●) was monitored throughout growth, and the concentration of glucose and sucrose in the culture supernatant was determined at intervals.

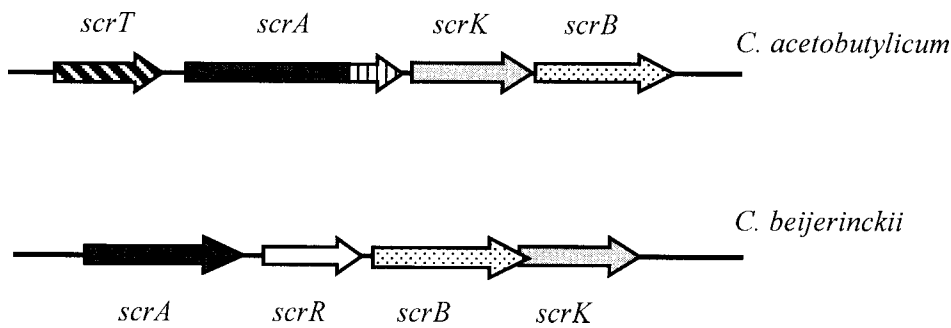


Figure 2. Diagrammatic representation of *scr* operons from *C. acetobutylicum* and *C. beijerinckii*. Homologous genes in the two operons are depicted in identical shading. The region encoding the EIIA^{scr} domain of the *C. acetobutylicum* *scrA* gene, and regulator genes *scrT* and *scrR* are each depicted in unique shading.

AGGAGG. It encodes a protein of 627 aa with a predicted molecular mass of 67,605 Da. The deduced sequence for the first 450 aa (approx.) of the protein has significant homology to IIBC^{scr} proteins, having 56%, 55% and 53% identity to the ScrA (IIBC^{scr}) proteins from *Salmonella typhimurium*, *Klebsiella pneumoniae* and *B. stearothermophilus* respectively (Ebner and Lengeler, 1988; Schmid *et al* 1991; Li and Ferenci, 1996). It also has (a relatively lower) 34% identity to the *C. beijerinckii* ScrA (IIBC^{scr}) protein. The C-terminal 180 aa (approx.) sequence has homology to IIA-glucose (IIA^{glc}) components, both discrete IIA^{glc} proteins - as in the case of the *Haemophilus influenzae* IIA^{glc} protein (40% identity) (Fleischmann *et al.*, 1995), and the IIA domain of Enzyme II^{glc} proteins - including GlcA from *B. stearothermophilus* (45% identity) and *B. subtilis* (41% identity) (Lai and Ingram, 1995; Sorokin *et al* 1996). Computer analysis of the hydrophobicity/hydrophilicity profile of the *C. acetobutylicum* ScrA protein reveals a typical N-terminal hydrophilic domain (of approximately 100 aa), corresponding to the cytoplasmically-orientated IIB domain; followed by a predominately hydrophobic domain, typical of the integral membrane IIC component (approx. 350 aa); and finally a C-terminal hydrophilic domain (approx. 180 aa), typical of the IIA component of Enzyme II (data not shown). We conclude therefore that the *C. acetobutylicum* ScrA protein is an Enzyme II of the type typically found in bacterial phosphotransferase systems (Saier Jr. and Reizer, 1992). Downstream of *scrA*, separated by 46 nucleotides, is the *scrK* gene. A putative rbs, AGGAGT, is located 8 nucleotides upstream of *scrK*. The *scrK* gene encodes a protein of 316 aa with a predicted molecular mass of 34,363 Da. It is homologous throughout its length to the *C. beijerinckii* ScrK protein (312 aa, 34,297 DA), with which it shares 50% identity. It is also homologous to other fructokinases, the most homologous (57% identity) being a putative fructokinase from *B. subtilis* (Kasahara *et al* 1997).

The *scrK* gene ends with a TAG stop codon which is immediately followed by the ATG start codon of the *scrB* gene, leaving no gap between these genes. A putative rbs, TGAAGG, lies within the coding sequence of *scrK*, and is 7 nucleotides upstream from the start of *scrB*. This gene encodes a protein of 490 aa with a predicted molecular mass of 57,580 Da. This is very similar in size to the *C. beijerinckii* ScrB protein (485 aa, 57,128 Da), although these proteins only share 38% identity. The *C. acetobutylicum* ScrB protein is most closely related to a

sucrase from *B. stearothermophilus* (Li and Ferenci, 1996) with 54% identity and is also homologous to equivalent proteins from *S. typhimurium* and *K. pneumoniae*, having 42% and 40% identity respectively (Titgemeyer *et al* 1996).

Identification of the *scrT* Gene

In *C. beijerinckii*, expression of the *scr* operon is modulated by a regulatory repressor protein ScrR (Reid *et al* 1999). The *scrR* gene lies between *scrA* and *scrB*, as depicted in Figure 2. There is no corresponding gene in the *C. acetobutylicum* *scr* gene system. However, we did identify a fourth ORF located 140 nucleotides upstream of *scrA* and in the same orientation as *scrAKB* (Figure 2). The ORF is preceded by a putative rbs, TGGAGT, 8 nucleotides upstream of an ATG initiation codon, and encodes a protein of 282 aa with a predicted molecular mass of 33,007 Da. The putative product is homologous to the BglG family of antiterminator proteins, being most closely related to the *B. subtilis* LicT (Schnetzer *et al* 1996), with 46% identity. It is also homologous to other well characterized antiterminator proteins, including BglG (Schnetzer *et al* 1987), SacT (Debarbouillé *et al* 1990) and SacY (Zukowski *et al* 1990) having 38%, 41% and 40% identity respectively to these proteins. An alignment with these antiterminators clearly identifies the *C. acetobutylicum* product as a member of this family of proteins (Figure 3A). We have assigned the name *scrT* to the gene. The ScrT protein contains all the conserved histidine residues which, in SacY and LicT of *B. subtilis* and BglG of *E. coli*, have been shown to be phosphorylated by components of the PTS (Tortosa *et al* 1997; Lindner *et al* 1999; Chen *et al* 1997).

