The EMBO Journal vol.14 no.16 pp.3958–3965, 1995

Co-ordinated regulation of amino sugar biosynthesis and degradation: the NagC repressor acts as both an activator and a repressor for the transcription of the *glmUS* operon and requires two separated NagC binding sites

Jacqueline Plumbridge

Institut de Biologie Physico-chimique (URA1139),
13 rue P. et M.Curie, 75005 Paris, France

The NagC repressor controls the expression of the divergently transcribed *nagE–nagBACD* operons involved in the uptake and degradation of the amino sugars, *N*-acetyl-D-glucosamine (GlcNAc) and glucosamine (GlcN). The *glmUS* operon, encoding proteins necessary for the synthesis of GlcN (*glmS*) and the formation of UDP-GlcNAc (*glmU*), is transcribed from two promoters located upstream of *glmU*. In the absence of amino sugars both promoters are active. However, in the presence of GlcNAc, the *glmU* proximal promoter, P1, is inactive while the upstream promoter, P2, is subject to weak induction. Two binding sites for the NagC repressor are located at –200 and –47 bp upstream of P1. Mutations which prevent NagC binding to either of these sites eliminate expression from the P1 promoter. This shows that binding of NagC is necessary for expression of the *glmU* P1 promoter and implies that NagC is playing the role of activator for this promoter. Moreover, the location of the distal NagC site suggests that this site is behaving like an upstream activating sequence (UAS).

Keywords: amino sugars/NagC/operators/transcriptional activator/transcriptional repressor

Introduction

There is an inherent dichotomy in the metabolism of many biologically important molecules, e.g. amino acids, nucleic acids, carbohydrates, lipids and their precursors, which are implicated in both anabolic and catabolic functions. In all cases, control mechanisms must exist to maintain correctly balanced growth and avoid a proverbial 'futile cycle' between biosynthesis and degradation. One logical way to achieve this is by the action of a 'switch' which, under specific conditions, can activate one pathway and/or inhibit the other.

The amino sugars, *N*-acetyl-D-glucosamine (GlcNAc) and D-glucosamine (GlcN) are good examples of compounds with two distinct functions in bacteria. They have structural roles as components of the peptidoglycan and the lipopolysaccharide in Gram-negative species and they are also valuable energy sources. The catabolic *nag* genes of the degradative cycle (Figure 1) are under classical negative control by a repressor, encoded by the gene *nagC*, which binds to operators overlapping the divergent *nagE* and *nagB* promoters forming a loop of DNA (Plumbridge and Kolb, 1991, 1993). As is the case for most catabolic operons, the *nag* genes are induced by the presence of the relevant sugars in the medium (White, 1968; Vogler and Lengeler, 1989; Plumbridge, 1990).

There is, in addition, a biosynthetic pathway (Figure 1) of which two of the genes, *glmU* and *glmS*, are located adjacent to each other on the *Escherichia coli* chromosome (Walker *et al.*, 1984; Mengin-Lecreulx and van Heijenoort, 1993, 1994). The expression of *glmS*, encoding glucosamine synthase, was shown to be transcriptionally regulated; both the protein and mRNA levels are decreased ~3-fold under conditions which induce the expression of the *nag* genes (White, 1968; Plumbridge *et al.*, 1993). There are no obvious promoter or termination signals in the *glmU–S* intergenic region (Walker *et al.*, 1984) and fusions with *lacZ* failed to detect any promoter activity in this region, implying that these two genes of the biosynthetic *glm* pathway are co-transcribed. Thus an analysis of the expression of *glmU* was undertaken.

Results

The glmU gene is expressed from two promoters

S1 mapping located two mRNA 5′ extremities upstream of the *glmU* gene. Using the Glm4–Glm5 fragment (Figure 2A) as probe and mRNA from a wild-type strain grown on glucose, two transcripts of unequal intensity, of ~250 and 360 nt, were detected. The shorter transcript is the more abundant (Figure 3, lane 1). Using mRNA from the same strain grown on GlcNAc (lane 3), this shorter transcript disappears, while with mRNA from a culture grown on GlcN it is considerably reduced (lane 2). The loss of the shorter transcript was not due just to the presence of amino sugars in the medium, because it was absent in a strain carrying a null mutation in the gene for the Nag repressor, *nagC* (lane 4), and in all other conditions which provoke a derepression of the *nagE–B* operons (see below).

The mRNA 5′ ends were located more precisely by primer extension. The labelled oligonucleotide used as primer, Glm3, hybridizes to the mRNA near the beginning of the *glmU* structural gene (Figure 2A). The reverse transcripts detected in glucose-grown mRNA preparations were 92/93 and 195 nt in length (Figure 4A, lane 1) and coincide with the positions of the 5′ ends predicted from the S1 experiment. Growth on GlcNAc or GlcN (lanes 2 and 3) or the presence of mutations in the genes *nagC*, *nagA*, *nagB* or a deletion of the *nagE–BACD* region (lanes 4, 5, 7 and 9), all of which provoke the derepression of the *nagE–nagB* operons (Plumbridge, 1991), also reduce or eliminate the shorter *glmU* transcript, while mutations in *nagE* or *nagD*, which have no effect on *nag* gene expression, do not affect *glmU* expression (Figure 4A, lanes 6 and 10).

