Brazilian Journal of Microbiology 45, 4, 1139-1144 (2014)
ISSN 1678-4405

Review

Regulation of gene expression: Cryptic β-glucoside (*bgl*) operon of *Escherichia coli* as a paradigm

Dharmesh Harwani

Department of Microbiology, Maharaja Ganga Singh University, Bikaner, India

Submitted: November 06, 2013; Approved: April 17, 2014.

Abstract

Bacteria have evolved various mechanisms to extract utilizable substrates from available resources and consequently acquire fitness advantage over competitors. One of the strategies is the exploitation of cryptic cellular functions encoded by genetic systems that are silent under laboratory conditions, such as the *bgl* (β-glucoside) operon of *E. coli.* The *bgl* operon of *Escherichia coli*, involved in the uptake and utilization of aromatic β-glucosides salicin and arbutin, is maintained in a silent state in the wild type organism by the presence of structural elements in the regulatory region. This operon can be activated by mutations that disrupt these negative elements. The fact that the silent *bgl* operon is retained without accumulating deleterious mutations seems paradoxical from an evolutionary view point. Although this operon appears to be silent, specific physiological conditions might be able to regulate its expression and/or the operon might be carrying out function(s) apart from the utilization of aromatic β-glucosides. This is consistent with the observations that the activated operon confers a *G*rowth *A*dvantage in *S*tationary *P*hase (GASP) phenotype to Bgl^+ cells and exerts its regulation on at least twelve downstream target genes.

Key words: cryptic, Bgl operon, beta-gluosides, GASP.

Introduction

Bacteria are the most successful and the most prevalent creatures on earth. In these natural habitats bacteria are subject to various kinds of stress, such as nutrient scarcity, with occasional availability of food, fluctuations in temperature, pH, osmolarity and severe competition for resources from other organisms. Various elaborate survival strategies are employed by microbes to sense and adjust to the external and internal milieu. According to the neutral theory of evolution, genes that do not contribute towards the fitness of an organism are not subjected to natural selection and are lost by genetic drift (*Kimura, 1968)*. In that way genomes are in a constant state of flux wherein pre-existing genes are lost by means of mutations and genetic drift and new genes are further acquired by horizontal gene transfer as well as mutations in preexisting genes. Cryptic genes are defined as genes that remain silent in the wild type organism but are capable of being activated and expressed by means of certain genetic changes (Hall *et al.*, 1983). These genes are different from pseudo genes since unlike pseudo genes they can be activated to a functional state. There are several known examples of cryptic genes in different organisms, such as the gene for citrate utilization in *E. coli* (Hall, 1982) and alcohol dehydrogenase gene in yeast (Paquin and Williamson 1986). In view of this, maintenance of such genes that do not contribute to the fitness of the organism is enigmatic. One possibility is that such genes are expressed under specific conditions and contribute to the organism's fitness (Thatcher *et al.*, 1998). In the present review the cryptic Bgl operon of *E coli* has been discussed to understand its contribution in conferring fitness advantage to Bgl^+ cells under stress physiological conditions.

Escherichia coli and β-glucosides Utilization

Identified in 1885 by Theodor Escherich, *E. coli* is one of the most well studied species of bacteria. While many strains of *E. coli* are non pathogenic, there are several strains that cause intestinal and extra intestinal infections.

Send correspondence to D. Harwani. Department of Microbiology, Maharaja Ganga Singh University, Bikaner, India. E-mail: dharmesh_harwani@hotmail.com.