Regulatory Features of the 824 *scr* Operon

The DNA sequence upstream of *scrT* (stretching approx. 200 nt) contains a number of discrete regions which resemble known regulatory features (Figure 4). There is a putative promoter region, which is in good agreement with the consensus clostridial promoter sequence (Young *et al* 1989). The putative -35 (TTGACA) and -10 (TATCAT) regions are separated by 17 nucleotides. Overlapping the putative -10 region is a sequence which closely resembles a catabolite responsive element (CRE). In other Gram positive bacteria the CRE sequence is the target site for catabolite repression via the catabolite control protein, CcpA (Hueck and Hillen, 1995). The potential CRE sequence at the putative -10 region has one mismatch relative to the consensus sequence (Weickert and Chambliss, 1990). Downstream of the putative promoter

A

```

LicT : MKIAKVINNNVISVVNEEQGK-ELVVMGRGLAFOKKSGDDVDEARTEKVFTL : 50
SacT : MKIYKVLNNAALIKEDDQ--EKIMVGPGLAFOKKNDLIPMNKVEKIFVV : 49
BglG : MQITKILNNVVVVIDDQQR-EKYVMGRGIGFOKRAGERINSSGTEKEYAL : 50
SacY : MKIKRILNHNAIVKQNE--EKILLCAGTAFNKKNDIVDPSKTEKTFIR : 49
ScrT : MVIKLLNNSAVTTIDDATRIEKVIMGKGLAFOKKPGDILNEEKTEKIESI : 51

LicT : DNKDVSEKEKFTLLYDTPTECMEVSEETISYAKLOLGKKLNDSTVSLTDHI : 101
SacT : RDENE--KFKQILQTLPEEHTETIAEDTISYAEGELAAPLSDHTALSDHI : 98
BglG : SSHELNGRLSELLSHTPLEVMATCDRIISLAQERLG-KLQDSTVSLTDHC : 100
SacY : KDTPDYKQFEEILETIPEDHTQISEQTISHAEKELNIKINERTHVAFSDHI : 100
ScrT : ENQENLKFQSLISETPIEHIKVSENTISYAKRKLDVKFDEHTVSLTDHL : 102

LicT : NFALQRNQGLDIKNALLWETKRLYRDEEFAIGKEALVMVKNKTVSLPPEDE : 152
SacT : SFAERIQNGLLVQNKLLHEIKALYKKEYEIGLWAIGHVKETLGVSLPEDE : 149
BglG : QFAKRFQQNVLLPNPLLDIQRLYPKFQLGEBALTIDKRIGVQLKDE : 151
SacY : SFAERLSNGMVIKNPLLEIKVLYPKEFQIGLWARALIKKGIHIPDE : 151
ScrT : SFAFRRYSKGIKIKNNMLWDIKRIYKEYNIGMWAVEYIKGELGIKMDEDE : 153

LicT : AGFIAHHIVNAELN-EEMPNIINTTKVMQEELSTVKYHEKLEFNEESLHYY : 202
SacT : AGYIAHHIHTAKMDAESMSYALKHTMIKEMIEKIKQYFNRKVDENSISYQ : 200
BglG : VGFIAMHLVSAQMS-GNMEDVAGVTQLMREMLQLIKFQFSLNYQEESLSYQ : 201
SacY : IGNIAMHIHTARNNAGDMTQTLDTTMIRDLIETEIQLSTINVEDTISYE : 202
ScrT : AGFIAHHIDASLN-ESMDNTINTEILDGILNIKYFFSIEFNEDDMSYD : 203

LicT : RFVTHLKFAQRLFNGTH-MESQDFLLDTVKEKYHRAYECTKKIQTYIER : 252
SacT : RLVTHLRYAVSRL-ESNEALHRMDEEMLYFIQKKYSFAYQCALELAEFLKN : 250
BglG : RLVTHLKFISWRILEHA-SINDSDESLQQAVKONYPQAWCAERTAIFIGL : 251
SacY : RLVTHLREFAIQHI-KAGESIYELDARMIDIEKKFKDAFLCALSIGTFVKK : 252
ScrT : RLLTHLKYFAQRVVSRKNAIDEEEKSFLEIVKTNYKEAYRCVGKIKSFIEK : 254

LicT : EYEHKLTSDELLLYTIHIRVV---KQA : 277
SacT : EYQLHLPESEAGYILHVORL----- : 271
BglG : QYQRKISPAEIMFLAINTERV----- : 272
SacY : EYGFEFPEKELCYIAMHIQR----- : 272
ScrT : NYDYEVKGGEIVYLLHVORVISSLRDK : 282

```

B

```

B.su licT : GGATTGTTACTGCTAAAGCAGGCAAACC 29
B.su sacP : GGATTGTTACTGGTAAAGCAGGCAAGACC 29
E.co bglG : GGATTGTTACTGCATTCGAGGCAAACC 29
B.su sacB : GGTTTGTTACTGATAAAGCAGGCAAGACC 29
C.ac RAT-1 : GGATTGTTACTGTTAA-GCAGGCAAACC 28

```

Figure 3. Sequence alignment of proteins and RAT sequences of the BglG family

A. Comparison of the deduced amino acid sequence of the *C. acetobutylicum* ScrT protein with regulatory proteins of the BglG family. Dashes within the sequences indicate gaps giving optimal alignment. Identical residues are within black boxes and conserved residues (present in at least three of the five sequences) are shaded. Conserved histidine residues, associated with the regulation of antiterminator proteins, are marked with an asterisk. Abbreviations: LicT, *Bacillus subtilis* LicT (Schnetz *et al* 1996); SacT, *B. subtilis* SacT (Debarbouillé *et al* 1990); BglG, *Escherichia coli* BglG (Schnetz *et al* 1987); SacY, *B. subtilis* SacY (Zukowski *et al* 1990); ScrT, *C. acetobutylicum* scrT gene product. B. Multiple alignment of the putative RAT-1 sequence from the *C. acetobutylicum* scr operon with antiterminator sequences. C.ac RAT-1 is the putative RAT sequence located upstream of the *C. acetobutylicum* scrT gene. The RAT sequences aligned with RAT-1 are the targets of the antiterminator proteins aligned with ScrT and lie upstream of the associated genes as follows; B.su licT, *B. subtilis* licT (Schnetz *et al* 1996); B.su sacP, *B. subtilis* sacP (Debarbouillé *et al* 1990); E.co bglG, *E. coli* bglG (Schnetz *et al* 1987); B.su sacB, *B. subtilis* sacB (Crutz and Steinmetz, 1992).