The 5′ ends of the P1 and P2 transcripts could correspond to transcription initiation sites, and inspection of

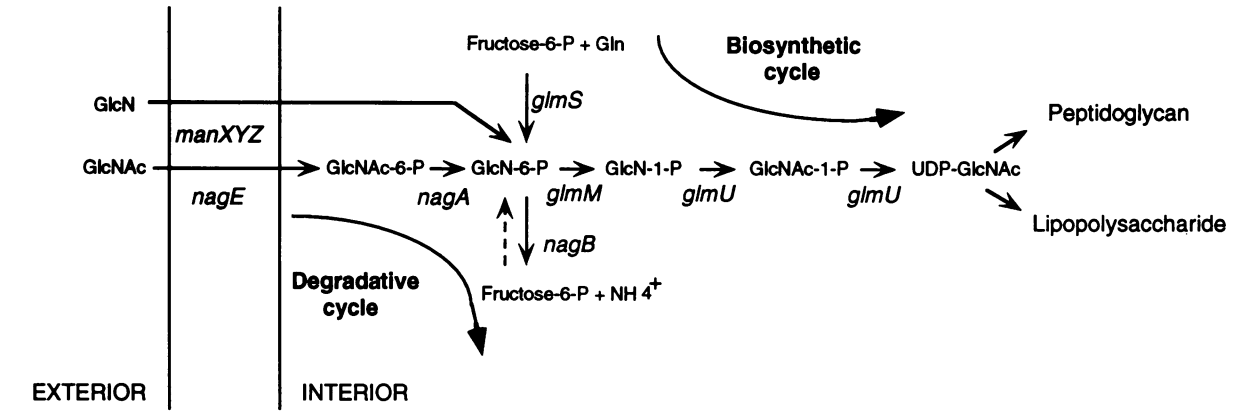

Fig. 1. Metabolism of *N*-acetylglucosamine. The scheme is presented to emphasize the separation of the metabolism into a catabolic process using the *nag* genes and an anabolic process involving the *glm* genes. In *E.coli*, both GlcN and GlcNAc are PTS sugars; they are taken up by the phosphoenol pyruvate-dependent phosphotransferase system (PTS, genes *manXYZ* and *nagE*) to give intracellular GlcN-6-P and GlcNAc-6-P (reviewed in Postma *et al.*, 1993). The phosphorylated sugars are subsequently degraded to fructose-6-P and ammonia by the products of the genes *nagA* and *nagB* (White, 1968). When GlcN or GlcNAc is present in the environment it is used for peptidoglycan and lipid A synthesis but, in the absence of external amino sugars, the gene *glmS*, whose product synthesizes GlcN-6-P from fructose-6-P and GlcN, is essential. The *glmU* gene encodes a bifunctional enzyme which is responsible for the acetylation of GlcN-1-P to give GlcNAc-1-P and the synthesis of UDP-GlcNAc (Mengin-Lecreulx and van Heijenoort, 1993, 1994). The *glmM* gene, encoding the enzyme responsible for the conversion of GlcN-6-P to GlcN-1-P, has been identified recently by D.Mengin-Lecreulx (manuscript in preparation).

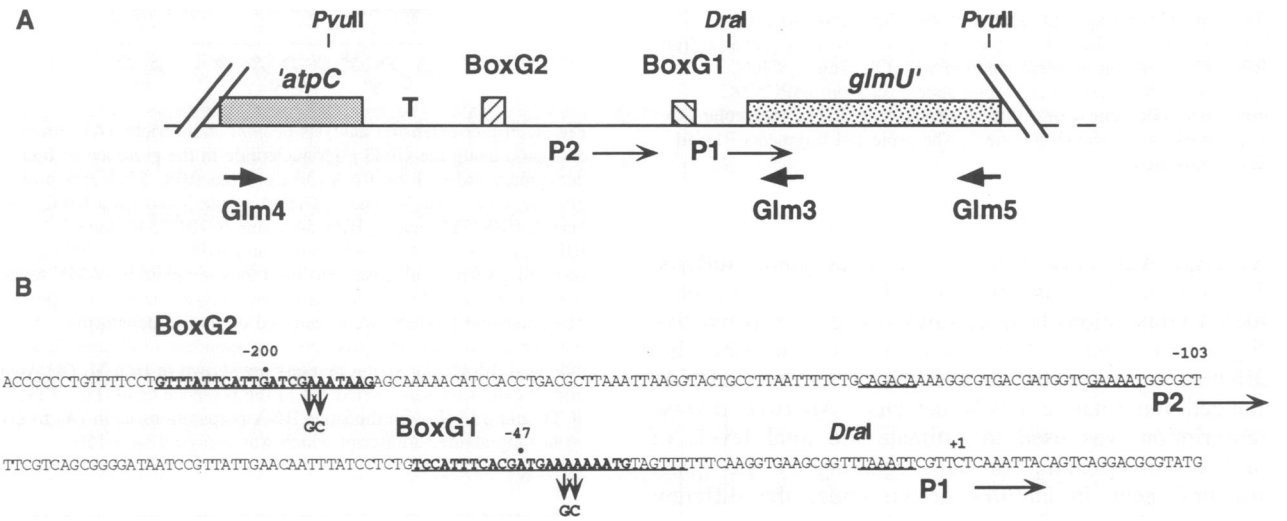

Fig. 2. Organization of the region upstream of *glmU*. (**A**) The intergenic region between *atpC* and *glmU* is shown with the relative locations of the two promoters, P1 and P2, and the two NagC binding sites, BoxG1 and BoxG2. T is the terminator of the *atp* operon. The location and orientation of the three oligonucleotides Glm3, Glm4 and Glm5 are shown. The *Pvu*II–*Pvu*II and *Pvu*II–*Dra*I fragments indicated were used in the construction of *glmU–lacZ* fusions. (**B**) Sequence of the *glmU* promoter–operator region. The sequence is numbered from the P1 transcription start site. The P2 start site is at –103 and the centres of the NagC boxes are at –47 and –200 on this scale. The mutations created in BoxG1 and BoxG2 are indicated.

the DNA sequence upstream of them identified a potential promoter sequence by homology with the consensus –10 and –35 sequences. The putative promoter for the longer but weaker transcript, P2, is CAGACA-18 bp-GAAAAT, while for the shorter transcript the best match with the consensus promoter is TAGTTT-18 bp-TAAATT (Figure 2). The underlined bases are those agreeing with the normally accepted consensuses for these two regions (Hawley and McClure, 1983; Harley and Reynolds, 1987). A mutation in the –10 region of the proposed P2 promoter (GAAAAT to GGACAT) eliminates the P2-derived transcript *in vivo* without affecting the P1-derived transcript, confirming that P1 and P2 are independent promoters (data not shown).