E. coli can utilize several carbohydrates, such as phosphorylated sugars, polyols, carboxylates, amino sugars, pentoses, hexoses, dissacharides, and polysaccharidesas as carbon source. However, wild type *E. coli*, like many other members of Enterobacteriaceae, is incapable of utilizing the β-glucosides as a sole source of carbon and energy. The β-glucosides are sugars mostly of plant origin that have a molecule of glucose linked through β-1, 4 linkage to an aliphatic or an aromatic side group. Some of the commonly found β-glucosides are salicin, arbutin, and cellobiose. The side groups in these sugars are 2-hydroxymethylphenyl, 4-hydroxyphenyl and glucose, respectively. Salicin is a secondary metabolite in the leaves of plants from the genus Salix; arbutin is found in the leaves of plants belonging to families Saxifragiceae, Rosaceae and Ericaceae, while cellobiose is a breakdown product of cellulose and lichenin and does not exist free in nature. There is heterogeneity among the members of the family Enterobacteriaceae with respect to their ability to utilize the β-glucosides as a carbon source. While members such as *E. coli, Shigella, and Salmonella* are incapable of fermenting these sugars, there are members such as *Klebsiella*, *Enterobacter*, *Erwinia* and *Citrobacter* which readily metabolize some or all of these sugars (Schaefler, 1967; Schaefler and Malamy, 1969).

Genetic Diversity of β-glucosides Utilization in *E. coli*

Wild type *E. coli* is unable to metabolize β-glucosides in spite of having three genetic systems for their utilization. These three genetic systems of *E. coli*: *bgl*, *asc* and *chb*, are classified as cryptic. Mutational activation of at least one of these systems is required to enable *E. coli* to metabolize these sugars. The *asc* operon, located at 58.7 min of *E. coli* chromosome (Hall *et al.*, 1991), upon being activated also enables the organism to utilize β-glucosides (Parker and Hall, 1988). This operon comprises a putative repressor, *ascG*, a PTS permease, *ascF* and a phospho-β-glucosidase, *ascB* (Hall and Xu, 1992). The *chb* operon of *E. coli,* located at 39 min on the chromosome, is a normal inducible operon for the uptake and utilization of chitobiose (Keyhani and Roseman, 1997). The *chb* operon comprises six ORFs, *chbBCARFG* and a regulatory region, *chbOP*. *chbBCA* encode three domains of the PTS permease, *chbR* encodes an activator that also acts as a repressor, *chbF* codes for phospho-glucosidase and *chbG* does not have any known function. ChbR, CAP and NagC have been implicated in the regulation of the *chb* operon by chitobiose (Plumbridge and Pellegrini, 2004). The *bgl* operon of *E. coli* (first studied by Schaefler is positioned at 83.8 min on the *E. coli* chromosome (Bachmann, 1990). The operon comprises three structural genes, *bglG*, *bglF* and *bglB* and a regulatory region *bglR* (Figure 1) (Mahadevan *et al.*, 1987; Schnetz *et al.*, 1987). The first gene of the operon, *bglG*, encodes an antiterminator that acts at two *rho* independent terminators flanking *bglG* (Mahadevan and Wright, 1987; Schnetz and Rak, 1988). The following gene, *bglF*, encodes a PTS permease that phosphorylates and transports the β-glucosides, salicin and arbutin. In the absence of the inducer BglF phosphorylates BglG, preventing its antiterminator function, thereby acting as a negative regulator of the *bgl* operon (Amster-Choder *et al.*, 1989; Schnetz and Rak, 1990). The last gene of the operon, *bglB,* encodes a phospho-β-glucosidase that cleaves phosphorylated salicin and arbutin. In addition to these three ORFs, the *bgl* operon also comprises another gene, *bglH*, which is not essential for the utilization of the β-glucosides. BglH is as sociated with outer membrane and is a porin, specific for the uptake of carbohydrates (Andersen *et al.*, 1999). In spite of being intact at the genetic level, the *bgl* operon is kept silent in the wild type organism due to the presence of certain negative structural elements in the regulatory region, *bglR* (Lopilato and Wright, 1990; Schnetz, 1995; Singh *et al.*, 1995; Schnetz and Wang, 1996; Mukerji and Mahadevan, 1997).