A

AAGCTTTTTTCAAATAAATATAATTTTCTTTTGACAAAAAATTTATGAAAATGTTATCATTAGTTTCATAGTATTGAATAC
 -35 -10

RAT-1

AG***GGATTGTTACTGTTAAGCAGGCAAAACC***CAGATTAAAGTAAATGAAGTAATCATATCGTTTATTTTAATTTGGGTTTT

ATTTATTGGTGTCTTTTAAATAATGGAGTGATATCAT **ATG GTT**.....
 M V ... (**ScrT**)

B

(scrT) **RAT-2**

AAG TAA GTAAATTTAATATCGA***GGATTGTTACTTTTTAGAAATTTATTTCTGTTAAGGCAATACCTAAATTAAGTGGTTT***
 K .

(scrA)

AGTTGAGCCATTTAGTTTAGGTATTTTTTTTTGTAATTTAAGATAAAACATAAAATT***AGGAGGG***AAAATAT **ATG GAT**..
 M D....

Figure 4. Nucleotide sequence of the DNA upstream and downstream of *scrT*

A. Nucleotide sequence of the regulatory region upstream of *scrT*. B. Nucleotide sequence of the intergenic region between *scrT* and *scrA*. The positions of the sequences shown in A and B are represented diagrammatically at the top of the figure. Potential -35 and -10 regions, marked accordingly, are underlined in bold text, and a putative CRE sequence is shaded. Transcription terminator sequences are indicated by arrows. Putative RAT sequences upstream of *scrT* and *scrA* are in bold italic text and marked as RAT-1 and RAT-2 respectively. Potential rbs sequences upstream of *scrT* and *scrA* are underlined and in italic text.

is a region of dyad symmetry, followed by several Ts, which resembles a ρ -independent transcription terminator. Preceding, and partially overlapping, this terminator region is a conserved RAT sequence, which is characteristic of operons controlled by antiterminators of the BglG family. An alignment of this *scr* sequence with other characterized RAT sequences clearly shows its strong relationship to these regulatory elements (Figure 3B). The most reasonable interpretation of the sequence data is that the *scr* sequence is a RAT which can interact with ScrT to prevent termination of transcription at the terminator in this region.

Situated in the intergenic region between *scrT* and *scrA* lies another potential transcription terminator which could prevent transcription of *scrAKB* (Figure 4). It seemed reasonable to expect that a RAT sequence would also be associated with this terminator, to facilitate antitermination via interaction with ScrT. However no typical RAT sequence could be identified. A key feature of RAT sequences is that they partially overlap the associated terminator. As shown in Figure 5 this overlapping 3'-region is particularly well conserved among RAT sequences. In fact, both the 5' and

the 3'-ends of the RAT sequences are well conserved, with most variation occurring in the middle of the sequence - a region which is thought to play a role in the individual specificity of RAT structures (Aymerich and Steinmetz, 1992).

An examination of the sequence surrounding the terminator preceding *scrA* revealed a sequence which is very similar to the 3'-end of a RAT, although the bases immediately upstream of this region do not show significant homology to RAT sequences. We also identified a second region, located 36 nt upstream of the putative terminator which has remarkable homology to the 5'-end of a RAT. If these two regions of homology do represent the beginning and end of a RAT, then in total this new putative RAT-like sequence would have a length of 42 nt, which is significantly longer than recognized RAT sequences. Nevertheless, by allowing the introduction of a large gap (for optimization), an alignment of this new RAT-like sequence (RAT-2) with other RAT sequences clearly reveals significant regions of homology (Figure 5A). Furthermore, the *arb* operon in *Erwinia chrysanthemi* (which is subject to BglG-type antiterminator regulation) also contains a putative

terminator for which no typical RAT could be identified (El Hassouni *et al* 1992). The inclusion in the alignment of a putative extended RAT-like sequence from this region (similar in size to the *scr* RAT-2), supports the hypothesis that RAT-2 is related to other RAT sequences (Figure 5A).

To further investigate the relationship between the *scr* RAT-2 sequence and established RAT sequences, we performed secondary structure analysis on the predicted RNA sequence for RAT-2. On the basis of both experimental evidence and computer analysis, a model has been proposed for the secondary structure of the *B.*

subtilis sacB RAT sequence (Aymerich and Steinmetz, 1992) which is depicted in Figure 5B. We found that the RAT sequence preceding *scrT* (RAT-1) can be folded to form a structure which is almost identical to the model RAT. We also discovered that RAT-2 has the potential to form a secondary structure which in part is equivalent to the *sacB* RAT, but which is extended by a further stem-loop fold (Figure 5B). Furthermore, the *E. chrysanthemi* sequence (shown in the alignment Figure 5A) can also be folded to form essentially the same type of structure (Figure 5B). The extra nucleotides required to extend the RAT-2