Levels of glmU mRNA under repressed and activated conditions

The pattern of transcription observed by the two techniques, S1 and primer extension, was identical and the relative intensities of the long and short transcripts in the different mRNA preparations were very similar by the two techniques. Quantitation of the radioactivity in the longer transcript compared with the total in lanes where both transcripts were detected suggested that the longer

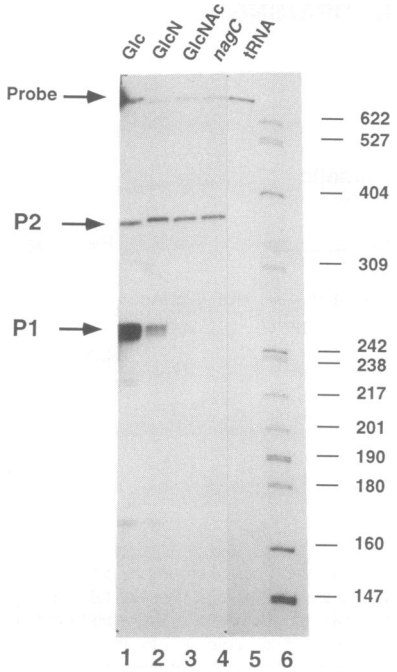

Fig. 3. S1 mapping of *glmU* transcripts. The probe used was Glm4–Glm5 (Figure 2A), it covers the C-terminal region of the preceding gene, *atpC*, the 352 bp intergenic region between *atp* and *glmU*, which includes a *rho*-independent transcription termination signal for the *atp* operon, and 180 nt of the *glmU* structural gene (Walker *et al.*, 1984). The Glm4–Glm5 fragment labelled at the Glm5 end (~0.2 pmol, 25 000 c.p.m.) was hybridized with 30 μg of total RNA prepared from IBPC5321 grown in minimal medium with Glc, lane 1; GlcN, lane 2 or GlcNAc, lane 3 as carbon source, or from IBPC529C grown with Glc, lane 4 or 30 μg of tRNA, lane 5. The molecular sizing markers are shown in lane 6. The protected fragments P1 and P2 are indicated.

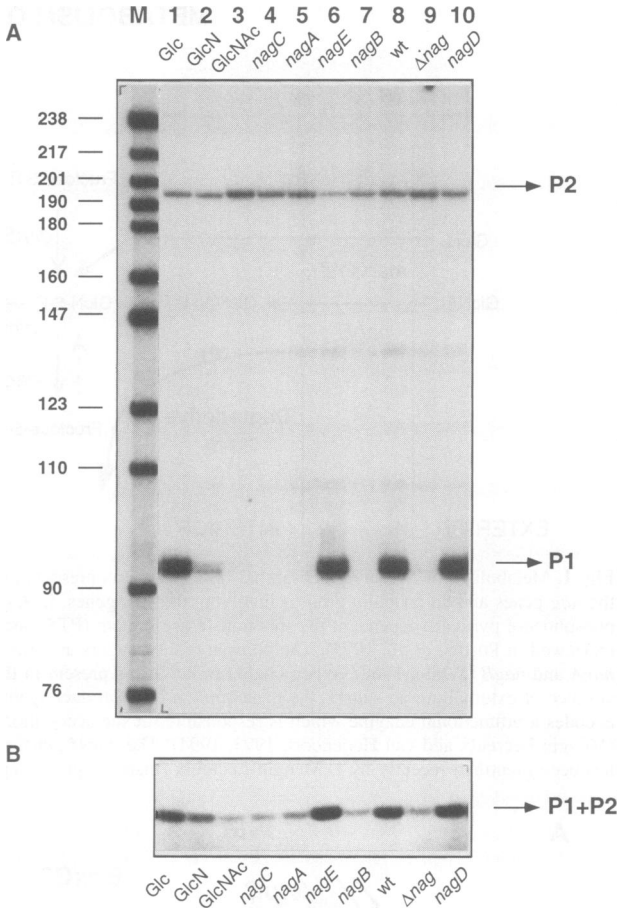

Fig. 4. Primer extension analysis of *glmU* transcripts. (**A**) Primer extension using the Glm3 oligonucleotide in the presence of four deoxynucleotides. Total RNA (30 μg) from IBPC 5321 grown on glucose, lane 1; GlcN, lane 2; GlcNAc, lane 3 and from IBPC529C, lane 4; IBPC524, lane 5; IBPC542, lane 6; IBPC546, lane 7; IBPC5321, lane 8; IBPC590, lane 9 and IBPC540, lane 10 (lanes 4–10 were all grown on glucose) was used as a template for AMV reverse transcriptase with the 5′- labelled Glm3 oligonucleotide (~1 pmol). The extension products were analysed on an 8% denaturing polyacrylamide gel. The products corresponding to P1 and P2 are indicated. Molecular sizing markers are shown in lane M. (**B**) Abortive primer extension was carried out in the presence of dGTP, dTTP, dCTP and ddATP with the same RNA preparations as in (A) to give a 36 nt long aborted transcript which was analysed on a 15% polyacrylamide gel.

transcript was only 10% of the total *glmU* mRNA. However, the P2 transcript was ~2-fold more abundant in mRNA preparations from cultures where P1 was inactive (Figure 4A, lanes 3, 4, 5, 7 and 9). It was possible that differences in structure and length of the two mRNAs changed the relative levels detected. Abortive reverse transcription was used to estimate the total levels of mRNA corresponding to the beginning of the *glmU* structural gene in cultures grown under the different conditions. The Glm3 oligonucleotide was used to prime reverse transcriptase in the presence of three deoxynucleotides and one dideoxynucleotide, to produce a short, aborted transcript of 36 nt, including the primer (Figure 2B). The amount of radioactivity in each band was quantitated and the relative levels are given in Table I. The level of *glmU* mRNA during growth on GlcNAc or in the presence of a mutation which derepresses the *nag* degradative genes is ~30% of that found in the wild-type strain grown on glucose.

