



HAL
open science

Escherichia coli glycogen genes are organized in a single glgBXCAP transcriptional unit possessing an alternative suboperonic promoter within glgC that directs glgAP expression

Manuel Montero, Goizeder Almagro, Gustavo Eydallin, Alejandro M Viale, Francisco José Muñoz, Abdellatif Bahaji, Jun Li, Mehdi Rahimpour, Edurne Baroja-Fernández, Javier Pozueta-Romero

► To cite this version:

Manuel Montero, Goizeder Almagro, Gustavo Eydallin, Alejandro M Viale, Francisco José Muñoz, et al.. Escherichia coli glycogen genes are organized in a single glgBXCAP transcriptional unit possessing an alternative suboperonic promoter within glgC that directs glgAP expression. *Biochemical Journal*, 2010, 433 (1), pp.107-117. 10.1042/BJ20101186 . hal-00547994

HAL Id: hal-00547994

<https://hal.science/hal-00547994>

Submitted on 18 Dec 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Title: *Escherichia coli* glycogen genes are organized in a single *glgBXCAP* transcriptional unit possessing an alternative suboperonic promoter within *glgC* that directs *glgAP* expression

Authors: Manuel Montero*, Goizeder Almagro*, Gustavo Eydallin, Alejandro M. Viale, Francisco José Muñoz, Abdellatif Bahaji, Jun Li, Mehdi Rahimpour, Edurne Baroja-Fernández and Javier Pozueta-Romero

Affiliation: Instituto de Agrobiotecnología (CSIC, UPNA, Gobierno de Navarra). Mutiloako etorbidea zenbaki gabe, 31192 Mutiloabeti, Nafarroa, Spain.

Key words: Gene regulation, glycogen metabolism, PhoP-PhoQ regulatory system

Abbreviations: ADPG, ADP-glucose; cAMP, cyclic AMP; CRP, cAMP receptor protein; GlgA, glycogen synthase; GlgB, branching enzyme; GlgC, ADPG pyrophosphorylase; GlgP, glycogen phosphorylase; GlgX, debranching enzyme; MOPS, morpholinepropanesulphonic acid; ppGpp, guanosine 5'-diphosphate 3'-diphosphate; RACE, rapid amplification of cDNA ends; RT, reverse transcriptase; TAP, tobacco acid pyrophosphatase; UTR, untranslated region; WT, wild type.

Corresponding author: Javier Pozueta Romero.

Address: Instituto de Agrobiotecnología. Mutiloako etorbidea zenbaki gabe, 31192 Mutiloabeti, Nafarroa, Spain. E-mail: javier.pozueta@unavarra.es. Phone: (34) 948168009. FAX: (34) 948232191

FOOTNOTE: A. M. V. present address is "Instituto de Biología Molecular y Celular de Rosario (IBR, CONICET), Departamento de Microbiología, Facultad de Ciencias Bioquímicas y Farmacéuticas, Universidad Nacional de Rosario, Suipacha 531, 2000 Rosario, Argentina". M. M. and G. A. have equally contributed to this work. A. B. present address is "Iden Biotechnology S. L., Pol. Ind. Mocholí, Plaza CEIN 5, nave B-3. 31110 Noáin, Nafarroa, Spain"

Short title: *E. coli* glycogen genes are organized in a single transcriptional unit

ABSTRACT

Although it is generally accepted that *E. coli* glycogen genes are organized in two tandemly arranged, differentially regulated *glgBX* and *glgCAP* operons, RT-PCR analyses carried out in this work showed that *E. coli* cells possess transcripts comprising the five *glgBXCAP* genes. *glg::lacZY* expression analyses in cells lacking the region immediately upstream of the *glgB* gene revealed an almost total abolishment of *glgB*, *glgX* and *glgC* expression, but only a 50-60% reduction of the wild type *glgA* and *glgP* expression levels. Furthermore, similar type of analyses showed that *glgA* and *glgP* expression was almost totally abolished in cells lacking *glgA* upstream sequences including *glgC*, *glgB* and the *asd-glgB* intergenic region upstream of *glgB*. The overall data thus indicated that *E. coli glgBXCAP* genes are organized in a single transcriptional unit controlled by promoter sequences occurring upstream of *glgB*, and that an alternative suboperonic promoter is located within *glgC* driving expression of *glgA* and *glgP* genes. Computer searches for consensus promoters, and analyses of *glgB::lacZY* and *glgA::lacZY* expression in cells containing deletions of *glgB* and *glgA* upstream sequences identified regions directing *glgBXCAP* and *glgAP* expression. 5' RACE analyses located a *glgBXCAP* transcription start site 155 bp upstream of the *glgB* initiation codon, and a *glgAP* transcription start site 359 bp upstream of the *glgA* initiation codon. Finally, *glg::lacZY* expression analyses on cells lacking *relA* or *phoP* regulatory genes indicated that both *glgBXCAP* operon and the suboperonic promoter driving *glgAP* expression form part of both the RelA and the PhoP-PhoQ regulons.

INTRODUCTION

Glycogen is a branched homopolysaccharide of α -1,4-linked glucose subunits with α -1,6-linkages at the branching points. The exact role of this polyglucan in bacteria is not as clear-cut as in animal and yeast cells [1-3], but several works have linked glycogen metabolism to environmental survival, symbiotic performance and colonization and virulence [4-9]. Synthesized by glycogen synthase (GlgA) using ADP-glucose (ADPG) as the glucosyl moiety donor, glycogen accumulation in *Escherichia coli* occurs when cellular carbon sources are in excess under conditions of other nutrients deficiency [10-14]. Although it has been widely accepted that ADPG is exclusively synthesized by ADPG pyrophosphorylase (GlgC) [12], recent reports have provided strong evidence about the occurrence of important sources, other than GlgC, of ADPG linked to glycogen biosynthesis in *E. coli* [15,16]. In this context, our experimental data revealed the existence of enzyme(s) capable of catalyzing the conversion of the glucose-1-phosphate synthesized by phosphoglucomutase into ADPG [15].

One of the salient features in bacterial gene expression is that genes of related functions are often clustered in a single transcriptional unit or operon. Unlike other bacterial species possessing a single glycogen operon [4,5,17,18], different experimental evidence has led to the hypothesis that genes involved in *E. coli* and *Salmonella enterica* glycogen metabolism are clustered in two tandemly arranged operons: *glgBX* (encompassing the genes coding for glycogen branching (GlgB) and debranching (GlgX) enzymes), and *glgCAP* (encoding the GlgC and GlgA anabolic enzymes as well as the catabolic glycogen phosphorylase (GlgP)) (reviewed in [11]). Main data supporting this hypothesis can be summarized as follows: First, S1 nuclease protection assays identified up to four different transcripts initiating within a 0.5 kbp region upstream of *glgC* [19], suggesting that *glgCAP* promoter sequences are located within *glgX*, although no *E. coli* consensus promoter sequences could be found in this region [19]. Second, *in vitro* assays employing S-30 *E. coli* extracts and the *glg* locus cloned in a plasmid vector showed that, unlike *glgB*, *glgC* and *glgA* coupled transcription-translation is enhanced by ppGpp [19]. These observations, however, conflict with recent transcriptomic analyses showing that expression of the five *glg* genes is regulated by ppGpp [20]. Third, expression of *glgC* and *glgA*, but not that of *glgB*, is stimulated by exogenously added cyclic AMP (cAMP) in maxicells derived from bacteria bearing the complete *E. coli glg* locus cloned in plasmids [21], which was consistent with mobility shift assays that identified a putative cAMP/cAMP receptor protein (CRP) binding site within the *glgX* gene [19]. Fourth, *in vitro* assays employing S-30 *E. coli* extracts and the *glg* locus cloned in a plasmid vector showed that, unlike *glgB*, *glgC* and *glgA* coupled transcription-translation is enhanced by the cAMP/CRP complex [19]. However, although the latter two observations pointed to the involvement of cAMP/CRP in regulating *E. coli glgCAP* expression, recent transcriptomic analyses conducted in two different *E. coli* K-12 genetic backgrounds failed to identify *glg* genes as members of the CRP regulon [22-24]. It is worth noting also that earlier experimental evidence obtained by other authors using chromosomally-located *glgA::lacZ* gene fusions and different *E. coli* mutant backgrounds argued against a significant *in vivo* role of cAMP and *relA* in regulating *E. coli glgCAP* expression at the transcriptional level [25].

Given the above contradictory data regarding regulatory aspects of the expression of glycogen genes in *E. coli*, we decided to characterize in detail *glg* genes transcription using reverse transcriptase (RT)-PCR and *glg::lacZY* fusion approaches. We have also identified *glgBXCAP* transcriptional regulatory regions by both deletion analyses on cells bearing *glg::lacZY* fusions and 5' rapid amplification of cDNA ends

(RACE). In addition, we have investigated the expression patterns of the five *glgBXCAP* genes in different transcriptional regulator mutants. With the overall data amassed, we show that the five *E. coli glgBXCAP* genes are transcribed in a single transcriptional unit under the control of promoter sequences occurring upstream of *glgB*. Moreover, we show the occurrence of an alternative promoter located within *glgC* that controls the expression of *glgA* and *glgP*.

Accepted Manuscript

THIS IS NOT THE VERSION OF RECORD - see doi:10.1042/BJ20101186

EXPERIMENTAL PROCEDURES

E. coli mutants and culture conditions

Strains, mutants and plasmids used in this work are shown in **Table 1**. *E. coli* K-12 derivative BW25113 single-gene knockout mutants were obtained from the Keio collection [26]. Unless otherwise indicated, cells were grown at 37°C with rapid gyratory shaking in liquid Kornberg medium (1.1% K₂HPO₄, 0.85% KH₂PO₄, 0.6% yeast extract from Duchefa, Haarlem, The Netherlands) supplemented with 250 µM MgCl₂, 50 mM glucose, and the appropriate selection antibiotic, or in morpholinepropanesulphonic acid (MOPS) medium supplemented with 50 mM glucose and amino acids at the concentrations described in Traxler et al. [20], after inoculation with 1 volume of an overnight culture for 50 volumes of fresh medium.

RNA extraction and RT-PCR

Total RNA from cells entering the stationary phase was extracted using the GenElute Bacterial Total RNA Miniprep kit (Sigma-Aldrich). For the synthesis of cDNA, 1 µg of DNase I (Takara) treated RNA was incubated 1 h at 43°C in 20 µL of Expand RT (Roche) assay mixture containing 5 µM oligonucleotide O1 (**Supplemental Table 1**). For PCR amplification of the resulting cDNA, 2 µL of the RT assay mixture was added to 20 µL of a PCR cocktail containing Taq polymerase (Biotools) and either 0.5 µM each of O2 and O3 (for PCRI), O4 and O5 (for PCRII), O6 and O7 (for PCRIII), or O8 and O9 (for PCRIV) (**Supplemental Table 1**). Using a programmable thermal cycler we performed amplification for 30 cycles, which consisted of denaturation at 94°C for 1 min, annealing at 55°C for 30 sec, and extension for 3 min, with a final extension at 72°C for 7 min. Two µL of reverse transcription reactions performed without RT were used as negative controls in the RT-PCR to rule out the possibility that our results are ascribed to contaminating DNA traces. See **Figure 1B** for schematic representation of the oligonucleotides annealing positions and the expected PCR products.