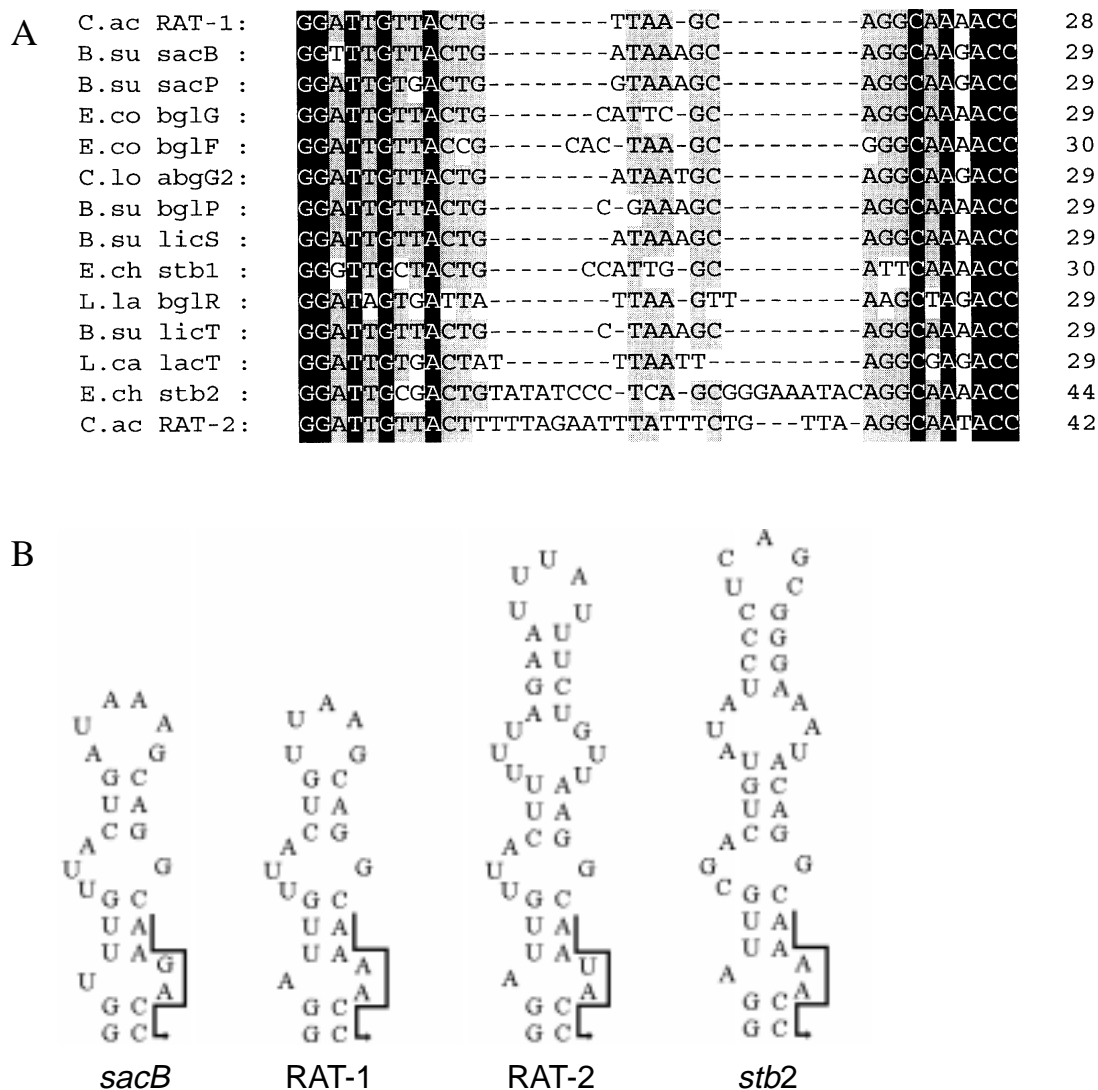


Figure 5. Comparative analysis of putative RAT sequences

A. Multiple alignment of putative RAT sequences from the *C. acetobutylicum scr* operon with other antiterminator sequences. Dashes within the sequences indicate gaps giving optimal alignment. Identical residues are within black boxes and highly conserved residues are shaded. C.ac RAT-1 and C.ac. RAT-2 are the putative RAT sequences located upstream and downstream (respectively) of the *C. acetobutylicum scrT* gene. E.ch *stb2* is the putative RAT-like sequence located downstream of the *Erwinia chrysanthemi arbG* gene, while E.ch *stb1* lies upstream of the same gene (El Hassouni *et al* 1992). C.lo *abgG2* lies downstream of the *C. longisporum abgG* gene (Brown and Thomson, 1998). The remaining RAT sequences lie upstream of the associated genes as follows; B.su *sacB*, *B. subtilis sacB* (Crutz and Steinmetz, 1992); B.su *sacP*, *B. subtilis sacP* (Debarbouillé *et al* 1990); E.co *bglG* and E.co *bglF*, *E. coli bglG* and *E. coli bglF* (Schnetz *et al* 1987); B.su *bglP*, *B. subtilis bglP* (Krüger and Hecker, 1995); B.su *licS* and B.su *licT*, *B. subtilis licS* and *B. subtilis licT* (Schnetz *et al* 1996); L.la *bglR*, *Lactococcus lactis bglR* (Bardowski *et al* 1994); L.ca *lacT*, *Lactococcus casei lacT* (Alpert and Siebers, 1997). B. Model of RAT and putative RAT-like structures. Arrows indicate the bases overlapping the associated terminator sequences.

structure are situated in the central region of the sequence, accounting for both the increased length of the RAT and the sequence conservation at both the 5' and 3' ends (Figure 5A).

Discussion

In *C. beijerinckii* we previously identified the pathway for uptake and catabolism of sucrose (Tangney *et al* 1998b). The pathway appears to be conserved in *C. acetobutylicum*, as we identified each of the required activities in cell-free extracts of cultures grown on sucrose. The sequencing of the *C. acetobutylicum* genome further allowed us to identify the genes for these enzymes and provided us with the first opportunity to compare paralogous catabolic gene systems from two different species of solventogenic clostridia. Sequence analysis revealed that while the two systems are clearly related, they are not particularly well conserved. The gene order in the two systems is different (Figure 2) and consequently the unusual overlap between *scrB* and *scrK* found in *C. beijerinckii* is not preserved in *C. acetobutylicum*. Neither is there a regulatory gene interrupting the structural genes in *C. acetobutylicum*. At the aa sequence level, the best conservation was observed between the *scrK* products, which exhibited 50% identity. The *scrB* products were less well conserved, with only 38% identity. In fact, the *C. acetobutylicum* ScrB was more closely related to equivalent genes from Gram negative bacteria than to the *C. beijerinckii* ScrB. The lowest homology was between the *scrA* products which share only 34% identity, and which also had significant structural differences.