Table I. Relative levels of *glmU* mRNA in different strains

Strain	Genotype	Medium	mRNA level
IBPC5321	wild-type	Glc	1.0
IBPC5321	wild-type	GlcN	0.64
IBPC5321	wild-type	GlcNAc	0.34
IBPC529C	*nagC*	Glc	0.28
IBPC524	*nagA*	Glc	0.35
IBPC546	*nagB*	Glc	0.28
IBPC590	Δ*nagEBACD*	Glc	0.35
IBPC542	*nagE*	Glc	1.4
IBPC540	*nagD*	Glc	1.2

There are two NagC operator sites upstream of glmU

The perfect inverse correlation between the induction of *nag* and repression of *glm* genes posed the question as to whether the regulation of *glmU* expression involved the NagC protein. DNase I footprinting experiments located two binding sites for the NagC repressor, one upstream of the P2 promoter and the second, which was only weakly protected, in the region between P2 and P1 (Figure 5, lanes 2 and 3, this footprint was obtained on a *glmU–lacZ* fusion, see below). Despite their comparatively weak binding of NagC, inspection of the DNA sequence in

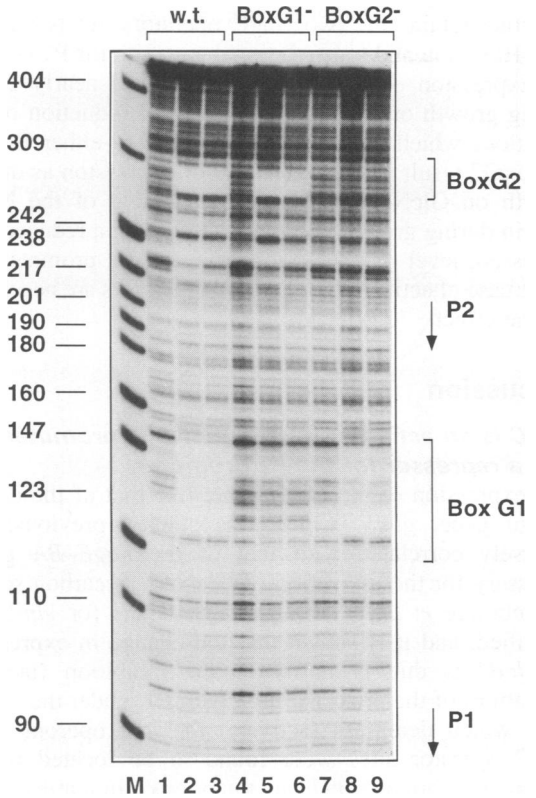

Fig. 5. Binding of NagC to wild-type and mutated *glmU* upstream sequences. The DNA fragment used was made by PCR using plasmids carrying wild-type or mutated *glmU–lacZ* fusions as templates. It is labelled at the Glm3 oligonucleotide and extends to the *Pvu*II site in *atpC* (Figure 2A and Materials and methods). DNase I protection experiments were performed with 100 μg/ml (lanes 2, 5 and 8) or 50 μg/ml (lanes 3, 6 and 9) of an extract from a NagC-overproducing strain and ~4 nM labelled DNA. Products were analysed on an 8% denaturing polyacrylamide gel. Molecular sizing markers are in lane M. The two protected regions corresponding to BoxG1 and BoxG2 are indicated, as are the locations of the P1 and P2 transcription start sites. Lanes 1–3, wild-type sequence; lanes 4–6, BoxG1⁻ fragment; lanes 7–9, BoxG2⁻ fragment.

these two regions identified sequences, called BoxG1 and BoxG2, in agreement with the consensus for NagC operators (Figure 2B). Previously, four NagC operators had been characterized, two in the intergenic *nagE–B* region and two upstream of the *manXYZ* operon (Plumbridge and Kolb, 1991). Recently, a detailed mutagenesis study of the *nagE* operator has identified the two AT base pairs at positions ±5 and ±6 as the most important recognition points for NagC binding and has shown that the base pairs at positions 3, 4 and 11 also contribute to the binding strength (Plumbridge and Kolb, 1995). The four previously characterized sites plus the two sites located upstream of *glmU* are compared in Figure 6. The AT base pairs at positions ±5, ±6 are strictly conserved. BoxG2, which binds repressor more strongly than BoxG1, shows higher homology with the deduced consensus. In particular, the BoxG2 has Gs at positions ±11, which produce stronger binding sites, and three out of four agreements with the deduced consensus at positions ±3 and ±4. BoxG1 is less like the consensus and seems to bind repressor rather weakly, as measured by DNase I footprinting.