Identification of transcription start sites by 5'-RACE

glgBXCAP transcription start site was determined using 5'-RACE on extracted bacterial RNA after treatment or not with tobacco acid pyrophosphatase (TAP) as indicated in [27]. Briefly, 10 µg of DNaseI (Takara) treated RNA were subjected to the action of TAP (Epicentre Technologies). After phenol chloroform extraction and ethanol precipitation, RNA was resuspended in 15 µL of water together with 500 pmol of an RNA adapter (5'-GAUAUGC GCGAAUCCUGUAGAACGAACACUAGAAGAAA-3'), which was ligated to the bacterial RNA using T4 RNA ligase (New England Biolabs) following the indications of the manufacturers. After phenol chloroform extraction, the RNA was resuspended in 20 µL of diethyl pyrocarbonate treated water. For identification of the main *glgBXCAP* transcription start site, 5 µL of the RNA preparation were included in 20 µL RT reaction using the *glgB* specific O10 oligonucleotide (**Supplemental Table 1**). Two µL of the RT reaction were then used for PCR using the O11 oligonucleotide annealing in *glgB*, and the O14 oligonucleotide annealing at the RNA adapter (**Supplemental Table 1**). For identification of the transcription start site at the *glgBXCAP* suboperonic promoter, 5 µL of the RNA preparation were included in 20 µL RT reaction using the *glgA* specific O12 oligonucleotide (**Supplemental Table 1**). Two µL of the RT reaction were then used for PCR using the O13 oligonucleotide annealing in *glgA*, and the O14 oligonucleotide annealing at the RNA adapter (**Supplemental Table 1**). Amplicons obtained from TAP-treated samples were cloned into pGEM-T (Invitrogen) and transformed into XL-1 blue

cells. The amplicon sequences present in different clones were then determined at the Secugen S.A. DNA sequence service (Madrid).

Disruption of genomic sequences

Δ *up-glgB* cells (lacking the 253 bp sequence located at positions -270 to -17 immediately upstream the *glgB* gene), Δ *glgBX** cells (Δ *glgBX* cells lacking the 270 bp region upstream *glgB*), Δ *glgBXC** cells (Δ *glgBXC* cells lacking the 270 bp region upstream *glgB*) and different deletions of the *glgB* and *glgA* upstream regions were produced following the method of Datsenko and Wanner [28] (see **Supplemental Figure 1**). A selectable antibiotic resistance gene was generated by PCR from a freshly isolated colony of *E. coli* MC4100 containing a spectinomycin resistance cassette, using 80 nucleotide-long primer pairs that included 60 nucleotides homology extensions for the targeted locus and 20 nucleotides priming sequences for the resistance gene (**Supplemental Table 2**).

LacZY transcriptional fusions

The kanamycin resistance cassette of Δ *glgB*, Δ *glgX*, Δ *glgC*, Δ *glgA* and Δ *glgP* cells of the Keio collection [26] was removed by using temperature-sensitive plasmid pCP20 carrying the FLP recombinase [29]. The scar sequence left after removal of the resistance cassette included a 34-nucleotide FRT site [26], which was used to build *lacZY* transcriptional fusions as reported in [30]. Briefly, Δ *glgB*, Δ *glgX*, Δ *glgC*, Δ *glgA* and Δ *glgP* cells of the Keio collection carrying pCP20 plasmid were transformed with pKG137, which has functional *lacZY* and Km^R cassette that integrated in the proper orientation at the FRT site by the action of the FLP recombinase, yielding to *lacZY* transcriptional fusions where the original resistance cassette of the Keio collection was previously placed. Transcriptional fusions were P1-transduced [31] into different mutants as needed. All fusions were verified by PCR using an oligonucleotide (5'-TTCAGGCTGCGCAACTGTTGG-3') annealing within *lacZ* (+150 bp reverse orientation) and oligonucleotides specifically annealing at positions 500 bp upstream of the insertion point. *lacZY* fusions inserted at the expected sites yielded ca. 750 bp PCR amplification products.

Western blot analyses

Bacterial extracts were separated by 10% SDS-PAGE, transferred to PVDF filters, and immunodecorated by using rabbit antisera raised against GlgC [16] and a goat anti-rabbit IgG alkaline phosphatase conjugate (Sigma).

Analytical procedures

Bacterial growth was followed spectrophotometrically by measuring the absorbance at 600 nm. β -galactosidase activity was measured and reported in the graphics as described by Miller [32]. Protein content was measured by the Bradford method using a Bio-Rad prepared reagent.

Computer analyses

Promoters were predicted using the BPROM program (SoftBerry, Mt. Kisco, NY; <http://linux1.softberry.com/berry.phtml?topic=bprom&group=programs&subgroup=gfi> ndb).

RESULTS

E. coli cells possess transcripts comprising the five *glg* genes

Genes involved in glycogen metabolism are clustered in a single chromosomal locus of the *E. coli* genome in the order *glgBXCAP* (**Figure 1A**). In this work we investigated the possible occurrence of transcripts comprising the five *glg* genes in *E. coli*. Towards this end, we reverse transcribed total *E. coli* RNA using as primer an oligonucleotide (O1) annealing into the 5' region of *glgP* transcripts (for further details see Experimental Procedures) (**Figure 1B**). The resulting cDNAs were then subjected to PCR analyses using as primers oligonucleotides corresponding to sequences occurring in *glgB* (O2 and O3 for PCR I), *glgX* and *glgC* (O4 and O5, respectively, for PCR II), *glgX* and *glgA* (O6 and O7, respectively, for PCR III), and *glgC* and *glgA* (O8 and O9, respectively, for PCR IV) (**Figure 1B**). As shown in **Figure 1C**, PCR I amplified a fragment whose size (ca. 2.1 kbp) corresponds to the whole *glgB* gene. PCR II amplified a ca. 2.0 kbp fragment corresponding to the whole *glgX* gene and a *glgC* fragment. PCR III amplified a ca. 1.8 kbp fragment corresponding to a *glgX* fragment, the whole *glgC* gene, and a *glgA* fragment. Finally, PCR IV amplified a ca. 2.7 kbp fragment corresponding to the whole *glgC* and *glgA* genes. The overall data thus showed that *E. coli* cells possess single transcripts comprising all the five *glg* genes.

The five *glg* genes are transcribed from a single operon under the control of promoter sequence(s) located upstream of *glgB*

Whether expression of the five *glg* genes is controlled by promoter sequences occurring upstream of the *glgB* gene was explored by analyzing the expression of chromosomal *glg::lacZY* transcriptional fusions on both wild type (WT) and on $\Delta_{up-glgB}$ cells lacking 253 bp of the *asd-glgB* intergenic region located upstream of *glgB* (nucleotide position -270 to -17 from the *glgB* translation initiation codon) (**Supplemental Figure 1**). As shown in **Supplemental Figure 2**, the five *glg* genes followed the same time-course expression pattern in WT cells, with the β -galactosidase activity increasing during the exponential growth phase, and reaching a plateau in the stationary phase. On the contrary, $\Delta_{up-glgB}$ cells displayed sharply reduced expression of *glgB*, *glgX* and *glgC* fusions ($\leq 15\%$ of that of WT cells at the onset of the stationary phase), and reduced but still very significant levels of expression of *glgA* and *glgP* (40-50% of those of WT cells) (**Figure 2A**, **Supplemental Figure 2**).

We also carried out western blot analyses of GlgC in Δ_{glgB}^* and Δ_{glgX}^* cells (Δ_{glgB} and Δ_{glgX} cells lacking the kanamycin resistance cassette, **Supplemental Figure 1**), and in $\Delta_{up-glgB}$ cells at the onset of the stationary phase. In agreement with the results obtained employing *glg::lacZY* fusions (cf. Fig. 2A), these analyses revealed that $\Delta_{up-glgB}$ cells accumulated drastically reduced levels of GlgC when compared with WT cells, whereas Δ_{glgB}^* and Δ_{glgX}^* cells accumulated nearly WT GlgC levels (**Figure 2B**). The small amount of GlgC accumulated by $\Delta_{up-glgB}$ cells (**Figure 2B**) is ascribed to the combination of both residual levels of *glgC* expression in these mutants, as judged by the analyses of β -galactosidase activity in *glgC::lacZY* expressing $\Delta_{up-glgB}$ cells (**Figure 2A**), and low GlgC proteolytic turnover, allowing this protein to accumulate within the cell (**Supplemental Figure 3**). Residual *glgC* expression occurring in $\Delta_{up-glgB}$ cells could be ascribed to low RNA polymerase affinity for nonspecific sites [33,34], rather than to active promoter sequences occurring within *glgX*. This conclusion is supported by comparative analyses of β -galactosidase activity in *glgC::lacZY* expressing $\Delta_{up-glgB}$ and Δ_{glgBX}^* cells (Δ_{glgBX} cells lacking the 270 bp sequence immediately upstream of *glgB*, **Supplemental Figure 1**), which revealed

marginally low and similar residual β -galactosidase activities in both strains (**Supplemental Figure 3**). The overall data thus provided strong indications that (a) the five *glgBXCAP* genes are organized in an operon mainly controlled by promoter sequence(s) located upstream of *glgB* and (b) *glgCAP* expression is only marginally controlled by sequences occurring within *glgX*.

Identification of a main promoter sequence in the intergenic *asd-glgB* region driving *E. coli glgBXCAP* expression

Computer searches for putative bacterial promoters in the *E. coli glg* locus identified -35 (TTACCG) and -10 (GGCTATTCT) sequences in the 273 bp *asd-glgB* intergenic region starting at 194 and 170 bp upstream of the *glgB* translation initiation codon, respectively (**Figure 3A**). The role of these sequences in directing expression of downstream *glgBXCAP* genes was analyzed by measuring the β -galactosidase activity levels of cells expressing *glgB::lacZY* fusions and containing different deletions of the *glgB* upstream region (**Figure 3B**). As seen in **Figure 3C**, cells containing deletions encompassing DNA regions located over 800 and 250 bp upstream of the *glgB* initiation codon (BI and BII) displayed similar β -galactosidase activities than cells bearing no deletions (B0). On the contrary, removal of the genomic region located between 250 and 150 bp upstream of *glgB* resulted in a sharp reduction in β -galactosidase levels (BIII), strongly indicating that this region contains the main promoter(s) directing expression of the *E. coli glgBXCAP* operon. To explore this possibility we identified by 5'RACE transcription start site(s) upstream of *glgB* (for details, see Experimental Procedures). Amplicons of ca. 350 bp were generated in TAP-treated (but not in non-treated) RNA samples, which were cloned and sequenced. Sequence analysis of different clones identified in all of them a single transcription start site located 155 bp upstream of the *glgB* translation initiation codon (**Figure 3A**), a result which fully agrees with both computer searches and deletion analyses for the location of *glgBXCAP* promoter sequences (see above).

Identification of an alternative suboperonic promoter region within *glgC* driving *glgA* and *glgP* expression

A salient feature of prokaryotic gene expression is the phenomenon of operon polarity (i.e., the reduced expression of promoter distal genes with respect to the proximal genes). This rule may sometimes be over-ridden at the suboperonic level by the presence of an internal promoter [35,36]. During the course of our investigations we realized that expression levels of *glgA* and *glgP* are higher than that of the upstream *glgC* gene (**Figure 2A**). This and the fact that the removal of the *glgB* upstream region only resulted in a 50-60% reduction of *glgA::lacZY* and *glgP::lacZY* fusions expression (cf. Fig. 2A) suggested that *glgA* and *glgP* may be under the control of an alternative suboperonic promoter located within *glgC*. To test this possibility we compared the expression of chromosomal *glgA::lacZY* and *glgP::lacZY* transcriptional fusions on $\Delta glgBX^*$ and $\Delta glgBXC^*$ cells ($\Delta glgBXC$ cells lacking the 270 bp sequence immediately upstream of *glgB*) (**Supplemental Figure 1**), and on WT cells. As illustrated in **Figure 4**, these analyses showed that, whereas *glgA* and *glgP* expression levels in $\Delta glgBX^*$ cells were still ca. 40-50% of that WT cells, *glgA* and *glgP* expression levels in $\Delta glgBXC^*$ cells further reduced to values $\leq 15\%$ of those of WT cells. The overall data thus strongly indicated that expression of *E. coli glgA* and *glgP* is not solely controlled by a promoter located upstream of *glgB*, but also by an alternative promoter located within *glgC*.