Mutations that Activate the *bgl* Operon

A variety of mutations, that act in cis or trans, can activate the silent *bgl* operon of *E. coli*, enabling the bacteria to utilize the β-glucosides, salicin and arbutin (Reynolds *et al.*, 1981; Reynolds *et al.*, 1986; Di Nardo *et al.*, 1982; Higgins *et al.*, 1988; Schnetz and Rak, 1992; Giel *et al.*, 1996). A single mutational event is sufficient to activate this operon. The most commonly occurring activating mutations for this operon are insertions of IS*1* or IS*5* in a 223 base pair

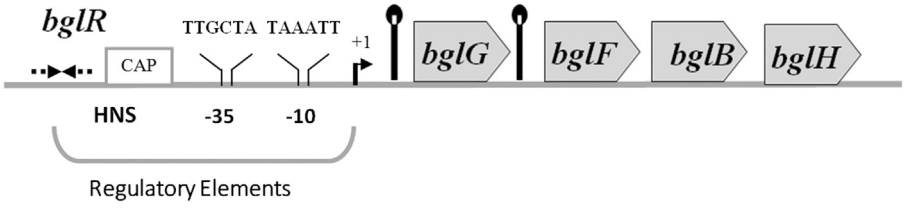

Figure 1 - Schematics for the *bgl* operon of *E. coli*. The operon comprises three structural genes, *bglG*, *bglF* and *bglB* and a regulatory region *bglR*. The first gene of the operon, *bglG* encodes an antiterminator that acts at two *rho* independent terminators. The next gene *bglF* encodes a PTS permease and a negative regulator of the *bgl* operon. The last gene of the operon, *bglB,* encodes a phospho-β-glucosidase. In addition to these three ORFs, the *bgl* operon also comprises another gene, *bglH*, which is not essential for the utilization of the β-glucosides.

sequence in the regulatory region of the *bgl* operon and also in some downstream sequences (Reynolds *et al.*, 1986; Di Nardo *et al.*, 1982; Higgins *et al.*, 1988). The insertion elements do not provide promoter element to the *bgl* operon (Di Nardo *et al.*, 1982) but activate the operon by disrupting the negative elements from the *bgl* promoter (Lopilato and Wright, 1990; Singh *et al.*, 1995). Point mutations in the binding site for *C*atabolite *A*ctivator *P*rotein (CAP), which brings this site closer to the consensus CAP binding sequence also have been shown to activate the *bgl* operon (Di Nardo *et al.*, 1982; Lopilato and Wright, 1990). These point mutations result in a higher affinity binding of CAP and exclusion of H-NS binding, since CAP and H-NS binding sites in the *bglR* are overlapping (Mukerji and Mahadevan, 1997). Mutations in the *hns* locus are also known to activate the *bgl* operon, since H-NS acts as a negative regulator of this operon (Defez and Felice, 1981; Higgins *et al.*, 1988). Change in the supercoiling status of DNA has also been shown to affect the expression of the *bgl* operon. Mutations in *gyrA* (48 min) and *gyrB* (83 min) loci, that are expected to reduce DNA supercoiling, are known to activate the *bgl* operon (Di Nardo *et al.*, 1982). Reduced super coiling is expected to destabilize the cruciform structure in the *bgl* regulatory region, thereby lifting the negative regulation from the *bgl* operon and allowing it to be expressed at a higher level. This is consistent with the observation that point mutations within the inverted repeat activate the *bgl* promoter and inhibition of gyrase fails to enhance the expression further (Mukerji and Mahadevan, 1997). In addition, mutations that lead to the over expression of LeuO or BglJ have been shown to activate the *bgl* operon (Giel *et al.*, 1996; Ueguchi *et al.*, 1998). The *bgl* operon is subject to induction by the β-glucosides after mutational activation. BglG and BglF encoded by the *bgl* operon bring about this second level of regulation (Mahadevan, 1997).

Growth Advantage in Stationary Phase (GASP)

It has been shown that bacterial population can be maintained at counts of about 10^6 colony forming units (CFUs) per ml for several years without the addition of fresh nutrients (Finkel, 2006). This is a highly dynamic phase wherein several population take over occur and the culture becomes highly heterogeneous. If bacteria are starved for prolonged periods of time, 99% of the population dies in a phase commonly known as the Death phase. The remaining 1% of the population not only remains alive but also grows during a phase now known as prolonged stationary phase. It has been demonstrated that the GASP phenotype of the older cultures is due to genetic changes in the population (Zambrano *et al.*, 1993). The occurrence of GASP is a continuous phenomenon wherein older culture will always take over the younger culture. For example a 10-day-old culture takes over a one day old culture and a twenty day old culture takes over a ten day old culture and so on (Zambrano and Kolter 1996; Finkel *et al.*, 1997).