Computer based analysis of the *C. acetobutylicum* ScrA revealed that it is a IBCA protein - *i.e.* all three functional domains of the sucrose Enzyme II are synthesized as a single protein. In contrast, the *scrA* gene in *C. beijerinckii* encodes a IBC protein. Generally such genes are linked to the gene encoding the associated IIA component. For example, there is a separate gene for the IIA component of the *C. beijerinckii* glucitol PTS (Tangney *et al* 1998a). However, there is no associated gene encoding the sucrose IIA component within the *C. beijerinckii scr* operon (Reid *et al* 1999). Furthermore, we had previously demonstrated that all of the specificity for sucrose transport via the PTS was contained within the membrane fraction of *C. beijerinckii* cells, demonstrating that there is no individual IIA^{scr} component in the cytosol in this organism (Tangney *et al* 1998b). The source of the IIA activity for sucrose in *C. beijerinckii* is therefore unknown. The finding that the IIA domain of the *C. acetobutylicum* ScrA is homologous to IIA glucose proteins (or domains) is therefore significant, as clearly this domain will function in the transport of sucrose via the PTS in *C. acetobutylicum*. One possible interpretation of these data is that in *C. beijerinckii* the IIA component of the glucose PTS (Mitchell *et al* 1991) can substitute for this activity in sucrose grown cells, as has also been demonstrated in *B. subtilis* (Sutrina *et al* 1990).

A striking and unexpected difference between the *C. acetobutylicum* and *C. beijerinckii scr* systems is the absence in the former of the repressor protein encoded by the *scrR* gene. In the absence of sucrose, transcription of the *C. beijerinckii scr* operon is proposed to be inhibited by binding of the repressor protein ScrR to an operator

upstream of the *scrA* gene; while in the presence of sucrose the operon is de-repressed and transcription proceeds (Reid *et al* 1999). Having detected a similar pathway for sucrose metabolism in both organisms and given the close relationship of *C. beijerinckii* to *C. acetobutylicum*, we had expected to find an equivalent regulatory gene in the latter. However, rather than a repressor we found evidence for the presence of the elements involved in the control of transcription by antiterminator proteins of the *E. coli* BglG family.

Sequence analysis provides compelling evidence that ScrT is an antiterminator of the BglG family. This is the first example of this type of regulator in solventogenic clostridia and further extends the range of organisms where such regulators have been identified. The *scrT* gene is located upstream of *scrAKB* and is flanked on either side by sequences which can function as ρ -independent transcription terminators. The upstream terminator sequence is preceded by, and partially overlaps with, a typical RAT sequence. This format essentially defines regulation of transcription by antiterminators of the BglG family, and strongly implies that ScrT can autoregulate its expression via an antitermination mechanism at this target site (RAT-1). Antiterminators of the BglG family are known to regulate expression of target genes which are themselves preceded by a RAT. Typically, but not exclusively, the target gene (or gene system) encodes an EII of the PTS which may in turn modulate the activity of the antiterminator protein by phosphorylation, providing the cell with a sensitive feedback loop for regulation (Rutberg, 1997; Stülke *et al* 1998). Upstream of *scrA* in *C. acetobutylicum* lies a second potential ρ -independent transcription terminator, but in this case there is no obvious associated RAT sequence. Nevertheless we did identify a novel sequence, which although longer than the currently accepted RAT sequences does have RAT-like features. The *scrA* RAT-like sequence (RAT-2) is homologous at the sequence level to other RATs at its termini but also, at least in part, in its secondary structure. Most importantly the sequence overlaps the terminator, which, should the RAT form a secondary structure, would preclude the formation of the terminator and allow transcription of *scrAKB*. While a promoter could not be clearly identified (a common difficulty due to the low GC content of clostridial DNA), Northern blot experiments revealed that these genes are expressed on a single transcript in sucrose grown cells. This implies that either the putative terminator sequence has no termination activity, or that it is somehow prevented from forming under these conditions. We favour the latter explanation and suggest that ScrT can interact with RAT-2 to prevent formation of the terminator. The *arb* gene system in *E. chrysanthemi* appears to follow a similar pattern (El Hassouni *et al* 1992). This operon is a paralog of the *E. coli bgl* operon (the original example for the BglG family of regulators) and comprises three genes, *arbG*, *arbF* and *arbB*, encoding an antiterminator, an Enzyme II and a phospho- β -glucosidase respectively. As with *scrT*, the *arbG* gene is flanked by two terminators. The upstream terminator is associated with a typical RAT sequence while the downstream terminator is overlapped by a sequence resembling the *scrA* RAT-2. These RAT-like sequences share extensive homology to each other (as well as to typical RAT sequences) at the 5' and 3' ends, although they differ from each other in the central part of the

sequence. It is therefore significant that they can fold to form similar structures, as depicted in Figure 5. It would be interesting to determine if ScrT and ArbG can recognise the opposite RAT-like target sequences. Antitermination activity at RAT sequences other than the associated target sequence has been demonstrated for several members of the BglG antiterminator family, using a gene reporter system (Le Coq *et al* 1995; Bardowski *et al* 1994; Schnetz *et al* 1996; Alpert and Siebers, 1997). This is perhaps not surprising given the high degree of conservation between typical RAT sequences. It is therefore conceivable that these new potential RAT-like structures could allow the associated antiterminator protein an extra degree of discrimination in target recognition. Alternatively, it is possible that subtle changes in the antiterminator/RAT complex formed with the putative new RAT-2 structure may serve to influence gene expression or mRNA stability, to further extend the sensitivity and versatility of regulation by this mechanism. It will be interesting to determine the true role of these sequences in the regulation of transcription.