NAGC OPERATOR SITES

```
1110 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 91011

 T T T A A T T T G C G A T A C G A A T T A A A  BOX E
 C T T A T T T T A T C A T T C A A A A A A T C  BOX B
 G A T A T T T T A C C T T T C G A A A T T T C  BOX M
 T T A T T T T T A G A T A T C G A A A A A A A T  BOX M'
 G T T T A T T C A T T G A T C G A A A T A A G  BOX G2
 T C C A T T T C A C G A T G A A A A A A A T G  BOX G1
 ─────────────────────────────────────────────
 g t t a t T T c g           c g A A a t a a c CONSENSUS
 c a a t a                             t a t t g
```

Fig. 6. Comparison of NagC operator sites. The four previously characterized NagC binding sites, BoxE, BoxB, BoxM and BoxM′ (Plumbridge and Kolb, 1991) are aligned with the two binding sites BoxG1 and G2 upstream of *glmU*. The consensus is the deduced optimal sequence found by a mutagenesis analysis of BoxE (Plumbridge and Kolb, 1995). The A residues hypermethylated by dimethyl sulfoxide (DMS) in the presence of NagC in boxes E, B and M′ are shown in bold. The effect of DMS on BoxG1 and G2 has not been investigated.

Table II. Relative expression levels from *glmU-lacZ* fusions carrying wild-type and mutant NagC boxes

Fusion	Strain	Glc (%)	GlcNAc (%)
glmU P1P2–lacZ			
Wild-type	wild-type	100	20
	nagC	17	ND
BoxG1⁻	wild-type	12	10
BoxG2⁻	wild-type	21	16
glmU P2–lacZ			
Wild-type	wild-type	100	180
BoxG1⁻	wild-type	180	175
BoxG2⁻	wild-type	170	175

Activities are expressed as the percentage of the wild-type fusion grown on glucose. The β-galactosidase activities from the wild-type P1P2 fusion in IBPC5321 grown on Glc is 1130 ± 30 Miller units and from the P2 fusion 26 ± 4 units.
ND, not determined.

Effect of the NagC boxes on activation of promoter P1

A transcriptional fusion between *glmU* and *lacZ* carried on single copy λ lysogens was used to quantitate *glmU* expression further under different growth conditions (Table II). There is a 5-fold drop in β-galactosidase activity in the presence of GlcNAc in the growth medium, or in a strain with a *nagC* mutation.

To investigate the importance of the NagC boxes for the regulation, mutations were introduced into each box to prevent NagC binding. They were the replacement of the two absolutely conserved AA nucleotides at positions +5 and +6 of the consensus with GC base pairs. The effect of these mutations in either of the NagC boxes was to reduce the level of expression during growth on glucose and thus eliminate the differential expression during growth on glucose or GlcNAc (Table II). The BoxG1 mutation reduced expression to 10% of the wild-type. fusion, while the BoxG2 mutation allowed nearer 20% expression.

DNase I footprinting confirmed that the mutations have indeed eliminated NagC binding to the *glmU* operators (Figure 5). The footprint corresponding to the wild-type sequence is shown in lanes 2 and 3. The mutation in BoxG1 (lanes 5 and 6) eliminated the NagC footprint at

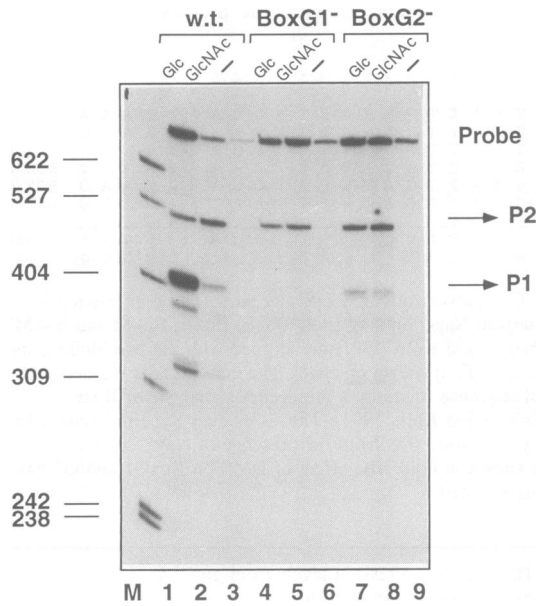

Fig. 7. S1 analysis of transcripts from wild-type and mutated *glmU–lacZ* fusions. The DNA probes used were made by PCR using plasmids carrying wild-type or mutated *glmU–lacZ* fusions as templates. They are labelled at the Lac22 oligonucleotide present in the beginning of the *lacZ* gene and extend to the *Pvu*II site in *atpC*. The 5′-end-labelled fragments were hybridized with RNA (30 μg) made from IBPC5321 carrying the same wild-type or mutated fusions on lysogens grown on glucose (lanes 1, 4 and 7) or GlcNAc (lanes 2, 5 or 8) or with 30 μg tRNA (lanes 3, 6 and 9). S1-resistant products were analysed on an 8% denaturing acrylamide gel. Molecular sizing markers are shown in lane M. Protected fragments corresponding to P1 and P2 transcripts are indicated.

BoxG1 but did not affect the binding to BoxG2. The mutation in BoxG2 eliminated binding to BoxG2 and seemed to reduce binding to BoxG1 (Figure 5, lanes 8 and 9).

The effect on the transcription from P1 and P2 was measured by S1 mapping of mRNA prepared from the lysogens and using a probe specific for the *glmU–lacZ* fusion (Figure 7). The wild-type construct from a lysogen grown in glucose showed the two transcripts corresponding to promoters P1 and P2. The P1 transcript was inhibited severely by growth on GlcNAc (Figure 7, compare lanes 1 and 2). For the lysogen carrying the mutation in BoxG1, the P1 transcript is not observed and there is no difference between glucose- and GlcNAc-grown mRNAs. For the mRNA from the lysogen carrying the BoxG2 mutation, the P1 transcript in mRNA made from a glucose-grown culture is greatly reduced in intensity but not eliminated and there is no difference between glucose- and GlcNAc-grown cultures. This correlates with the slightly higher levels of expression from the *glmU–lacZ* fusion carrying the BoxG2 mutation than the BoxG1 mutation (Table II). Thus both NagC boxes are necessary for efficient transcription from the P1 promoter.