The Cryptic Bgl Operon Regulates *oppA* an Oligopeptide Transporter

Global analysis of intracellular proteins from Bgl^+ and Bgl^- strains revealed that the operon exerts regulation on at least twelve downstream target genes. Of these, *oppA*, which encodes an oligo-peptide transporter, was confirmed to be up-regulated in the Bgl^+ condition (Harwani *et al.*, 2012). Since the oligopeptide transporter (*oppA*) has been shown to be up-regulated in the Bgl^+ strain, it is conceivable that the functions encoded by *oppA* contribute to the GASP phenotype exhibited by Bgl^+ strains (Harwani *et al.*, 2012). Interestingly the ZK819-97TΔ*oppA* cells have lost the strong fitness advantage shown by the parent strain ZK819-199 97T in co-culture experiment. This suggests that a part of the growth advantage in stationary phase of the Bgl^+ strain is contributed by OppA. The involvement of the *bgl* operon in the regulation of OppA expression could be direct or indirect. OppA is regulated negatively by a small regulatory RNA (sRNA) *gcvB* (Argaman *et al.*, 2001) which has been shown to inhibit translation initiation by binding to the *oppA* mRNA (Sharma *et al.*, 2007). In turn, the transcription of *gcvB* is positively regulated by the GcvA protein, the major transcription factor of the glycine cleavage system (Urbanowski *et al.*, 2000). Expression of *gcvB* is high during early log phase, but its level decreases during cell growth (Argaman *et al.*, 2001). This reduction in *gcvB* expression was much more pronounced in Bgl^+ cells. Similarly, a significant decrease in *gcvA* transcription in Bgl^+ cells was also registered in the stationary phase (Harwani *et al.*, 2012).

These observations suggest that the regulation of *oppA* by the *bgl* operon is via its regulators *gcvA* and *gcvB* (Figure 2). In view of this it has been proposed that the ability to transport oligo-peptides, mediated by the over expression of *oppA,* is partly responsible for the GASP phenotype exhibited by Bgl^+ strains. Down-regulation of *oppA* in a strain carrying a deletion of *bglG* may be one of the reasons for the loss of the GASP phenotype of the ΔbglG strain. The complete loss of the GASP phenotype in the ΔbglG mutant and its partial rescue in the ΔbglGΔgcvA double mutant suggest that BglG is a master regulator involved in modulating the expression of downstream genes important in stationary phase survival and *oppA* is one such locus (Harwani *et al.*, 2012). It has been shown that the BglG decreases *gcvA* mRNA stability and suggests a specific role **for** BglG-mediated post transcriptional regulation at this locus and yet at other unexplored loci (Figure 2). The molecular mechanism by which BglG/$gcvA^{RAT}$ duplex is exposed to the degradation machinery is still unknown. The regulatory role exerted by BglG on *gcvA* to control OppA translation