To identify promoter sequences located within *glgC* directing expression of *glgAP* a deletion approach similar to that used above (cf. Figure 3) was employed using *glgA::lacZY* fusions, except that the genomic region located between *asd* and part of the *glgC* gene in these cells was removed to eliminate the main *glgBXCAP* promoter regions and any other potential promoter sequences located upstream of the *glgC* gene. Putative -35 (TCGCAATT) and -10 (ACCTACAAT) consensus promoter sequences were identified starting at 394 and 369 bp upstream of the *glgA* translation initiation codon, respectively (Figure 5A). As seen in Figure 5B and 5C, similar levels of β -galactosidase activity were detected in cells expressing *glgA::lacZY* fusions under control of regions located over 450 bp upstream of the *glgA* initiation codon (deletions AI and AII). On the contrary, removal of the genomic region between 450 and 300 bp upstream of the *glgA* initiation (deletion AIII) resulted in a drastic reduction in β -galactosidase activity, strongly suggesting that suboperonic promoter sequences directing *glgAP* expression are located in this region. To explore this possibility we identified by 5'RACE the transcription start site upstream of *glgA* (for details, see Experimental Procedures). Amplicons of ca. 450 bp were generated in TAP-treated (but not in non-treated) RNA samples, which were cloned and sequenced. Sequence analyses of different clones identified in all of them a single transcription start site located 359 bp upstream of the *glgA* translation initiation codon (Figure 5A), a result which fully agrees with both computer searches and deletion analyses for the location of alternative suboperonic sequences driving *glgAP* expression (see above).

The *glgBXCAP* operon belongs to the RelA and PhoP-PhoQ regulons

During nutrient starvation *E. coli* elicits the so-called stringent response that switches the cell from a growth-related mode to a maintenance/survival/biosynthesis mode [37]. The hallmark of this pleiotropic response is the accumulation of the alarmone (p)ppGpp, which is synthesized by the *relA* product [38]. This nucleotide binds prokaryotic RNA polymerase to increase transcription of amino acid biosynthesis genes during nutritional starvation, and to down-regulate the production of ribosomes, nucleic acids and proteins [39]. Therefore, when ppGpp accumulates, growth is arrested and an important pool of ATP diverts towards glycogen biosynthesis under conditions of carbon source excess. Consistent with this view, experimental evidences provided by different authors have shown that (a) *relA* positively affects glycogen accumulation [13,14,40,41] and (b) *glgCAP* expression is under control of the RelA product, ppGpp [14,20,21]. Whether *glgBXCAP* operon is under *relA* control was investigated by comparing the expression levels of chromosomal *glg::lacZY* transcriptional fusions on WT cells and $\Delta relA$ cells cultured in either Kornberg or in MOPS media supplemented with glucose (Figure 6, Supplemental Figure 4). As shown in Figure 6A and 6B, these analyses clearly showed that the expression of the five *glg* genes was largely abolished in $\Delta relA$ cells. The overall data thus not only reinforce the notion that *glgBXCAP* compose a single transcriptional unit, but also indicate that it forms part of the *E. coli* RelA regulon.

PhoP-PhoQ is a two-component regulatory system occurring in *E. coli* and *Salmonella spp.* that monitors the availability of extracellular Mg^{2+} , and transcriptionally controls the expression of many genes [42,43]. We have recently shown that PhoP-PhoQ highly regulates glycogen metabolism at submillimolar environmental Mg^{2+} concentrations since (a) cells impaired in the PhoP and PhoQ functions accumulate low glycogen [13,14], and (b) expression of *glgC::lacZY* was positively regulated by the PhoP-PhoQ regulatory system at low environmental Mg^{2+} [14]. Whether *glgBXCAP* operon is under PhoP-PhoQ control was investigated by comparing the expression levels of chromosomal *glg::lacZY* transcriptional fusions on

WT and $\Delta phoP$ cells cultured in liquid Kornberg medium supplemented with glucose under limiting environmental Mg^{2+} concentrations. As shown in **Figure 6C**, these analyses clearly showed that the expression of the five *glg* genes was lower in $\Delta phoP$ cells than in WT cells, the overall data further reinforcing the idea that the five *glg* genes belong to a single *glgBXCAP* operon that is under positive control of the PhoP-PhoQ regulatory system under limiting environmental Mg^{2+} concentrations.

The suboperonic promoter driving *glgAP* expression belongs to the RelA and PhoP-PhoQ regulons

Whether the alternative suboperonic promoter region driving *glgAP* expression is also under control of the *relA* and *phoP/phoQ* gene products was investigated by comparing the expression levels of chromosomal *glgA::lacZY* transcriptional fusion on $\Delta glgBX^*$ and $\Delta glgBX^*\Delta relA$ cells cultured in either Kornberg or in MOPS media supplemented with glucose. Furthermore, we also compared the expression levels of chromosomal *glgA::lacZY* transcriptional fusion on $\Delta glgBX^*$ and $\Delta glgBX^*\Delta phoP$ cells cultured in liquid Kornberg medium supplemented with glucose under limiting environmental Mg^{2+} concentrations. As shown in **Figure 7**, these analyses clearly showed that *glgA* expression was lower in $\Delta glgBX^*\Delta relA$ and $\Delta glgBX^*\Delta phoP$ cells than in $\Delta glgBX^*$ cells, the overall data showing that the suboperonic promoter driving *glgAP* expression belongs to both the RelA and PhoP-PhoQ regulons.

Expression of *E. coli glgBXCAP* genes is not altered in a Δcya background

cAMP produced by the product of the *cya* gene is an important cellular mediator of glucose effects in *E. coli*. The cAMP/CRP complex is a positive transcriptional regulator of a number of catabolic operons, and as such, plays a role in catabolite repression, whereby secondary carbon sources are not catabolized in the presence of glucose. Both cAMP and CRP are required for expression of *glgS* and PTS-related genes required for normal glycogen production [22,44]. On the basis of *in vitro* and *in vivo* experimental data, Preiss and co-workers have proposed that expression of *E. coli glgCAP* genes is positively regulated at the transcriptional level by the cAMP/CRP complex [19,21,45] (reviewed in [11]). However, a role of cAMP in regulating *E. coli glgCAP* expression has been questioned by other authors on the basis of results obtained using cells bearing chromosomally-located *glgA::lacZ* gene fusions [25,44]. Moreover, recent transcriptome analysis conducted in different *E. coli* strains also failed to include *glg* genes into the CRP regulon [22,23]. In addition, our own analysis using *glgC::lacZY* fusions also indicated that *glgC* expression was not significantly affected in an *E. coli* BW25113 Δcya mutant background lacking the ability to produce cAMP [14]. To investigate whether *glgBXCAP* may be part of the cAMP regulon we compared the expression levels of chromosomal *glg::lacZY* transcriptional fusions on WT cells and Δcya cells cultured in liquid Kornberg medium supplemented with 5 mM glucose. Furthermore, to investigate whether the alternative suboperonic promoter driving *glgAP* expression is regulated by cAMP we compared the expression levels of chromosomal *glgA::lacZY* transcriptional fusion on $\Delta glgBX^*$ cells and $\Delta glgBX^*\Delta cya$ cells cultured in liquid Kornberg medium supplemented with 5 mM glucose. As shown in **Supplemental Figure 5A** and **5B**, these analyses revealed that the expression of *glgBXCAP* on WT cells, and of *glgA* on $\Delta glgBX^*$ cells was not affected by the lack of the *cya* function.

We also compared the *glgB::lacZY* expression levels on WT cells cultured in glucose and in other carbon sources such as maltose, fructose and galactose, conditions that are known to induce cAMP synthesis. As shown in **Supplemental Figure 5C**, these analyses revealed that *glgB* expression on WT cells was not affected by the nature of the

carbon source, the overall data reinforcing the idea that cAMP does not control *E. coli* glycogen metabolism at the level of transcription of *glg* genes.

DISCUSSION

In this work we have shown that the five *E. coli glgBXCAP* genes are transcribed in a single transcriptional unit under control of promoter sequences located upstream of *glgB*. Moreover, we have also shown the occurrence of an alternative promoter located within *glgC* that controls the expression of downstream *glgA* and *glgP* genes. Finally, we have provided evidence that *glgBXCAP* operon forms part of both the RelA and PhoP-PhoQ regulons. Previous works by other authors using S1 nuclease protection assays indicated the existence in *E. coli* of *glgCAP* transcripts initiating within the *glgX* coding region [19,46]. However, computer searches failed to identify putative consensus bacterial promoter sequences within *glgX* [19, also confirmed in this work]. Moreover, 5'RACE analyses using *glgC* specific oligonucleotides failed in our hands to detect any putative transcription start site within *glgX* (not shown). Finally, comparative analyses of β -galactosidase activity in *glgC::lacZY* expressing $\Delta up-glgB$ and $\Delta glgBX^*$ cells revealed marginally low and similar residual activities in both strains (**Supplemental Figure 3**). It is thus likely that the results obtained by other authors [19,46] could be ascribed to low basal levels of non-specific transcription initiation occurring in regions immediately upstream of *glgC*, which is a situation reported for many *E. coli* genes [34]. Another possible explanation is the nucleolytic processing of longer transcripts. In this context, it is worth noting that longer transcripts attributed to read-through transcription from the upstream *glgBX* genes were also detected in S1 protection assay assays [46].

The existence of *glg* genes clustered in a single chromosomal locus forming part of a single operon, and showing suboperonic promoters has been previously reported in α -proteobacterial species such as *Agrobacterium tumefaciens*, *Mesorhizobium loti*, and *Rhizobium tropici* [4,5,18]. In some of these cases, however, the order of *glg* genes and the position of the suboperonic promoters in the corresponding operons are different to those of *E. coli* reported here, possibly responding to different metabolic needs. Internal promoters in large bacterial operons reinforce expression of downstream genes, which would otherwise be transcribed in lower amounts owing to the effects of operon polarity. Therefore, the suboperonic promoter within *glgC* may have evolved to guarantee adequate *glgAP* expression levels. Alternatively, internal promoters may also temporally allow the expression of certain genes of the operon under reduced expression of other members of the same operon whose functions are redundantly fulfilled by other loci [35,36]. The relevance of *glgA* expression when *glgC* is lowly expressed can be well exemplified by the occurrence of *glgA* expressing *glgC* mutants accumulating substantial glycogen content [16] due to the occurrence of important sources, other than GlgC, of ADPG [15]. Since glycogen plays relevant roles in the survival of *E. coli* and *S. enterica* to sporadic periods of famine, and/or in colonization and virulence, and because glycogen metabolism is highly interconnected with multiple and important cellular processes [13,14], it is tempting to speculate that redundancy of ADPG sources and regulatory mechanism(s) guaranteeing *glgA* expression have been selected during evolution to warrant glycogen production in these particular species.