Effect of the NagC boxes on expression of P2

The primer extension and S1 experiments (Figures 3 and 4) suggested that when P1 was activated the transcript from P2 was reduced in intensity. To test this *in vivo*, a fusion between *glmU* P2 and *lacZ* was constructed using the *Dra*I site between BoxG1 and the P1 promoter. This construct retains the two NagC operators but is missing the –10 region and transcriptional start site for P1 (Figure 2). Expression of this fusion is increased nearly 2-fold during growth on GlcNAc (Table II). Introduction of the mutations which abolish NagC binding to either BoxG1 or BoxG2 result in the same level of expression as during growth on GlcNAc. In this case, binding of the NagC protein during growth on glucose results in a reduced, i.e. repressed, level of expression from the P2 promoter. As in the case of activation of P1, both operators are necessary for the effect.

Discussion

NagC is an activator for the glmU P1 promoter and a repressor for the P2 promoter

The expression of the *glmU* gene, like that of the downstream gene, *glmS*, which was studied previously, is inversely correlated with that of the *nagE–BA* genes necessary for the utilization of GlcNAc as carbon source (Plumbridge *et al.*, 1993). Two promoters for *glmU* are identified, and it is shown that the change in expression of *glmU* is due primarily to the inhibition (lack of activation) of the proximal promoter, P1, under the conditions which derepress (activate) the *nag* operons. Two NagC operator sites were found to be located in the region upstream of *glmU*, and mutations eliminating NagC binding to either one of these sites prevent the activation of the P1 promoter. Taken together, these results show that the NagC repressor is acting as an activator for the P1 promoter and that the two binding sites are necessary for the activation.

The activity of the P2 promoter is also sensitive to the presence of GlcNAc in the medium, but in the opposite sense to P1. The expression from P2 was activated ~2-fold when GlcNAc was present. The same mutations which eliminate the activation of P1 also eliminate the repression of P2 (Table II). The overall change in *glmU*-specific mRNA, as measured by the abortive reverse transcriptase technique, is only ~3-fold. It reflects the opposing effects, the repression of P2 and the activation of P1, but is somewhat less than predicted by the *glmU–lacZ* fusions (5-fold). This could indicate that there is some post-transcriptional regulation affecting mRNA levels. The comparatively feeble level of regulation, compared with the 40-fold regulation of the *nagE–BA* genes, is presumably because *glmU* encodes an essential biosynthetic enzyme.

Comparison of NagC as activator and repressor

The centres of the NagC boxes are at –47 and –200 compared with the P1 start site. The majority of transcriptional activators bind in the region –40 to –50 from the +1 of the activated promoters, e.g. CAP, FNR (for reviews, see Collado-Vides *et al.*, 1991; Gralla, 1991). At this location, it is presumed that the activator makes protein–protein contacts with RNA polymerase bound at the promoter (Ishihama, 1993; Busby and Ebright, 1994). The location of a NagC box at –47 relative to P1 is consistent with it being an activation site for the P1 promoter when NagC is bound.

The P2 promoter is induced ~2-fold by growth on GlcNAc. Moreover, both operators are again necessary

for this regulation. The location of the two operators ~100 bp upstream and 50 bp downstream of the P2 promoter suggests that repression does not occur by direct inhibition of RNA polymerase binding to the P2 promoter. An alternative explanation is that repression occurs by the 'roadblock' mechanism described for PurR repression of the *purB* promoter (He and Zalkin, 1992). The presence of NagC bound to BoxG1 could impede the passage of polymerase from P2. In this case, the role of the upstream operator is to enhance the binding of NagC to the weaker downstream site.

It is interesting to compare the characteristics of NagC-mediated repression of *nag* degradative genes with the NagC-mediated activation of the *glm* genes. In both cases, two operator sites are involved but their positions relative to the promoters are very different and the distance between the boxes is also different: 94 bp for *nag* compared with 153 bp for *glm*. The 94 bp for the *nagE–B* operator corresponds to nine turns of B-form DNA (helix step of 10.4 bp) so that the two operators are located in-phase on the DNA and produce a characteristic pattern for the footprint of NagC on the *nagE–B* DNA. The two operators are protected from DNase I attack over at least 22 bp and the pattern of attack between the boxes is modified, producing a series of cleavages every 10–11 bp diagnostic of DNA loop formation (Hochschild and Ptashne, 1986; Plumbridge and Kolb, 1991). On the other hand, NagC binding to the two operators upstream of *glmU* did not give any indication of DNA loop formation *in vitro*, the two boxes were protected to different extents and there was no change in the pattern of attack in the intervening region. The separation between BoxG1 and G2 (153 bp) corresponds to a non-integral number of turns (14.7) of the helix with a pitch of 10.4 bp.

However, the *in vivo* results with the *glmU–lacZ* fusions demonstrate that the two NagC boxes are necessary to activate transcription from P1 and it seems reasonable to assume that an interaction between NagC molecules bound to the two sites is necessary for activation *in vivo*. One reason why no loop formation is detected *in vitro* could be because some other factor is necessary to permit NagC protein molecules bound to the two sites to approach. This factor could be a supercoiled template or an extra protein component. An accessory protein is postulated to be necessary for co-operative repression *in vivo* by the Gal repressor of two *gal* operators (Kuhnke *et al.*, 1989; Mandal *et al.*, 1990). It has been shown that integration host factor (IHF) binding in between the upstream activator site and the promoter stimulates activation of some NifA- and NtrC-regulated genes (Hoover *et al.*, 1990; Claverie-Martin and Magasanik, 1992) and also of some NarL-regulated genes (Rabin *et al.*, 1992; Schroder *et al.*, 1993). Binding of additional proteins could change the pitch of the helix. The helical spacing for AraC-induced looping of the *araBAD* promoter region was found *in vivo* to be 11.1 bp (Lee and Schleif, 1989) and for LacI binding to two *lac* operators to be 11.28 bp (Law *et al.*, 1993).