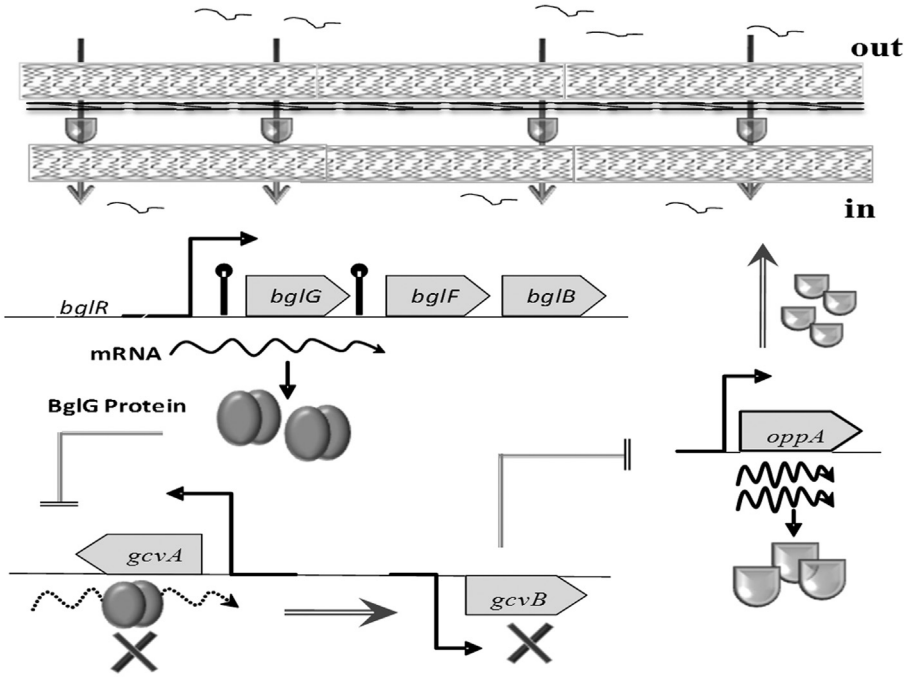

Figure 2 - Schematics for the genetic regulation mediated by the activated *bgl* operon of an *E. coli* on *oppA.* In a co-culture experiment Bgl^+ cells exhibit GASP phenotype over Bgl^- cells under stationary phase growth condition. BglG interacts at *gcvA* mRNA to destabilize it which is known to positively regulates GcvB. In absence of GcvB, its negative regulation on *oppA* is relieved (GcvB is a translational repressor of *oppA*). The elevated level of oligopeptide transporter can now facilitate transport of small peptides.

would enhance our understanding of the BglG-mediated signalling process.

Post-transcriptional regulation mediated by BglG on *gcvA* mRNA leads to its destabilization that affects GcvA translation negatively. The reduced level of GcvA leads to the reduced transcription of *gcvB* (translational repressor of OppA), and the increased translation of OppA. Thus, elevated levels of OppA in the Bgl^+ strain facilitate transport of oligo-peptides, conferring a GASP phenotype over the wild type Bgl^- counterpart (Figure 2) (Harwani *et al.*, 2012).

Conclusion

The maintenance of genetic systems which are apparently of no selective advantage to the organism is an evolutionary paradox. This applies to `cryptic' genes, since they do not seem to function in the wild type organism and require mutational activation for expression. Does their retention in the wild type organism have any physiological significance? Are these genes truly silent or there are specific physiological conditions that can induce their expression? A possible mechanism has been documented by which *oppA* confers a competitive advantage to Bgl^+ cells relative to Bgl^- cells in the stationary phase (as described above). The involvement of the *bgl* operon in functions unrelated to the catabolism of β-glucosides suggests that selection for elevated expression of the operon can occur even in the absence of β-glucosides. This could be achieved by either by mutations or by overriding its negative regulation under specific growth conditions such as stationary phase. Though such elevated expression may not be sufficient to allow utilization of β-glucosides, it may be sufficient for the regulation of the downstream target genes, providing a selective force for the maintenance of the *bgl* genes over evolutionary time. Conclusively *bgl* operon has been highlighted for its involvement in the functions unrelated to the β-glucosides utilization. This could be one of the possible signal transduction mechanisms by which bacteria might modulate gene expression upon starvation stimuli. The present review help strengthen the notion that rather than a "cryptic" genetic element, the *bgl* operon should be considered as a dynamic component of the *E. coli* genome.

Acknowledgments

Author is highly thankful to Prof. S. Mahadeven for his constant encouragement and research guidance. Thanks are also due to the DBT, India for providing research fellowship (RA-III).

References

Amster-Choder O, Houman F, Wright A(1989). Protein phosphorylation regulates transcription of the β-glucoside utilization operon in *E. coli*. Cell 58:847-855.

Andersen C, Rak B, Benz R (1999) The gene *bglH* present in the *bgl* operon of *Escherichia coli*, responsible for uptake and

fermentation of β-glucosides encodes for a carbohydrate-specific outer membrane porin. Mol Microbiol 31:499-510.