Regulation of *glg* gene expression involves a still not well defined assemblage of factors that are adjusted to the physiological and nutritional status of the cell. Results presented in this work showing that the five *E. coli glgBXCAP* genes can be transcribed in a single mRNA unit open new clues on the mechanisms regulating glycogen metabolism in enterobacterial species. Depending on both external and internal

conditions, complex operons may be regulated by the presence and/or combination of various factors including differential transcriptional or translational efficiency, differential localized instability in the polycistronic mRNA, presence of internal promoter(s), regulatory structures at the 5' untranslated regions (UTR) of the transcripts, etc. [34,47,48]. Notably, the positions of the two promoter regions identified in this work regulating expression of *glg* genes with respect to that of the initiation codons (cf. Fig. 3A and Fig. 5A) point to the occurrence of relatively long UTRs on each of the corresponding transcripts. 5' UTRs serve as targets of regulatory factors including RNA binding proteins, RNase E, and regulatory small RNAs, and can form riboswitch structures interacting with low molecular weight effectors. All these elements can control the fate and stability of transcripts [47]. It is therefore tempting to speculate that these regions may play important roles in regulating translation capability and/or stability of *glgBXCAP* and *glgAP* transcripts. In this context, we have recently shown that glycogen accumulation in *E. coli* is severely compromised in a Δhfq genetic background [14]. Hfq is a chaperone that stabilizes many regulatory sRNAs, and facilitates the base pairing between sRNAs and their target mRNAs thus affecting both transcript translation and degradation [47,49]. Thus, it is highly conceivable that the glycogen-deficient phenotype of Δhfq cells may result from small RNAs- and/or Hfq-mediated regulation of *glgBXCAP* expression. In fact, recent analyses of mRNA targets of Hfq in *Salmonella* have shown that *glgB*, *glgX*, *glgC* and *glgA* transcript levels are similarly reduced in a Δhfq background [50]. The global regulator CsrA is a RNA binding protein that controls the expression of genes involved in carbohydrate metabolism. *glgB::lacZ* and *glgC::lacZ* translational fusion expression analyses revealed that CsrA exerts a negative effect on glycogen accumulation through its negative control on *glgB* and *glgC* expression [51]. Although more recent studies focused their attention on the effect of CsrA in *glgC* expression, and showed that CsrA binds sequences immediately upstream of *glgC* to promote mRNA decay and prevent *glgC* translation [52], further studies will be necessary to elucidate whether CsrA can also bind the 5' UTR of *glgBXCAP* transcripts to prevent glycogen biosynthesis by regulating their nucleolytic processing. Needless to say, further investigations are necessary to understand the many regulatory aspects of bacterial glycogen metabolism at both transcriptional and post-transcriptional levels.

Accepted Article

AUTHORS CONTRIBUTION

Manuel Montero, Goizeder Almagro, Gustavo Eydallin, Francisco J. Muñoz, Abdellatif Bahaji, Jun Li, Mehdi Rahimpour, and Edurne Baroja-Fernandez were responsible for performing experiments. Manuel Montero, Goizeder Almagro, Alejandro M. Viale and Javier Pozueta-Romero were responsible for planning experiments and analyzing data. Manuel Montero, Goizeder Almagro, Alejandro M. Viale and Javier Pozueta-Romero were responsible for writing the manuscript.

ACKNOWLEDGEMENTS

This research was partially supported by the grant BIO2007-63915 from the Comisión Interministerial de Ciencia y Tecnología and Fondo Europeo de Desarrollo Regional (Spain) and by Iden Biotechnology S.L.. M.M. acknowledges a post-doctoral contract from I3P program of Consejo Superior de Investigaciones Científicas. M. R. acknowledges a pre-doctoral JAE fellowship from the Consejo Superior de Investigaciones Científicas. A. M. V. expresses his gratitude to the Ministerio de Educación y Cultura, the Consejo Superior de Investigaciones Científicas and the Public University of Nafarroa for their support. G. A. acknowledges a pre-doctoral fellowship from the Public University of Nafarroa. We thank Cristina Cirauqui, Maite Sesma and Maite Hidalgo (Institute of Agrobiotechnology) for technical support.

Accepted Manuscript

REFERENCES

- 1 Lillie, S. H. and Pringle, J. R. (1980) Reserve carbohydrate metabolism in *Saccharomyces cerevisiae*: responses to nutrient limitation. *J. Bacteriol.* **143**, 1384-1394
- 2 Wilson W. A, Wang Z. and Roach P. J. (2002a) Systematic identification of the genes affecting glycogen storage in the yeast *Saccharomyces cerevisiae*: implication of the vacuole as a determinant of glycogen level. *Mol. Cell Proteomics* **1**, 232-242
- 3 Ferrer, J. C., Favre, C., Gomis, R. R., Fernández-Novell, J. M., García-Rocha, M., de la Iglesia, N., Cid, E. and Guinovart, J. J. (2003) Control of glycogen deposition. *FEBS Lett.* **546**, 127-132
- 4 Marroquí, S., Zorreguieta, A., Santamaría, C., Temprano, F., Soberón, M., Megías, M. and Downie, J. A. (2001) Enhanced symbiotic performance by *Rhizobium tropici* glycogen synthase mutants. *J. Bacteriol.* **183**, 854-864
- 5 Lepek, V. C., D'Antuono, A. L., Tomatis, P. E., Ugalde, J. E., Giambiagi, S. and Ugalde, R. A. (2002) Analysis of *Mesorhizobium loti* glycogen operon: effect of phosphoglucomutase (*pgm*) and glycogen synthase (*glgA*) null mutants on nodulation of *Lotus tenuis*. *Mol. Plant Microb. Interact.* **15**, 368-375
- 6 Chang, D. E., Smalley, D. J., Tucker, D. L., Leatham, M. P., Norris, W. E., Stevenson, S. J., Anderson, A. B., Grissom, J. E., Laux, D. C., Cohen, P. S. and Conway, T. (2004) Carbon nutrition of *Escherichia coli* in the mouse intestine. *Proc. Natl. Acad. Sci. USA* **101**, 7427-7432
- 7 Jones, S. A., Jorgensen, M., Chowdhury, F. Z., Rodgers, R., Hartline, J., Leatham, M. P., Struve, C., Krogfelt, K. A., Cohen, P. S. and Conway, T. (2008) Glycogen and maltose utilization by *Escherichia coli* O157:H7 in the mouse intestine. *Infect. Immun.* **76**, 2531-2540
- 8 Sambou T., Dinadayala P., Stadthagen G., Barilone N., Bordat Y., Constant P., Levillain, F., Neyrolles, O., Gicquel, B., Lemassu, A. and Jackson, M. (2008) Capsular glucan and intracellular glycogen of *Mycobacterium tuberculosis*: biosynthesis and impact on the persistence in mice. *Mol. Microbiol.* **70**, 762-774
- 9 Bourassa, L. and Camilli, A. (2009) Glycogen contributes to the environmental persistence and transmission of *Vibrio cholerae*. *Mol. Microbiol.* **72**, 124-138
- 10 Dietzler, D. N., Leckie, M. P. and Lais, C. J. (1973) Rates of glycogen synthesis and the cellular levels of ATP and FDP during exponential growth and the nitrogen-limited stationary phase of *Escherichia coli* W4957(K). *Arch. Biochem. Biophys.* **156**, 684-693
- 11 Preiss J. (2009) Glycogen: biosynthesis and regulation. In, *EcoSal-Escherichia coli and Salmonella: Cellular and Molecular Biology* (Böck, A., Curtiss III, R., Kaper, J. B., Karp, P. D., Neidhardt, F. C., Nyström, T., Slauch, J. M., Squires, C. L. and Ussery D., eds.) <http://www.ecosal.org>. ASM Press, Washington, DC. doi:10.1128/ecosal.4.7.4
- 12 Ballicora, M. A., Iglesias, A. A. and Preiss, J. (2003) ADP-Glucose pyrophosphorylase, a regulatory enzyme for bacterial glycogen synthesis. *Microbiol. Mol. Biol. Rev.* **67**, 213-225
- 13 Eydallin, G., Viale, A. M., Moran-Zorzano, M. T., Munoz, F. J., Montero, M., Baroja-Fernandez, E. and Pozueta-Romero, J. (2007) Genome-wide screening of genes affecting glycogen metabolism in *Escherichia coli* K-12. *FEBS Lett.* **581**, 2947-2953
- 14 Montero, M., Eydallin, G., Almagro, G., Muñoz, F. J., Viale, A. M., Rahimpour, M., Sesma, M. T., Baroja-Fernández, E. and Pozueta-Romero, J. (2009) *Escherichia coli* glycogen metabolism is controlled by the PhoP-PhoQ regulatory system at

submillimolar environmental Mg^{2+} concentrations, and is highly interconnected with a wide variety of cellular processes. *Biochem. J.* **424**, 129-241

15 Morán-Zorzano, M. T., Alonso-Casajús, N., Muñoz, F. J., Viale, A. M., Baroja-Fernández, Eydallin G. and Pozueta-Romero, J. (2007) Occurrence of more than one important source of ADPglucose linked to glycogen biosynthesis in *Escherichia coli* and *Salmonella enterica*. *FEBS Lett.* **581**, 4423-4429

16 Eydallin, G., Morán-Zorzano, M. T., Muñoz, F. J., Baroja-Fernández, E., Montero, M., Alonso-Casajús, N., Viale, A. M. and Pozueta-Romero, J. (2007) An *Escherichia coli* mutant producing a truncated inactive form of GlgC synthesizes glycogen: further evidences for the occurrence of various important sources of ADPglucose in enterobacteria. *FEBS Lett.* **581**, 4417-4422

17 Kiel, J. A. K. W., Boels, J. M., Geldman, G. and Venema, G. (1994) Glycogen in *Bacillus subtilis*: molecular characterization of an operon encoding enzymes involved in glycogen biosynthesis and degradation. *Mol. Microbiol.* **11**, 203-218

18 Ugalde, J. E., Lepek, V., Uttaro, A., Estrella, J., Iglesias, A. and Ugalde, R. A. (1988) Gene organization and transcription analysis of the *Agrobacterium tumefaciens* glycogen (*glg*) operon: two transcripts for the single phosphoglucomutase gene. *J. Bacteriol.* **180**, 6557-6564

19 Romeo, T. and Preiss, J. (1989) Genetic regulation of glycogen biosynthesis in *Escherichia coli*: *In vitro* effects of cyclic AMP and guanosine 5'-diphosphate 3'-diphosphate and analysis of *in vivo* transcripts. *J. Bacteriol.* **171**, 2773-2782

20 Traxler, M. F., Summers, S. M., Nguyen, H-T., Zacharia, V. M., Hightower, G. A., Smith, J. T. and Conway, T. (2008) The global, ppGpp-mediated stringent response to amino acid starvation in *Escherichia coli*. *Mol. Microbiol.* **68**, 1128-1148

21 Romeo, T., Black, J. and Preiss, J. (1990) Genetic regulation of glycogen biosynthesis in *Escherichia coli*: *in vivo* effects of the catabolite repression and stringent response systems in *glg* gene expression. *Current Microbiol.* **21**, 131-137

22 Gosset, G., Zhang, Z., Nayyar, S., Cuevas, W. A. and Saier, M. H. (2004) Transcriptome analysis of Crp-dependent catabolite control of gene expression in *Escherichia coli*. *J. Bacteriol.* **186**, 3516-3524