We demonstrated that the metabolism of sucrose in *C. acetobutylicum* is regulated by glucose, as is the case for many substrates in bacteria where glucose is a preferred carbon source. The mechanism of catabolite repression in solventogenic clostridia is currently not understood. In *B. subtilis*, a pleiotropic form of catabolite repression is achieved by glucose mediated repression of transcription of catabolic operons via the PTS phosphocarrier protein HPr and the catabolite control protein, CcpA. In the presence of glucose, HPr is phosphorylated by a metabolite-activated, ATP-dependent kinase, and a complex of CcpA, phosphorylated HPr and fructose bisphosphate regulates gene expression by binding to a target sequence called a catabolite responsive element or CRE (Weickert and Chambliss, 1990; Hueck and Hillen, 1995). Significantly, there is a potential CRE sequence overlapping the putative *scrT* promoter shown in Figure 4, and we have also identified the putative elements of a *B. subtilis*-like regulatory system in the *C. acetobutylicum* genome. A gene encoding a CcpA homolog has been cloned from *C. beijerinckii* and a putative CRE has also been identified for the *scr* operon in this organism (Reid *et al* 1999). Furthermore we have demonstrated ATP-dependent phosphorylation of HPr in cell-free extracts of *C. beijerinckii* (unpublished results). Therefore despite the differences between the two *scr* operons, it is possible that repression by glucose occurs by a mechanism similar to that in *B. subtilis*. In the case of *C. acetobutylicum*, repression of the *scr* operon by glucose may also result through alterations in the phosphorylation of ScrT. Such dual regulation of transcription has been demonstrated for the *bglPH* operon in *B. subtilis*, where in addition to regulation by antitermination via LicT, transcription is also subject to independent catabolite repression at a CRE sequence via CcpA (Krüger *et al* 1996).

Over the years it has been apparent that data were often at variance when obtained from strains of *C. acetobutylicum* which were presumed to be identical, including the type strains ATCC 824, DSM 792 and NCIMB 8052 which were lodged in different culture collections. Recent taxonomic and phylogenetic studies of a large number of solventogenic strains showed that there are four major taxonomic groups, and while strains ATCC 824 and

DSM 792 still represent *C. acetobutylicum* the NCIMB 8052 strain has been reclassified as *C. beijerinckii* (Dürre, 1998). Our analysis of equivalent gene systems from *C. acetobutylicum* and *C. beijerinckii* further highlights the significant differences between these species, where, despite the phenotypic similarities the regulation of sucrose metabolism is remarkably different in the two organisms. The fact that the *C. acetobutylicum* genome has been sequenced will facilitate further experimental analysis, and should help to significantly improve our general understanding of this industrially important organism. However, from the evidence presented in this work, caution would be advised in the extrapolation to other solventogenic clostridia of any conclusions from analysis of the *C. acetobutylicum* genome, without further experimental verification.

Experimental Procedures

Organism and Growth Conditions

C. acetobutylicum ATCC 824 was maintained as a spore suspension at 4°C. Spores were heat shocked at 80°C for 10 min. and inoculated into 20 ml of Reinforced Clostridial Medium (RCM, Oxoid). Starter cultures, grown overnight at 37°C in an anaerobic cabinet (Forma Scientific, Marietta, Ohio) under an N₂-H₂-CO₂ (80:10:10) atmosphere, were subcultured into Clostridial Basal Medium (O' Brien and Morris, 1971), supplemented with the appropriate carbon source, and incubated anaerobically at 37°C to provide working cultures. The defined medium contained the following (per litre): carbon source, 10 g; casein hydrolysate, 4 g; KH₂PO₄, 0.5 g; K₂HPO₄, 0.5 g; MgSO₄·7H₂O, 0.2 g; MnSO₄·4H₂O, 10 mg; NaCl, 10 mg; FeSO₄·7H₂O, 10 mg; *p*-aminobenzoic acid, 1 mg; thiamine HCL, 1 mg; d-biotin, 2 µg. Phosphates were sterilised separately and added to the medium after cooling.

Assay of Sugar Concentration in Culture Supernatants

Culture samples (1 ml) were removed and centrifuged at stated intervals. Glucose concentration in supernatants was determined using a Sigma assay kit No. 510-A. Sucrose was hydrolysed by diluting 0.2 ml of supernatant to 1 ml with 50 mM potassium phosphate buffer, pH 6.0, and incubating with 50 units of invertase at 37°C for 1 hour. Glucose was then assayed and the amount of sucrose estimated by difference.

Enzyme Assays in Cell-Free Extracts

Cell-free extracts were prepared essentially by the method of Mitchell and Booth (1984) as described previously (Tangney *et al* 1992). Enzyme assays were carried out as described previously (Tangney *et al* 1998b).

Northern Blot Experiments

RNA was isolated from cells using the RNeasy Total RNA kit from Qiagen. The protocol was slightly modified for use with *C. acetobutylicum*, in that the concentration of lysozyme in the cell lysis solution was increased from 1 mg/ml to 10 mg/ml. DNA probes (0.6 kb internal fragments of each ORF) were prepared by PCR and labelled with digoxigenin (DIG)-dUTP, which was obtained from Boehringer Mannheim. Northern gels were run according to the method of Goda and Minton (Goda and Minton, 1995). Membranes were hybridised with DIG-labelled DNA probes at 50°C overnight. They were washed with low stringency (2xSSC; 0.1% (w/v) SDS at room temperature) and high stringency (0.5xSSC; 0.1% (w/v) SDS at 62°C) washes and bound RNA:DNA hybrids were detected by chemiluminescence using CSPD[®] as described by the manufacturers, Boehringer Manneheim.