Argaman L, Hershberg R, Vogel J, Bejerano G, Wagner EG, Margalit H, Altuvia S (2001) Novel small RNA-encoding genes in the intergenic regions of *Escherichia coli*. Curr Biol 11:941-950.

Bachmann BJ (1990) Linkage Map of *Escherichia coli* K-12. Microbiol Rev 54:130-197.

Defez R, de Felice M (1981) Cryptic operon for ß-glucoside metabolism in *Escherichia coli* K12: genetic evidence for a regulatory protein. Genetics 97:11-25.

Di Nardo S, Voelkel KA, Sternglanz R, Reynolds AE, Wright A (1982) *Escherichia coli* DNA topoisomerase I mutants have compensatory mutations in the DNA gyrase genes. Cell 31:43-51.

Finkel SE (2006) Long-term survival during stationary phase: evolution and the GASP phenotype. Nature Rev Microbiol 4:113-120.

Finkel SE, Zinser E, Gupta S, Kolter R (1997) Life and death in stationary phase. *In:* Busby, S.J.W., Thomas, C.M., Brown, N.L. (eds). Molecular microbiology. Springer-Verlag, Berlin, Germany, 3-16.

Giel M, Desnoyer M, Lopilato J (1996) A mutation in a new gene, *bgl J*, activates the *bgl* operon in *Escherichia coli* K-12. Genetics 143:627-635.

Hall BG (1982) A chromosomal mutation for citrate utilization by *Escherichia coli* K-12. J Bacteriol 152:269-273.

Hall BG, Xu L (1992) Nucleotide sequence, function, activation, and evolution of the cryptic asc operon of *Escherichia coli* K-12. Mol Biol Evol 9:688-706.

Hall BG, Xu L, Ochman H (1991) Physical map location of the *asc* operon of *Escherichia coli* K-12. J Bacteriol 173:5250.

Hall BG, Yokoyama S, and Calhoun DH (1983) Role of cryptic genes in microbial evolution. Mol Biol Evol 1:109-124.

Harwani D, Zangoui P, Mahadevan S (2012) The β-glucoside (bgl) operon of *Escherichia coli* is involved in the regulation of *oppA* encoding an oligo-peptide transporter. J Bacteriol 194:90-99.

Higgins CF, Dorman CJ, Stirling DA, Waddell L, Booth IR, May G, Bremer E (1988) A physiological role for DNA supercoiling in the osmotic regulation of gene expression in *S. typhimurium* and *E. coli*. Cell 52:569-584.

Keyhani NO, Roseman S (1997) Wild type *Escherichia coli* grows on the chitin disaccharide, N, N'-diactylchitobiose, by expressing the *cel* operon. Proc Natl Acad Sci 94:14367-14371.

Kimura M (1968) Evolutionary rate at the molecular level. Nature 217:624-626.

Lopilato J, and Wright A (1990) Mechanisms of activation of the cryptic *bgl* operon of Escherichia coli K-12. *In:* Edited by Drilca, K., Riley, M. (eds). The Bacterial Chromosome. Washington DC, American Society for Microbiology, pp 435-444.

Madan R, Kolter R, Mahadevan S (2005) Mutations that activate the silent *bgl* operon of *Escherichia coli* confer a growth advantage in stationary phase. J Bacteriol 187:7912-7917.

Mahadevan S (1997) The BglG group of antiterminators: a growing family of bacterial regulators. J Biosci 22:505-513.

Mahadevan S, Wright A (1987) A bacterial gene involved in transcription antitermination: regulation at a rho-independent terminator in the *bgl* operon of *E. coli*. Mahadevan. Cell 50:485-494.

Mahadevan SA, Reynolds E, Wright A (1987) Positive and negative regulation of the *bgl* operon in *Escherichia coli.* J Bacteriol 169:2570-2578.

Mukerji M. Mahadevan S (1997) Characterization of the negative elements involved in silencing the *bgl* operon of *Escherichia coli*: possible roles for DNA gyrase, H-NS, and CRP-cAMP in regulation. Mol Microbiol 24:617-627.