23 Zheng, D., Constantinidou, C., Hobman, J. L. and Minchin, S. D. (2004) Identification of the CRP regulon using *in vitro* and *in vivo* transcriptional profiling. *Nucleic Acids Res.* **32**, 5874-5893

24 Hollands, K., Busby, S. J. W. and Lloyd, G. S. (2007) New targets for the cyclic AMP receptor protein in the *Escherichia coli* K-12 genome. *FEMS Microbiol. Lett.* **274**, 89-94

25 Weichart D., Lange R., Henneberg N. and Hengge-Aronis R. (1993) Identification and characterization of stationary phase-inducible genes in *Escherichia coli*. *Mol. Microbiol.* **10**,407-420

26 Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L. and Mori, H. (2006) Construction of *Escherichia coli* K-12 in frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* doi:10.1038/msb4100050

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

28 Datsenko, K. A. and Wanner, B. L. (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. USA* **97**, 6640-6645

- 29 Cherepanov, P. P. and Wackernagel, W. (1995) Gene disruption in *Escherichia coli*: Tc^R and Km^R cassettes with the option of Flp-catalyzed excision of the antibiotic-resistance determinant. *Gene* **158**, 9-14
- 30 Ellermeier, C. D., Janakiraman, A. and Slauch, J. M. (2002) Construction of targeted single copy *lac* fusions using lambda Red and FLP-mediated site-specific recombination in bacteria. *Gene* **290**, 153-161
- 31 Miller, J. H. (1992) A short course in bacterial genetics: A laboratory manual and handbook for *Escherichia coli* and related bacteria. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.
- 32 Miller, J. H. (1972) Experiments in molecular genetics. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.
- 33 deHaseh, P. L., Lohman, T. M., Burgess, R. R. and Record, M. T. (1978) Nonspecific interactions of *Escherichia coli* RNA polymerase with native and denatured DNA: differences in the binding behavior of core and holoenzyme. *Biochemistry* **17**, 1612-1622.
34. Mendoza-Vargas, A., Olvera, L., Olvera, M., Grande, R., Vega-Alvarado, L., Taboada, B., Jimenez-Jacinto, V., Salgado, H., Juárez, K., Contreras-Moreira, B., Huerta, A. M., Collado-Vides, J. and Morett, E. (2009) Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in *E. coli*. *PLoS One*. **4**(10):e7526
- 35 Adhya, S. (2003) Suboperonic regulatory signals. *Sci. STKE* **185**, pe22
- 36 Lee, H. J., Jeon, H. J., Ji, S. C., Yun, S. H. and Lim, H. M. (2008) Establishment of an mRNA gradient depends on the promoter: an investigation of polarity in gene expression. *J. Mol. Biol.* **378**, 318-327
- 37 Chatterji, D. and Ojha, A. K. (2001) Revisiting the stringent response, ppGpp and starvation signalling. *Curr. Opin. Microbiol.* **4**, 160-165
- 38 Cashel, M., Gentry, D. R., Hernández, V. J. and Vinella, D. (1996) The Stringent Response, In: *Escherichia coli* and *Salmonella* cellular and molecular biology, 2nd Ed (Neidhart, F. C., Curtis, III., R., Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., Schaechter, M. and Umberger, H. E., eds.), pp. 1458-1496, ASM Press, Washington, DC
- 39 Magnusson, L. U., Farewell, A. and Nyström, T. (2005) ppGpp: a global regulator in *Escherichia coli*. *Trends Microbiol.* **13**, 236-242
- 40 Bridger, W. A. and Paranchych, W. (1978) *relA* gene control of bacterial glycogen synthesis. *Can. J. Biochem.* **56**, 403-406
- 41 Taguchi, M., Izui, K. and Katsuki, H. (1980) Augmentation of glycogen synthesis under stringent control in *Escherichia coli*. *J. Biochem. (Tokyo)* **88**, 379-387
- 42 Garcia-Vescovi, E., Soncini, F. C. and Groisman, E. A. (1996) Mg²⁺ as an extracellular signal: environmental regulation of *Salmonella* virulence. *Cell* **84**, 165-174
- 43 Spinelli S. V., Pontel L. B., García Vescovi E. and Soncini F. C. (2008) Regulation of magnesium homeostasis in *Salmonella*: Mg²⁺ targets the *mgtA* transcript for degradation by RNase E. *FEMS Microbiol. Lett.* **280**, 226-234
- 44 Hengge-Aronis, R. and Fischer, D. (1992) Identification and molecular analysis of *glgS*, a novel growth-phase-regulated and *rpoS*-dependent gene involved in glycogen synthesis in *Escherichia coli*. *Mol. Microbiol.* **6**, 1877-1886
- 45 Urbanowski, J., Leung, P., Weissbach, H. and Preiss, J. (1983) The *in vitro* expression of the gene for *Escherichia coli* ADPglucose pyrophosphorylase is stimulated by cyclic AMP and cyclic AMP receptor protein. *J. Biol. Chem.* **258**, 2782-2784

- 46 Liu M. Y., Yang H. and Romeo T. (1995) The product of the pleiotropic *Escherichia coli* gene *csrA* modulates glycogen biosynthesis via effects on mRNA stability. *J. Bacteriol.* **177**, 2663-2672
- 47 Kaberdin V. R. and Bläsi U. (2006) Translation initiation and the fate of bacterial mRNAs. *FEMS Microbiol. Rev.* **30**, 967-979
- 48 Roberts J. W., Shankar S. and Filter J. J. (2008) RNA polymerase elongation factors. *Annu. Rev. Microbiol.* **62**, 211-233
- 49 Brennan, R. G. and Link, T. M. (2007) Hfq structure, function and ligand binding. *Curr. Opin. Microbiol.* **10**, 125-133
- 50 Sittka, A., Lucchini, S., Papenfort, K., Sharma, C. M., Rolle, K., Binnewies, T. T., Hinton, J. C. D. and Vogel, J. (2008) Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PloS Genet.* **4**, e1000163.
- 51 Romeo, T., Gong, M., Liu, M.Y., Brun-Zinkernagel, A-M. (1993) Identification and molecular characterization of *csrA*, a pleiotropic gene from *Escherichia coli* that affects glycogen biosynthesis, gluconeogenesis, cell size, and surface properties. *J. Bacteriol.* **175**, 4744-4755.
52. Baker, C.S., Morozov, I., Suzuki, K., Romeo, T. and Babitzke, P. (2002) CsrA regulates glycogen biosynthesis by preventing translation of *glgC* in *Escherichia coli*. *Mol. Microbiol.* **44**, 1599-1610.
53. Feng, Y. and Cronan, J.E. (2010) Overlapping repressor binding sites result in additive regulation of *Escherichia coli* FadH by FadR and ArcA. *J. Bacteriol.* **192**, 4289-4299.

FIGURE LEGENDS

Figure 1: Occurrence of *glgBXCAP* transcripts in *E. coli*. (A) Physical arrangement of *glgB*, *glgX*, *glgC*, *glgA* and *glgP* genes in the *E. coli* genome. The number of bp separating each of the corresponding genes is indicated in the upper part of the scheme. A negative number implies gene overlapping. (B) Schematic illustration of the localization of the O1 primer necessary to reverse transcribe *glgBXCAP* transcripts, and of primers necessary to amplify by PCR fragments of the resulting cDNAs (O2 and O3 for PCRI; O4 and O5 for PCRII; O6 and O7 for PCRIII, and O8 and O9 for PCRIV. (C) PCR amplification analyses of the cDNAs produced using O1 as primer. Lane 1: size markers (hyperladder I, Bioline); lane 2: PCR using RNA extracts; lane 3: PCR on cDNAs obtained using O1 in reverse transcription reaction.

Figure 2: *glgBXCAP* expression is under control of promoter sequence(s) occurring upstream of the *glgB* gene. (A) Expression of *glg::lacZY* fusions in WT cells (white bars), and in cells lacking 253 bp of the *asd-glgB* intergenic region located upstream of *glgB* (nucleotide position -270 to -17 from the *glgB* translation initiation codon) (grey bars). (B) Western blot of GlgC in WT cells, Δ *up-glgB* cells, Δ *glgB** cells and Δ *glgX** cells. The gel was loaded with 60 μ g per lane of total soluble proteins. In “A” and “B”, cells were cultured in liquid Kornberg-glucose medium, and harvested at the onset of the stationary phase. The results are the mean \pm SE of 5 independent experiments.

Figure 3: Identification of a main promoter sequence upstream of *glgB*. (A) Nucleotide sequence of the 273 bp *asd-glgB* intergenic region upstream of *glgB*. Predicted -35 (TTACCG) and -10 (GGCTATTCT) sequences starting at 194 and 170 bp, respectively, upstream of the *glgB* translation initiation codon (in bold) are indicated in grey boxes. The transcription initiation start localized at position 155 bp upstream of the *glgB* translation initiation codon is indicated in bold and with an asterisk. Arrows indicate the end of BII and BIII deletions (250 bp and 150 bp upstream of the *glgB* translation initiation codon, respectively). The AGGA Shine-Dalgarno (SD) sequence upstream of *glgB* is boxed. The *asd* TAA termination codon is indicated in bold italics. (B) Schematic representation of the different deletions used to map the promoter regions located at the *glgB* upstream region in *glgB::lacZY* fusions. (C) β -galactosidase activity (in Miller units) of cells expressing *glgB::lacZY* fusions under control of B0, BI, BII and BIII deletions. Cells were cultured in liquid Kornberg-glucose medium, and harvested at the onset of the stationary phase to determine β -galactosidase activity levels. For details see Experimental Procedures.

Figure 4: Expression of *E. coli glgA* and *glgP* is controlled by promoter sequences occurring upstream of *glgB*, and by an alternative suboperonic promoter located within *glgC*. The graphic represents β -galactosidase activity of WT cells (white bars), Δ *glgBX** and Δ *glgBXC** cells (light grey and dark grey bars, respectively) expressing *glgA::lacZY* and *glgP::lacZY* fusions. Cells were cultured in liquid Kornberg-glucose medium, and harvested at the onset of the stationary phase for β -galactosidase activity measurements. For details see Experimental Procedures.

Figure 5: Identification of an alternative promoter within *glgC* driving *glgA* and *glgP* expression in *E. coli*. (A) Nucleotide sequence of the 695 bp *glgA* upstream region occurring within *glgC*. Putative -35 (TCGCAATT) and -10 (ACCTACAAT) sequences starting at 394 and 369 bp, respectively, upstream of the *glgA* translation initiation codon (+1, in bold) are indicated in grey boxes. The transcription initiation start

localized at position 359 bp upstream of the *glgA* translation initiation codon is indicated in bold and with an asterisk. Arrows indicate the end of AI, AII and AIII deletions (600, 450 and 300 bp, respectively, upstream of the *glgA* translation initiation codon). The AGGA Shine-Dalgarno (SD) sequence upstream of *glgA* is boxed. (B) Schematic representation of the different deletions used to map the promoter regions located at the *glgA* upstream region in *glgA::lacZY* fusions. (C) β -galactosidase activity (Miller units) of cells expressing *glgA::lacZY* fusions under control of AI, AII and AIII deletions. Cells were cultured in liquid Kornberg-glucose medium, and harvested at the onset of the stationary phase.