Role of the upstream operator

Upstream activating sequences (UAS) or enhancers are common in eukaryotes but comparatively rare (or not yet identified) in prokaryotes. Two systems are well characterized in Gram-negative bacteria and both involve other specialized transcription factors. For the nitrogen-regulated genes expressed by σ^{54}-carrying RNA polymerase, binding sites for the activator NtrC or NifA are located in the range of 100–200 bp upstream of the transcription start site. The role of these sites seems to be to increase the local concentration of the activator NtrC, as is characteristic of eukaryotic enhancers (reviewed in Kustu *et al.*, 1991). Expression of the FNR-dependent nitrate regulon is further activated by NarL binding to an upstream site, situated 70–200 bp upstream of the transcription start site, depending upon the promoter (reviewed in Stewart, 1993).

The case of NagC is somewhat different, in the sense that the same protein is binding adjacent to the RNA polymerase at position –47 (BoxG1) and also at the remote enhancer site (BoxG2). Although the proximal site, BoxG1, is located at a position typical of prokaryotic activators (Collado-Vides *et al.*, 1991; Gralla, 1991), it is not sufficient to allow NagC to to activate *glmU* by itself; this could be because NagC occupation of this lower affinity site is inadequate in the absence of the higher affinity upstream site. In this case the upstream site is needed to increase the local NagC concentration at BoxG1 (a chelation effect). However, the possibility that NagC bound to both sites is making direct contacts with RNA polymerase cannot be ruled out. It has been shown that a second cAMP/CAP site, suitably located upstream of another functional CAP site, synergistically enhances transcription, probably via protein–protein contacts of both CAP complexes with RNA polymerase (Joung *et al.*, 1993; Busby *et al.*, 1994).

Dual function regulatory protein

There are many examples of DNA binding proteins which are either activators or repressors according to the location of their binding sites. For example, the TyrR protein is, in most situations, a repressor but activates a certain subset of the *tyr* regulon (reviewed in Pittard and Davidson, 1991). The AraC protein activates and represses the *araBAD* genes, depending upon the presence of the inducer and cAMP/CAP which change the occupancy of the different sites (Reeder and Schleif, 1993 and references therein). The NagC protein shows strong homology to a family of xylose repressors from Gram-positive species (Titgemeyer *et al.*, 1994) and it would be interesting to know whether these repressors also possess an activator function.

In the case of NagC, the regulatory protein has evolved to control both biosynthetic and catabolic pathways for amino sugars. A similar situation exists for fatty acid biosynthesis and degradation. The FadR protein is a repressor for the *fad* genes involved in fatty acid degradation but is an activator for the *fab* genes of the biosynthetic pathway (Henry and Cronan, 1991a,b; reviewed in Black and DiRusso, 1994). The fructose repressor, FruR, appears to be a pleiotropic regulatory protein; it controls gluconeogenesis by activating the expression of a series of genes necessary for growth on Krebs cycle intermediates and repressing a number of glycolytic enzymes (Chin *et al.*, 1987; Geerse *et al.*, 1989; Ramseier *et al.*, 1993). For all three systems, NagC, FadR and FruR, release of the bound regulatory protein is achieved by the presence of the

Table III. Bacterial strains

Name	Relevant genotype	Reference
IBPC5321	*thi-1, his-4, argE3, argG6, xyl-5, mtl-1, rpsL, ΔlacX74*	Plumbridge (1991)
IBPC5321R	IBPC5321 *recA1, srl*::Tn*10*	
IBPC529C	IBPC5321 *nagC*::cm	Plumbridge (1991)
IBPC524	IBPC5321 *nagA*::cm	Plumbridge (1991)
IBPC546	IBPC5321 *nagB*::km	Plumbridge (1991)
IBPC542	IBPC5321 *nagE*::km	Plumbridge (1991)
IBPC540	IBPC5321 *nagD*::tc	Plumbridge (1991)
IBPC590	IBPC5321 Δ*nagEBACD*::tc	Plumbridge (1993)

inducing metabolite in the cell, GlcNAc-6-phosphate for NagC, long chain acyl-CoAs in the case of FadR and fructose-1-phosphate for FruR (Ramseier *et al.*, 1993). All these examples represent an economical use of a regulatory protein acting as a switch to co-ordinate expression of genes required under different conditions. The strategic positioning of the binding sites can be used to achieve opposite effects.

Materials and methods

Bacteriological methods

The bacterial strains used are listed in Table III. Bacteria were grown routinely in LB medium or minimal MOPS medium (Neidhardt *et al.*, 1974) supplemented with 50 μg/ml arginine and histidine and 0.2% glucose, GlcN or GlcNAc.

S1 mapping

Total RNA was extracted from exponentially growing cultures by the hot phenol method. To map chromosomal *glmU* mRNA, the probe used was the 720 bp PCR-synthesized fragment, Glm4–Glm5 (Figure 2A) with the Glm5 oligonucleotide 5′-end-labelled with [γ-^{32}P]ATP and polynucleotide kinase. To map mRNAs derived from the *glmU–lacZ* fusions, the probe used was a PCR-synthesized fragment using the *glmU–lacZ* fusion plasmid (see below) carrying the wild-type or mutated *glmU* region as template. The oligonucleotides used were RBP22 homologous to the *bla* gene of pRS415, and Lac22 homologous to the beginning of *lacZ*. The Lac22 oligo was labelled at its 5′ end with [γ-^{32}P]ATP and polynucleotide kinase prior to the PCR. The PCR-synthesized fragment was cut with *Eco*RI to remove the upstream sequences derived from the plasmid vector, pRS415, and the ~800 bp *glmU–lacZ* fragments (wild-type or carrying mutations) purified from an agarose gel. Hybridization, S1 digestion and analysis of the S1-resistant fragments by denaturing polyacrylamide gel electrophoresis were performed as described previously (Plumbridge, 1990).