Paquin CE, and Williamson VM (1986) Ty insertions at two loci account for most of the spontaneous antimycin A resistance mutations during growth at 15 °C of *Saccharomyces cerevisiae* strains lacking ADH1. Mol Cell Biol 6:70-79.

Parker LL, Hall BG (1988) A fourth *Escherichia coli* gene system with the potential to evolve β-glucoside utilization. Genetics 119:485-490.

Plumbridge J, Pellegrini O (2004) Expression of the chitobiose operon of *Escherichia coli* is regulated by three transcription factors: NagC, ChbR and CAP. Mol Microbiol 52:437-449.

Reynolds A, Mahadevan S, LeGrice SF, Wright A (1986) Enhancement of bacterial gene expression by insertion elements or by mutation in a CAP-cAMP binding site. J Mol Biol 191:85-95.

Reynolds A. Felton EJ, Wright A (1981) Insertion of DNA activates the cryptic *bgl* operon in *E. coli* K12. Nature. 293:625-629.

Schaefler S (1967) Inducible system for the utilization of β-glucosides in *Escherichia coli*, active transport and utilization of β-glucosides. J Bacteriol 93:254-263.

Schaefler S and Malamy A (1969) Taxonomic investigation on expressed and cryptic phospho-β-glucosidases in Enterobacteriaceae. J Bacteriol 99:422-433.

Schnetz K (1995) Silencing of *Escherichia coli bgl* promoter by flanking sequence elements. EMBO J 14:2545-2550.

Schnetz K, and Rak B (1988) Regulation of the *bgl* operon of *Escherichia coli* by transcriptional antitermination. EMBO J. 7:3271-3277.

Schnetz K, and Rak B (1990). β-glucoside permease represses the *bgl* operon of *Escherichia coli* by phosphorylation of the antiterminator protein and also interacts with glucose-specific enzyme III, the key element in catabolite control. Proc Natl Acad Sci USA 87:5074-5078.

Schnetz K, Rak B (1992) IS5: a mobile enhancer of transcription in *Escherichia coli*. Proc Natl Acad Sci USA 89:1244-1248.

Schnetz K, Tolocyzki C, Rak B (1987) β-glucoside (*bgl*) operon of *Escherichia coli* K12: nucleotide sequence, genetic organisation, and possible evolutionary relationship to regulatory components of two *Bacillus subtilis* genes. J Bacteriol 169:2579-2590.

Schnetz K, Wang JC (1996). Silencing of the *Escherichia coli bgl* promoter: effects of template supercoiling and cell extracts on promoter activity in vitro. Nucleic Acids Res 24:2422-8.

Sharma CM, Darfeuille F, Plantinga TH, Vogel J (2007) A small RNA regulates multiple ABC transporter mRNAs by targeting C/A-rich elements inside and upstream of ribosome-binding sites. Genes Dev 21:2804-2817.

Singh J, Mukerji M, Mahadevan S (1995) Transcriptional activation of the *Escherichia coli bgl* operon: negative regulation by DNA structural elements near the promoter. Mol Microbiol 17:1085-1092.

Thatcher J, Shaw W, Janet M, Dickinson WJ (1998) Marginal fitness contributions of non essential genes in yeast. Proc Natl Acad Sci USA 95:253-257.

Ueguchi C, Ohta T, Seto C, Suzuki T, Mizuno T (1998) The *leuO* gene product has a latent ability to relieve *bgl* silencing in *Escherichia coli*. J Bacteriol 180:190-193.

Urbanowski ML, Stauffer LT, and Stauffer GV (2000) The *gcvB* gene encodes a small untranslated RNA involved in expression of the dipeptide and oligopeptide transport systems in *Escherichia coli*. Mol Microbiol 37:856-868.

Zambrano MM, Kolter R (1996) GASPing for life in stationary phase. Cell 86:181-184.

Zambrano MM, Siegele DA, Almirón M, Tormo A, Kolter R (1993) Microbial competition: *Escherichia coli* mutants that take over stationary phase cultures. Science 259:1757-1760.