Figure 6: The *glgBXCAP* operon forms part of the RelA and PhoP-PhoQ regulons. Expression levels of *glgB::lacZY*, *glgX::lacZY*, *glgC::lacZY*, *glgA::lacZY* and *glgP::lacZY* in WT cells (white bars) and cells lacking the *relA* (A and B) and *phoP* (C) genes (grey bars). In “A”, cells were cultured in liquid Kornberg-glucose medium (see growth curve in **Supplemental Figure 4**). In “B”, cells were cultured in MOPS-glucose medium. In “C”, cells were cultured in Kornberg medium lacking MgCl₂ supplementation. In all cases, cells were harvested at the onset of the stationary phase. The results are the mean \pm SE of 5 independent experiments. For details see Experimental Procedures.

Figure 7: The suboperonic promoter driving *glgAP* expression forms part of the RelA and PhoP-PhoQ regulons. In “A” and “B”, expression levels of *glgA::lacZY* in Δ *glgBX** cells and Δ *glgBX***relA* cells cultured in liquid Kornberg-glucose medium and MOPS-glucose medium, respectively. In “C”, expression levels of *glgA::lacZY* in Δ *glgBX** cells and Δ *glgBX***phoP* cells cultured in Kornberg medium lacking MgCl₂ supplementation. In all cases, cells were harvested at the onset of the stationary phase. For details see Experimental Procedures.

Table 1: *E. coli* strains and plasmids used in this work

Designation	Description	Source
Bacteria		
BW25113	<i>lacI^r rrmB_{T14} ΔlacZ_{WJ16} hsdR514 ΔaraBAD_{AH33} ΔrhaBAD_{LD78}</i>	Keio collection [26]
<i>Δup-glgB</i>	<i>glgB</i> upstream region (-17 to -270 bp) replaced by a Sm ^R cassette in reversed orientation	This work
<i>ΔglgB</i>	BW25113 <i>ΔglgB::Km^R</i>	Keio collection [26]
<i>ΔglgB*</i>	BW25113 <i>ΔglgB::Km^R</i> where Km ^R cassette was removed using FRT sites	This work
<i>ΔglgX</i>	BW25113 <i>ΔglgX::Km^R</i>	Keio collection [26]
<i>ΔglgX*</i>	BW25113 <i>ΔglgX::Km^R</i> where Km ^R cassette was removed using FRT sites	This work
<i>ΔrpoS</i>	BW25113 <i>ΔrpoS::Km^R</i>	Keio collection [26]
<i>ΔrelA</i>	BW25113 <i>ΔrelA::Km^R</i>	Keio collection [26]
<i>ΔrelA*</i>	BW25113 <i>ΔrelA::Km^R</i> where Km ^R cassette was removed using FRT sites	This work
<i>ΔphoP</i>	BW25113 <i>ΔphoP::Km^R</i>	Keio collection [26]
<i>ΔphoP*</i>	BW25113 <i>ΔphoP::Km^R</i> where Km ^R cassette was removed using FRT sites	This work
<i>glgB::lacZY</i> (B0)	BW25113 <i>glgB::lacZY</i> transcriptional fusion	This work
<i>glgX::lacZY</i>	BW25113 <i>glgX::lacZY</i> transcriptional fusion	This work
<i>glgC::lacZY</i>	BW25113 <i>glgC::lacZY</i> transcriptional fusion	This work
<i>glgA::lacZY</i>	BW25113 <i>glgA::lacZY</i> transcriptional fusion	This work
<i>glgP::lacZY</i>	BW25113 <i>glgP::lacZY</i> transcriptional fusion	This work
<i>Δup-glgB glgB::lacZY</i>	BW25113 <i>glgB::lacZY</i> transcriptional fusion in <i>Δup-glgB</i>	This work
<i>Δup-glgB glgX::lacZY</i>	BW25113 <i>glgX::lacZY</i> transcriptional fusion in <i>Δup-glgB</i>	This work
<i>Δup-glgB glgC::lacZY</i>	BW25113 <i>glgC::lacZY</i> transcriptional fusion in <i>Δup-glgB</i>	This work
<i>Δup-glgB glgA::lacZY</i>	BW25113 <i>glgA::lacZY</i> transcriptional fusion in <i>Δup-glgB</i>	This work
<i>Δup-glgB glgP::lacZY</i>	BW25113 <i>glgP::lacZY</i> transcriptional fusion in <i>Δup-glgB</i>	This work
<i>ΔrelA glgB::lacZY</i>	BW25113 <i>glgB::lacZY</i> transcriptional fusion in <i>ΔrelA*</i>	This work
<i>ΔrelA glgX::lacZY</i>	BW25113 <i>glgX::lacZY</i> transcriptional fusion in <i>ΔrelA*</i>	This work
<i>ΔrelA glgC::lacZY</i>	BW25113 <i>glgC::lacZY</i> transcriptional fusion in <i>ΔrelA*</i>	This work

<i>ΔrelA glgA::lacZY</i>	BW25113 <i>glgA::lacZY</i> transcriptional fusion in <i>ΔrelA</i> *	This work
<i>ΔrelA glgP::lacZY</i>	BW25113 <i>glgP::lacZY</i> transcriptional fusion in <i>ΔrelA</i> *	This work
<i>ΔphoP glgB::lacZY</i>	BW25113 <i>glgB::lacZY</i> transcriptional fusion in <i>ΔphoP</i> *	This work
<i>ΔphoP glgX::lacZY</i>	BW25113 <i>glgX::lacZY</i> transcriptional fusion in <i>ΔphoP</i> *	This work
<i>ΔphoP glgC::lacZY</i>	BW25113 <i>glgC::lacZY</i> transcriptional fusion in <i>ΔphoP</i> *	This work
<i>ΔphoP glgA::lacZY</i>	BW25113 <i>glgA::lacZY</i> transcriptional fusion in <i>ΔphoP</i> *	This work
<i>ΔphoP glgP::lacZY</i>	BW25113 <i>glgP::lacZY</i> transcriptional fusion in <i>ΔphoP</i> *	This work
<i>ΔglgBX</i> *	BW25113 <i>ΔglgBX::Sp^R</i> lacking the 270 bp <i>glgB</i> upstream region	This work
<i>ΔglgBX* glgA::lacZY</i>	BW25113 <i>glgA::lacZY</i> transcriptional fusion in <i>ΔglgBX</i> *	This work
<i>ΔglgBX* glgC::lacZY</i>	BW25113 <i>glgC::lacZY</i> transcriptional fusion in <i>ΔglgBX</i> *	This work
<i>ΔglgBX* glgP::lacZY</i>	BW25113 <i>glgP::lacZY</i> transcriptional fusion in <i>ΔglgBX</i> *	This work
<i>ΔglgBXC</i> *	BW25113 <i>ΔglgBXC::Sp^R</i> lacking the 270 bp <i>glgB</i> upstream region	This work
<i>ΔglgBXC* glgA::lacZY</i>	BW25113 <i>glgA::lacZY</i> transcriptional fusion in <i>ΔglgBXC</i> *	This work
<i>ΔglgBXC* glgP::lacZY</i>	BW25113 <i>glgP::lacZY</i> transcriptional fusion in <i>ΔglgBXC</i> *	This work
<i>Δup-glgB</i> -800 (BI)	<i>glgB</i> upstream region (from -800 to -1327 bp) replaced by a <i>Sp^R</i> cassette in reversed orientation	This work
<i>Δup-glgB</i> -250 (BII)	<i>glgB</i> upstream region (from -250 to -1327 bp) replaced by a <i>Sp^R</i> cassette in reversed orientation	This work
<i>Δup-glgB</i> -150 (BIII)	<i>glgB</i> upstream region (from -150 to -1327 bp) replaced by a <i>Sm^R</i> cassette in reversed orientation	This work
<i>Δup-glgA</i> -600 (AI)	<i>ΔglgBX</i> * with upstream 696 bp of <i>glgC</i> (-1296 to -600 of <i>glgA</i>) replaced by a <i>Sp^R</i> cassette in reversed orientation	This work
<i>Δup-glgA</i> -450 (AII)	<i>ΔglgBX</i> * with upstream 846 bp of <i>glgC</i> (-1296 to -450 of <i>glgA</i>) replaced by a <i>Sp^R</i> cassette in reversed orientation	This work
<i>Δup-glgA</i> -300 (AIII)	<i>ΔglgBX</i> * with upstream 996 bp of <i>glgC</i> (-1296 to -300 of <i>glgA</i>) replaced by a <i>Sp^R</i> cassette in reversed orientation	This work
Plasmids		
pCP20	Plasmid expressing FLP recombinase. <i>Amp^R</i> , used for removal of <i>Km^R</i> cassettes	[28]
pKG137	Plasmid including <i>lacZY</i> and <i>Km^R</i> cassette, used for transcriptional fusions construction	[53]

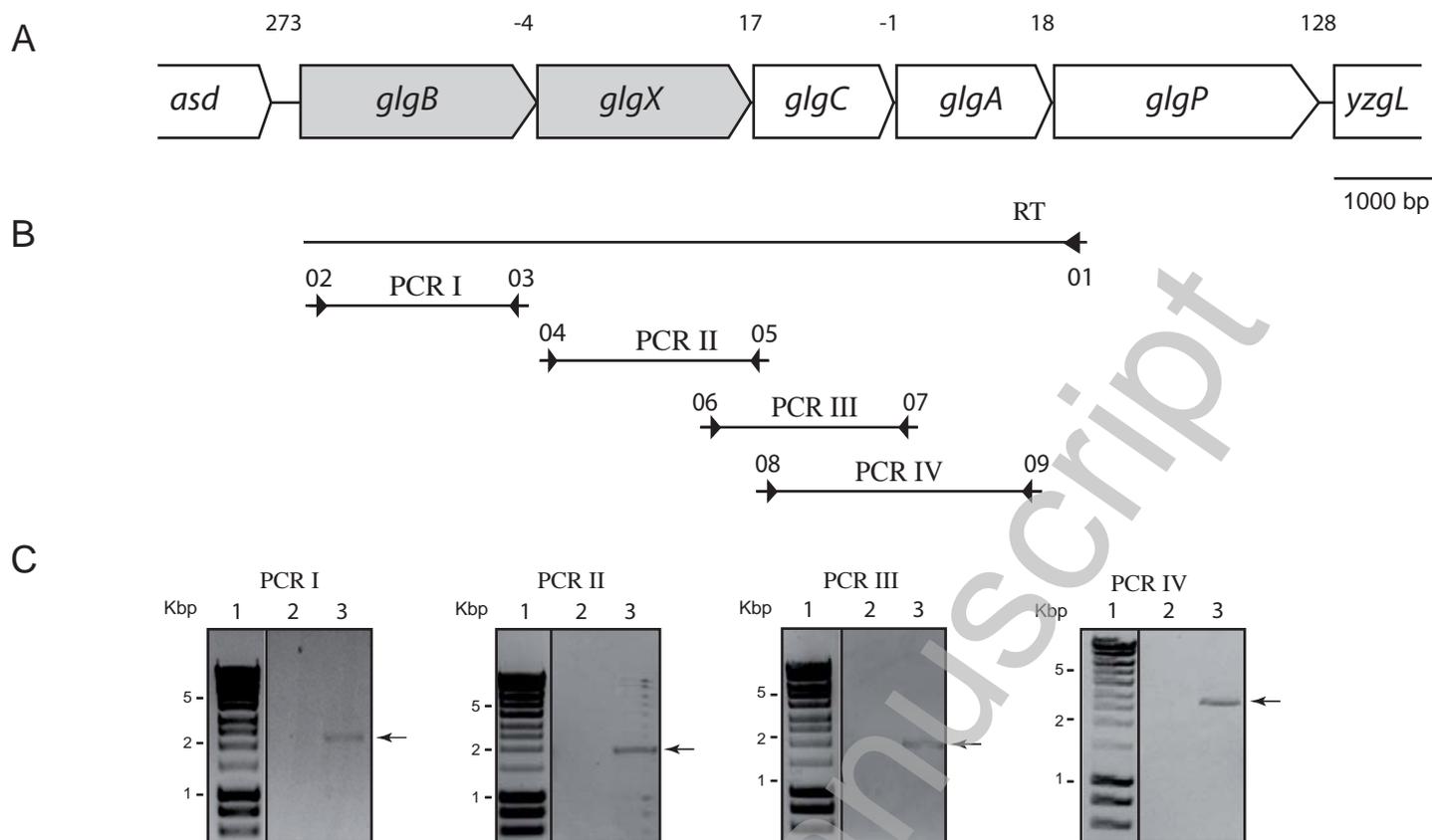
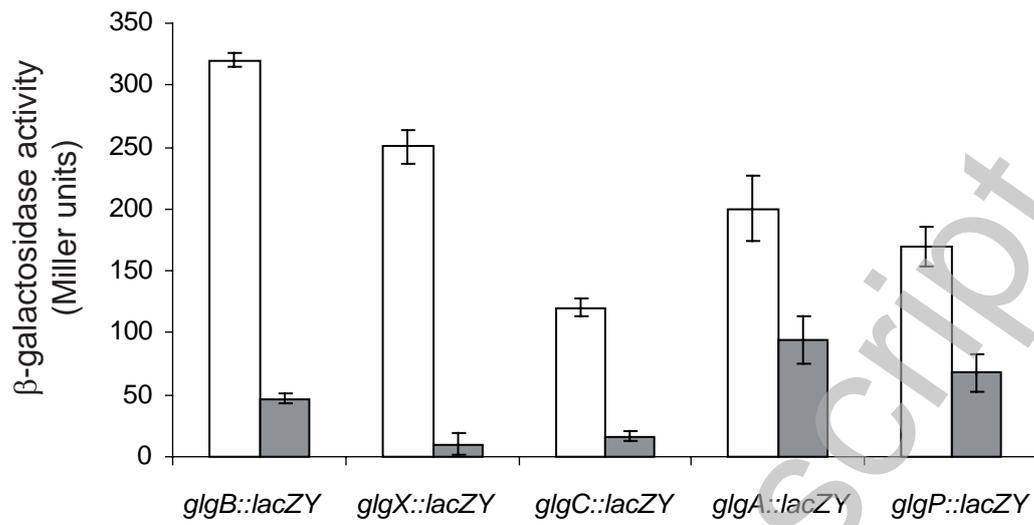