DNA/RNA Computational Analysis Methods

Data bank searches were performed using the BLAST service at the National Center for Biotechnology Information (Altschul *et al* 1997). The *C. acetobutylicum* ATCC 824 *scr* operon sequence was identified from the genome sequence deposited by the Genome Therapeutics Corporation, USA. RNA folding was predicted using the Mfold server, based on the version 3.0 package by Zuker and Turner (Mathews *et al* 1999).

Nucleotide Sequence Accession Number

The DNA sequence data described in this paper have been deposited in GenBank with the accession number AF205034.

Acknowledgements

We are grateful to Margaret-Anne Allred for expert technical assistance. Work in the authors' laboratory is supported by BBSRC Research Grant PO9675.

References

- Alpert, C.A., and Siebers, U. 1997. The *lac* operon of *Lactobacillus casei* contains *lacT*, a gene coding for a protein of the BglG family of transcriptional antiterminators. *J. Bacteriol.* 179: 1555-1562.
- Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25: 3389-3402.
- Aymerich, S., and Steinmetz, M. 1992. Specificity determinants and structural features in the RNA target of the bacterial antiterminator proteins of the BglG/SacY family. *Proc. Natl. Acad. Sci. USA.* 89: 10410-10414.
- Bardowski, J., Ehrlich, S.D., and Chopin, A. 1994. BglR protein, which belongs to the BglG family of transcriptional antiterminators, is involved in beta-glucoside utilization in *Lactococcus lactis*. *J. Bacteriol.* 176: 5681-5685.
- Brown, G.D., and Thomson, J.A. 1998. Isolation and characterisation of an aryl-beta-D-glucoside uptake and utilisation system (*abg*) from the gram-positive ruminal *Clostridium* species *C. longisporum*. *Mol. Gen. Genet.* 257: 213-218.
- Chen, Q., Arents, J.C., Bader, R., Postma, P.W., and Amster-Choder, O. 1997. BglF, the sensor of the *Escherichia coli* *bgl* system, uses the same site to phosphorylate both a sugar and a regulatory protein. *EMBO J.* 16: 4617-4627.
- Crutz, A.M., and Steinmetz, M. 1992. Transcription of the *Bacillus subtilis* *sacX* and *sacY* genes, encoding regulators of sucrose metabolism, is both inducible by sucrose and controlled by the DegS-DegU signaling system. *J. Bacteriol.* 174: 6087-6095.
- Debarbouillé, M., Arnaud, M., Fouet, A., Klier, A., and Rapoport, G. 1990. The *sacT* gene regulating the *sacPA* operon in *Bacillus subtilis* shares strong homology with transcriptional antiterminators. *J. Bacteriol.* 172: 3966-3973.
- Dürre. 1998. New insights and novel developments in clostridial acetone/butanol/isopropanol fermentation. *Appl. Microbiol. Biotechnol.* 49: 639-648.
- Ebner, R., and Lengeler, J.W. 1988. DNA sequence of the gene *scrA* encoding the sucrose transport protein Enzyme II (Scr) of the phosphotransferase system from enteric bacteria: homology of the Enzyme II (Scr) and Enzyme II (Bgl) proteins. *Mol. Microbiol.* 2: 9-17.
- El Hassouni, M., Henrissat, B., Chippaux, M., and Barras, F. 1992. Nucleotide sequences of the *arb* genes, which control β -glucoside utilization in *Erwinia chrysanthemi* - comparison with the *Escherichia coli* *bgl* operon and evidence for a new β -glycohydrolase family including enzymes from eubacteria, archaeobacteria, and humans. *J. Bacteriol.* 174: 765-777.
- Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A., Merrick, J.M., Mckenney, K., Sutton, G., Fitzhugh, W., Fields, C.A., Gocayne, J.D., Scott, J.D., Shirley, R., Liu, L.-I., Glodek, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T., Hedblom, E., Cotton, M.D., Utterback, T.R., Hanna, M.C., Nguyen, D.T., Saudek, D.M., Brandon, R.C., Fine, L.D., Fritchman, J.L., Fuhrmann, J.L., Geoghegan, N.S.M., Gnehm, C.L., McDonald, L.A., Small, K.V., Fraser, C.M., Smith, H.O., and Venter, J.C. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269: 496-512.
- Goda, S.K., and Minton, N.P. 1995. A simple procedure for gel electrophoresis and Northern blotting of RNA. *Nucleic Acids Res.* 23: 3357-3358.
- Gorke, B., and Rak, B. 1999. Catabolite control of *Escherichia coli* regulatory protein BglG activity by antagonistically acting phosphorylations. *EMBO J.* 18: 3370-3379.
- Hueck, C.J., and Hillen, W. 1995. Catabolite repression in *Bacillus subtilis*: a global regulatory mechanism for the Gram-positive bacteria? *Mol. Microbiol.* 15: 395-401.
- Kasahara, Y., Nakai, S., Ogasawara, N., Yata, K., and Sadaie, Y. 1997. Sequence analysis of the *groESL-cotA* region of the *Bacillus subtilis* genome, containing the restriction/modification system genes. *DNA Res.* 4: 335-339.
- Krüger, S., Gertz, S., and Hecker, M. 1996. Transcriptional analysis of *bglIPH* expression in *Bacillus subtilis*: Evidence for two distinct pathways mediating carbon catabolite repression. *J. Bacteriol.* 178: 2637-2644.
- Krüger, S., and Hecker, M. 1995. Regulation of the putative *bglIPH* operon for aryl-beta-glucoside utilization in *Bacillus subtilis*. *J. Bacteriol.* 177: 5590-5597.
- Lai, X., and Ingram, L.O. 1995. Discovery of a *ptsHI* operon, which includes a third gene (*ptsT*), in the thermophile *Bacillus stearothermophilus*. *Microbiol.* 141: 1443-1449.
- Le Coq, D., Lindner, C., Krüger, S., Steinmetz, M., and Stülke, J. 1995. New β -glucoside (*bgl*) genes in *Bacillus subtilis*: the *bglP* gene product has both transport and regulatory functions similar to those of BglF, its *Escherichia coli* homolog. *J. Bacteriol.* 177: 1527-1535.
- Li, Y., and Ferenci, T. 1996. The *Bacillus stearothermophilus* NUB36 *surA* gene encodes a thermophilic sucrose related to *Bacillus subtilis* *SacA*. *Microbiol.* 142: 1651-1657.
- Lindner, C., Galinier, A., Hecker, M., and Deutscher, J. 1999. Regulation of the activity of the *Bacillus subtilis* antiterminator LicT by multiple PEP-dependent, Enzyme I- and HPr-catalysed phosphorylation. *Mol. Microbiol.* 31: 995-1006.
- Mathews, D.H., Sabina, J., Zuker, M., and Turner, D.H. 1999. Expanded sequence dependence of thermodynamic parameters provides robust prediction of RNA secondary structure. *J. Mol. Biol.* 288: 911-940.
- Mitchell, W.J., and Booth, I.R. 1984. Characterisation of the *Clostridium pasteurianum* phosphotransferase system. *J. Gen. Microbiol.* 130: 2193-2200.
- Mitchell, W.J., Shaw, J.E., and Andrews, L. 1991. Properties of the glucose phosphotransferase system of *Clostridium acetobutylicum* NCIB 8052. *Appl. Environ. Microbiol.* 57: 2534-2539.
- O' Brien, R.W., and Morris, J.G. 1971. Oxygen and the growth and metabolism of *Clostridium acetobutylicum*. *J. Gen. Microbiol.* 68: 307-318.
- Reid, S.J., Rafudeen, M.S., and Leat, N.G. 1999. The genes controlling sucrose utilization in *Clostridium beijerinckii* NCIMB 8052 constitute an operon. *Microbiol.* 145: 1461-1472.
- Rutberg, B. 1997. Antitermination of transcription of catabolic operons. *Mol. Microbiol.* 23: 413-421.
- Saier Jr., M.H., and Reizer, J. 1992. Proposed uniform nomenclature for the proteins and protein domains of the bacterial phosphoenolpyruvate:sugar phosphotransferase system. *J. Bacteriol.* 174: 1433-1438.
- Schmid, K., Ebner, R., Jahreis, K., Lengeler, J.W., and Titgemeyer, F. 1991. A sugar-specific porin, ScrY, is involved in sucrose uptake in enteric bacteria. *Mol. Microbiol.* 5: 941-950.
- Schnetz, K., Stülke, J., Gertz, S., Krüger, S., Krieg, M., Hecker, M., and Rak, B. 1996. LicT, a *Bacillus subtilis* transcriptional antiterminator protein of the BglG family. *J. Bacteriol.* 178: 1971-1979.
- Schnetz, K., Toloczky, C., and Rak, B. 1987. Beta-glucoside (*bgl*) operon of *Escherichia coli* K12 - nucleotide sequence, genetic organization, and possible evolutionary relationship to regulatory components of two *Bacillus subtilis* genes. *J. Bacteriol.* 169: 2579-2590.
- Sorokin, A., Azevedo, V., Zumstein, E., Galleron, N., Ehrlich, S.D., and Serró, P. 1996. Sequence analysis of the *Bacillus subtilis* chromosome region between the *serA* and *kdg* loci cloned in a yeast artificial chromosome. *Microbiol.* 142: 2005-2016.
- Stülke, J., Arnaud, M., Rapoport, G., and Martin-Verstraete, I. 1998. PRD - a protein domain involved in PTS-dependent induction and carbon catabolite repression of catabolic operons in bacteria. *Mol. Microbiol.* 28: 865-874.
- Sutrina, S.L., Reddy, P., Saier Jr., M.H., and Reizer, J. 1990. The glucose permease of *Bacillus subtilis* is a single polypeptide chain that functions to energise the sucrose permease. *J. Biol. Chem.* 265: 18581-18589.
- Tangney, M., Brehm, J.K., Minton, N.P., and Mitchell, W.J. 1998a. A gene system for glucitol transport and metabolism in *Clostridium beijerinckii* NCIMB 8052. *Appl. Environ. Microbiol.* 64: 1612-1619.
- Tangney, M., Rouse, C., Yazdani, M., and Mitchell, W.J. 1998b. Sucrose transport and metabolism in *Clostridium beijerinckii* NCIMB 8052. *Journal of Applied Microbiol.* 84: 914-919.
- Tangney, M., Smith, P., Priest, F.G., and Mitchell, W.J. 1992. Maltose transport in *Bacillus licheniformis* NCIB 6346. *J. Gen. Microbiol.* 138: 1821-1827.
- Titgemeyer, F., Jahreis, K., Ebner, R., and Lengeler, J.W. 1996. Molecular analysis of the *scrA* and *scrB* genes from *Klebsiella pneumoniae* and plasmid pUR400, which encode the sucrose transport protein Enzyme II Scr of the phosphotransferase system and a sucrose-6-phosphate invertase. *Mol. Gen. Genet.* 250: 197-206.
- Tortosa, P., Aymerich, S., Lindner, C., Saier, M.H., Reizer, J., and LeCoq, D. 1997. Multiple phosphorylation of SacY, a *Bacillus subtilis* transcriptional antiterminator negatively controlled by the phosphotransferase system. *J. Biol. Chem.* 272: 17230-17237.
- Weickert, M.J., and Chambliss, G.H. 1990. Site-directed mutagenesis of a catabolite repression operator sequence in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. USA* 87: 6238-6242.

- Young, M., Minton, N.P., and Staudenbauer, W.L. 1989. Recent advances in the genetics of the clostridia. *FEMS Microbiol. Rev.* 63: 301-326.
- Zukowski, M.M., Miller, L., Cosgwell, P., Chen, K., Aymerich, S., and Steinmetz, M. 1990. Nucleotide sequence of the *sacS* locus of *Bacillus subtilis* reveals the presence of two regulatory genes. *Gene* 90: 153-155.