Figure 1

A



B

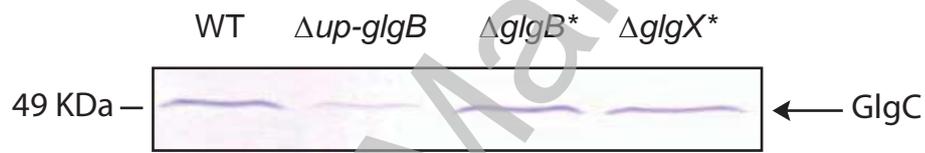


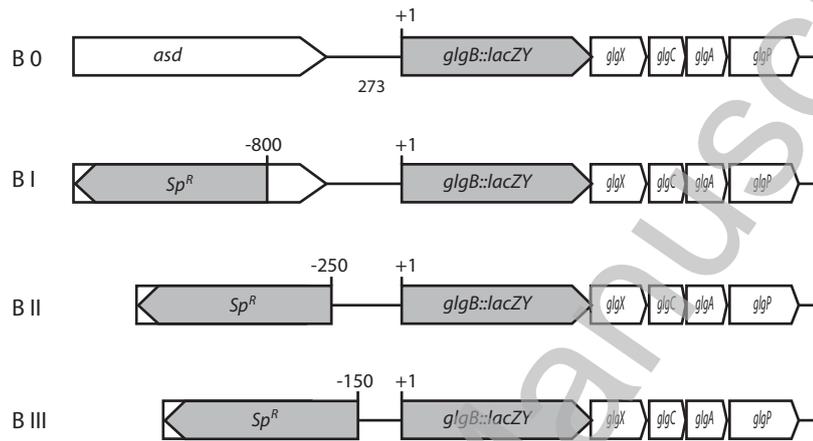
Figure 2

A

```

-276  GTAATCTTTTATTTCATTAAAATCTGGGGCGCGATGCCGCCCTGTTAGTGCGTAATACAGGAGTAAGCGCAGATGTTTCATGATTTACCGGGAGTTAAATAG
      -10      *      BII      -35
-176  AGCATTGGCTATTCTTTAAGGGTGGCTGAATACATGAGTATTACAGCCTTACCTGAAGTGAGGACGACGCAGAGAGGATGCACAGAGTGTGCGCCGTT
      -10      *      BIII
-76   CAGGTCAAAAAATGTCACAACCAGAAGTCAAAAAATCCAATTGGATGGGGTGACACAATAAAACAGGAAGAGACAAGCAATGTCCGATCGTATCGATAGAGAC
      SD      +1
  
```

B



C

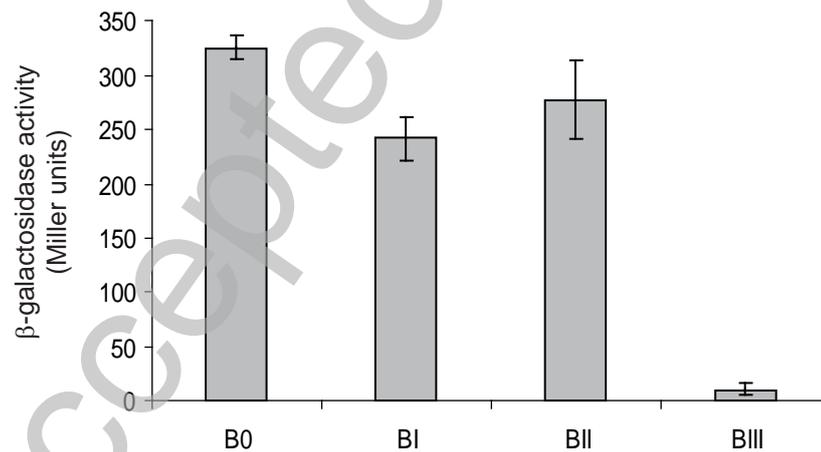


Figure 3

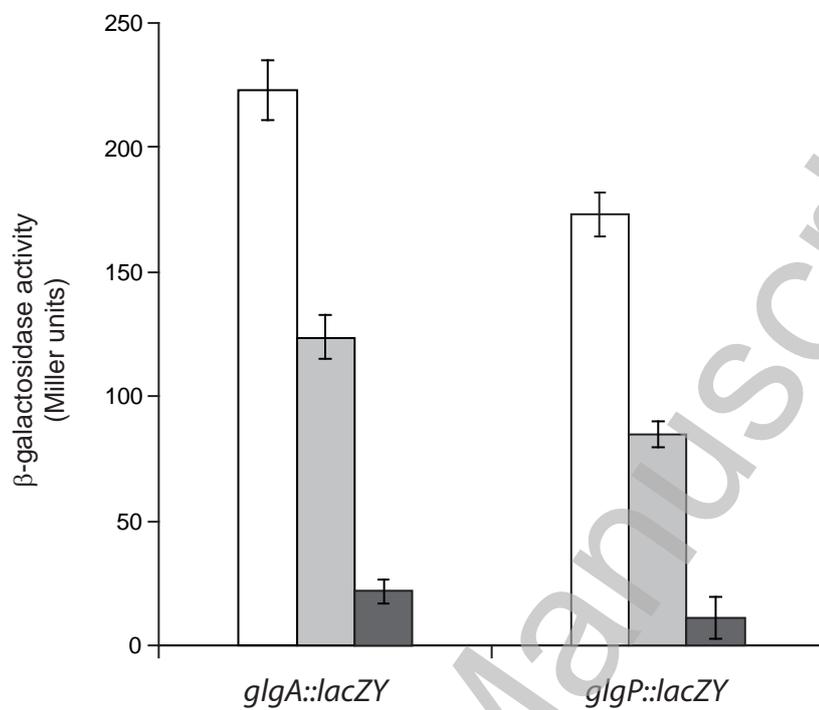


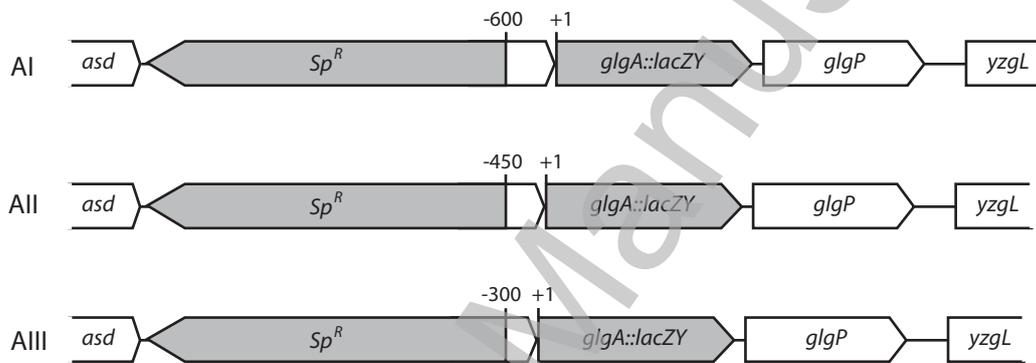
Figure 4

A

```

-695 TCAATGCCGAACGATCCGAGCAAATCTCTGGCGAGTATGGGTATCTACGTCTTTGACGCCGACTATCTGTATGAACTGCTGGAAGAAGACGATCGCGATG
-595 AGAACTCCAGCCACGACTTTGGCAAAGATTTGATTCCTCAAGATCACCGAAGCCGGTCTGGCCTATGCGCACCCGTTCCCGCTCTCTTTCGCTACAATCCGA
-495 CCCGGATGCCGAGCCGTACTGGCGCGATGTGGGTACGCTGGAAGCTTACTGGAAAAGCGAACTCGATCTGGCCTCTGTGGTGCCGGAAGTGGATATGTAC
-395 GATCGCAATTGGCCAATTTCGCACCTACAATGAATCATTTACC GCCAGCGAAAATTCGTGCAGGATCGCTCCGGTAGCCACGGGATGACCCCTTAACTCACTGG
-295 TTTCCGGCGGTTGTGTGATCTCCGTTCCGGTGGTGGTGCAGTCCGTTCTGTTCGCGCGTTCCGCGTGAATTCATTCTGCAACATTGATTCGCGCGTATT
-195 GTTACCGGAAGTATGGGTAGGTCGCTCGTGCCTGCGCCGCTGCGTCATCGATCGTCTGTGTTATTCCGGAAGGCATGGTATTGGTGAAAACGCA
-95 GAGGAAGATGCACGTCGTTTCTATCGTTT CAGAAGAAGGCATCGTCTGGTAACGCGCGAAAATGCTACGGAAGTTAGGGCATAAAAGGAGCGATAATGCA
    
```

B



C

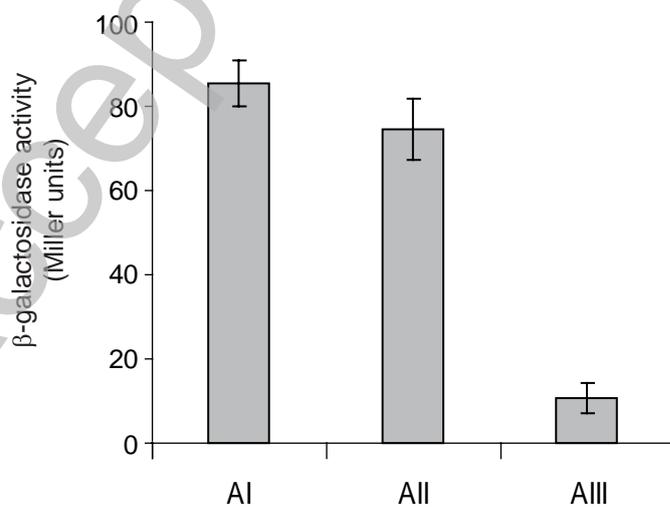
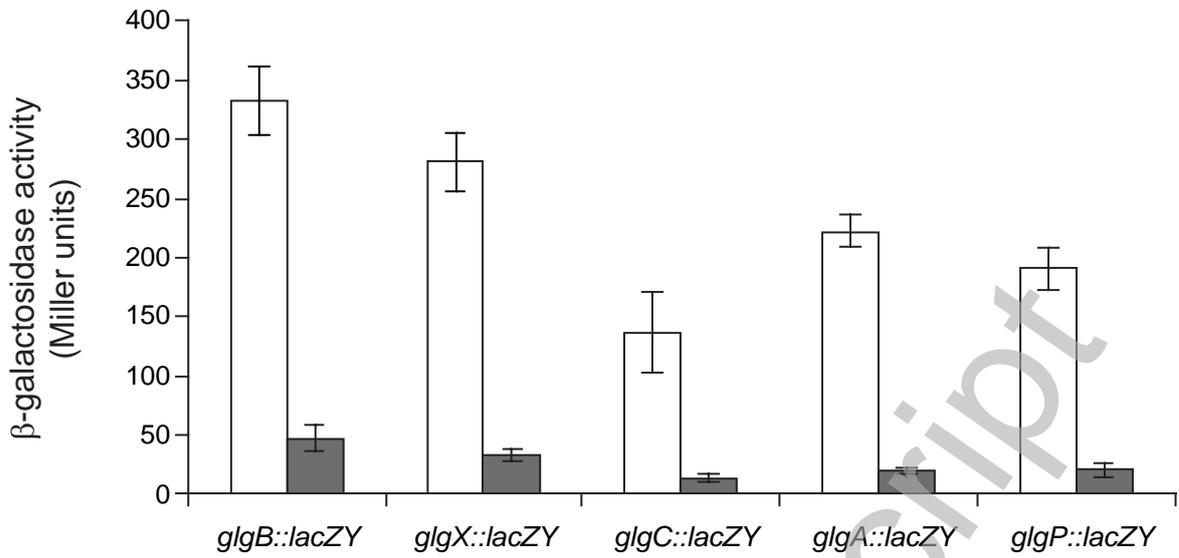
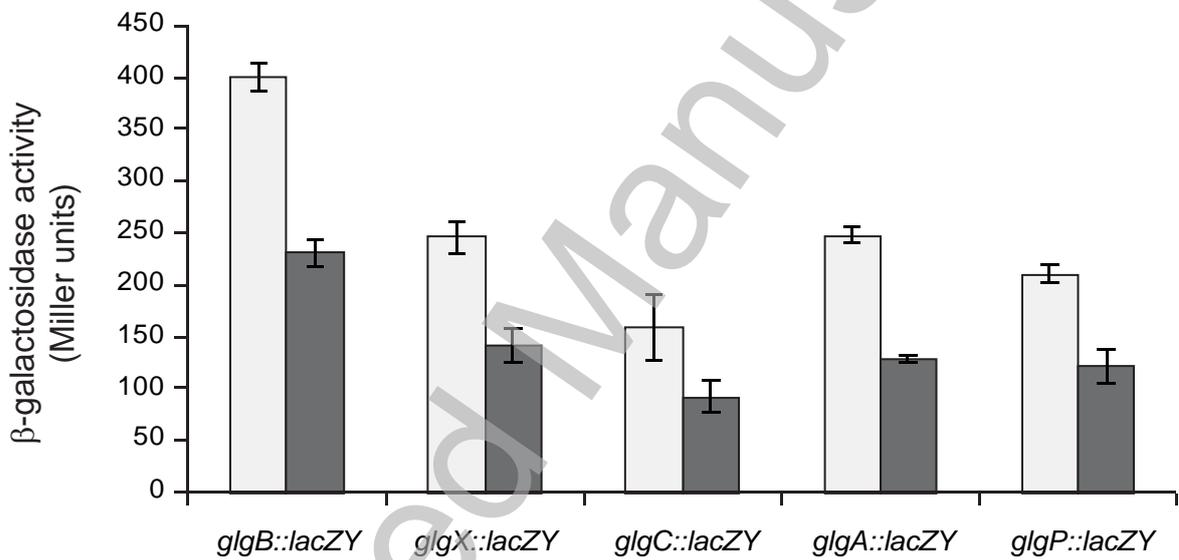


Figure 5

A



B



C

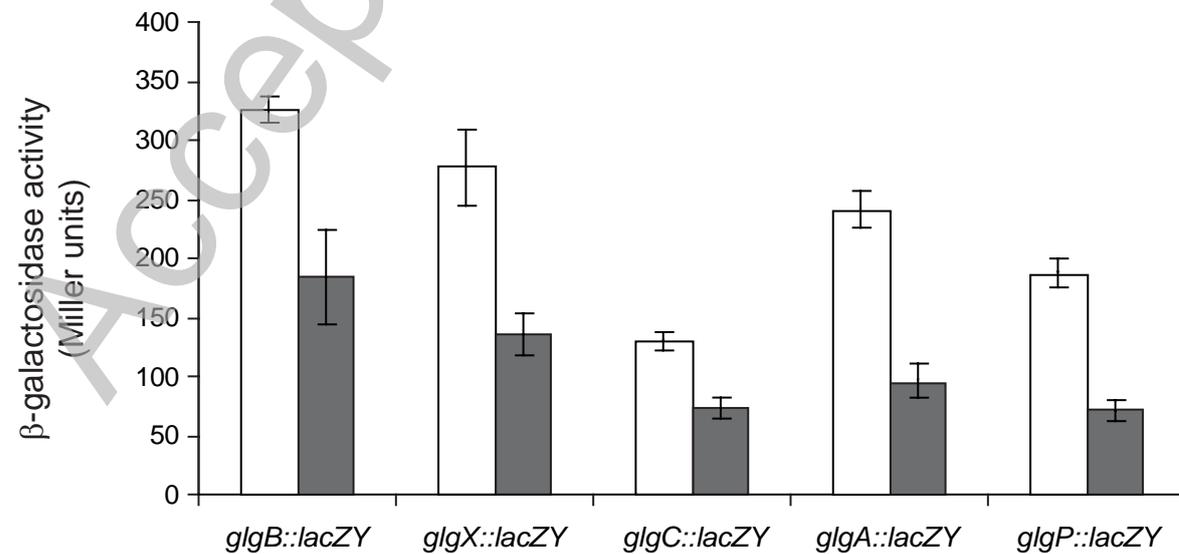


Figure 6

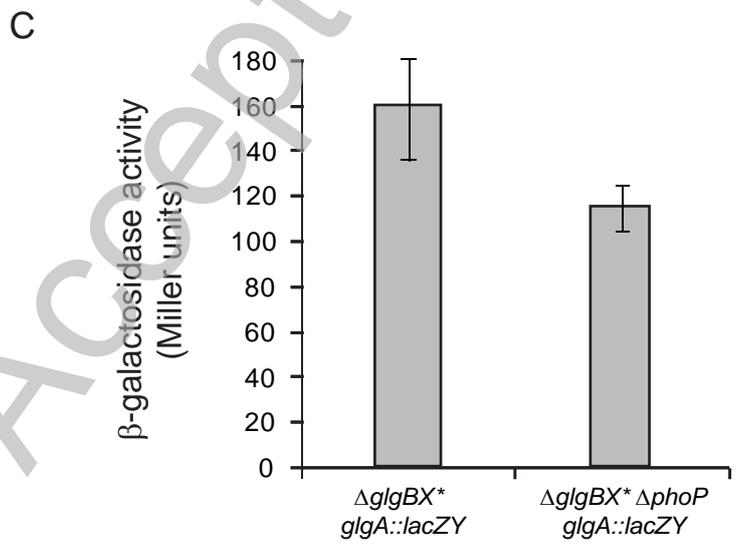
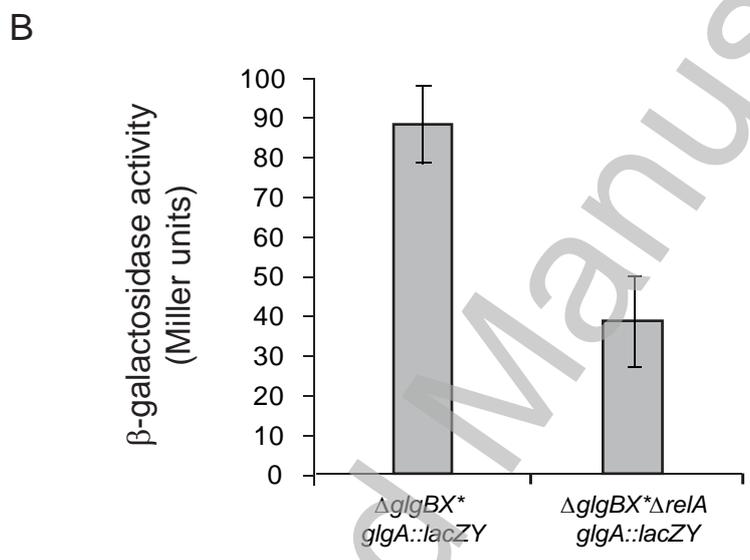
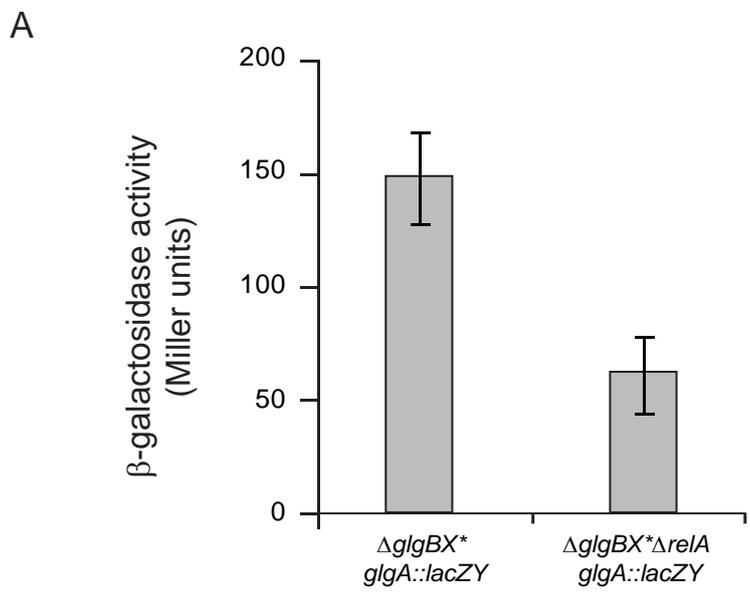


Figure 7