Reverse transcriptase mapping

The primer used was the Glm3 oligonucleotide (Figure 2A) labelled at its 5′ end by [γ-^{32}P]ATP and polynucleotide kinase. It was used both for complete extension experiments with all four deoxynucleotides and for abortive extension in the presence of dGTP, dTTP, dCTP and dideoxyATP as described previously (Plumbridge *et al.*, 1993). The radioactivity present in the different bands of the gels was quantitated with a PhosphorImager (Molecular Dynamics).

DNase I footprinting

Initially the probe used was the Glm4–Glm3 PCR-generated fragment (Figure 2) with the Glm3 oligonucleotide 5′-end-labelled with [γ-^{32}P]ATP and polynucleotide kinase. For the footprinting of mutant *glmU* operators the probe was generated on the pRS415-derived plasmids as template using Glm3 (5′-end-labelled) and RBP22, homologous to the *bla* gene on pRS415 (because the Glm4 oligonucleotide region is not present on the fusions). The PCR product was digested with *Eco*RI and the 460 bp fragment purified on a native 5% acrylamide gel. Complex formation using an extract from a NagC-overproducing strain and DNase I digestion were carried out as described previously (Plumbridge and Kolb, 1991).

Construction of glmU–lacZ fusions

The 625 bp *Pvu*II fragment (Figure 2A) was inserted into the *Sma*I site of pRS415 (Simons *et al.*, 1987) to produce a fusion carrying P1 and P2. To construct a fusion carrying just P2, the 356 bp *Pvu*II–*Dra*I fragment was inserted into the same vector. The *glmU–lacZ* transcriptional fusions were transferred to bacteriophage λ as described, to give λRS/GUP1P2 and λRS/GUP2, and used to lysogenize bacteria IBPC5321R (wild-type) and IBPC529C (*nagC*). Monolysogens were identified and β-galactosidase activities measured on four aliquots of each culture during exponential growth as described previously (Miller, 1972; Plumbridge and Kolb, 1993). Activities are given in the units described by Miller (1972). Two or three independently constructed plasmids carrying the wild-type or mutated fusions were transferred to λRS45, and β-galactosidase activities of several lysogens were measured for each phage carrying a fusion.

Oligonucleotide-directed mutagenesis

The *glmU* insert of the *glmU–lacZ* fusions corresponding to the *Pvu*II–*Pvu*II or *Pvu*II–*Dra*I fragments were excised as *Eco*RI–*Bam*HI fragments from pRS/GUP1P2 and pRS/GUP2, using the sites of the pRS415 linker, and inserted into pT7T3U18 (Pharmacia). Oligonucleotide mutagenesis, by the method of Kunkel, was performed as described (Sambrook *et al.*, 1989). Oligonucleotides Glm6 and Glm7 were 26 nucleotides in length and homologous to the complementary strand of that shown in Figure 2. The Glm6 oligonucleotide changed the conserved AA at positions +5 and +6 of NagC operator, BoxG1, to GC. The Glm7 oligonucleotide produced the same change in BoxG2. Mutations were identified by sequencing and then used to construct *glmU–lacZ* fusions by inserting the *Eco*RI–*Bam*HI fragments carrying the mutations into pRS415. The presence of the mutations on the fusions was verified by sequencing. Asymmetric PCR was performed using lysates of the phages as a source of template DNA and oligonucleotides complementary to the beginning of *lacZ* and the *bla* gene region present in λRS45 (Lac22 and RBP22, see above) to amplify a fragment covering the *glmU* region cloned in the fusion as described (Plumbridge and Kolb, 1995). This predominately single-stranded product was then sequenced using suitable internal oligonucleotides.

Acknowledgements

I am very grateful to Annie Kolb and Henri Buc for their constant interest and critical reading of the manuscript. I thank Dominique Mengin-Lecreulx for communicating unpublished results and Steve Busby, Dominique Mengin-Lecreulx, Bernard Badet and Marianne Grunberg-Manago for useful discussions and comments on the manuscript. This work was supported by grants from the CNRS, INSERM, and the EEC

References

Black,P.N. and DiRusso,C.C. (1994) Molecular and biochemical analyses of fatty acid transport, metabolism and gene regulation in *Escherichia coli*. *Biochim. Biophys. Acta*, **1210**, 123–145.

Busby,S. and Ebright,R.H. (1994) Promoter structure, promoter recognition and transcription activation in prokaryotes. *Cell*, **79**, 743–746.

Busby,S., West,D., Lawes,M., Webster,C., Ishihama,A. and Kolb,A. (1994) Transcription activation by the *Escherichia coli* cyclic AMP receptor protein. Receptors bound in tandem at promoters can interact synergistically. *J. Mol. Biol.*, **241**, 341–352.

Chin,A.M., Feucht,B.U. and Saier,M.H. (1987) Evidence for regulation of gluconeogenesis by the fructose phosphotransferase system in *Salmonella typhimurium*. *J. Bacteriol.*, **169**, 897–899.

Claverie-Martin,F. and Magasanik,B. (1992) Positive effects of DNA bending on activation of transcription from a distant site. *J. Mol. Biol.*, **227**, 996–1008.

Collado-Vides,J., Magasanik,B. and Gralla,J.D. (1991) Control site location and transcriptional regulation in *Escherichia coli*. *Microbiol. Rev.*, **55**, 371–394.

Geerse,R.H., van der Pluijm,J. and Postma,P.W. (1989) The repressor of the PEP: fructose phosphotransferase system is required for the transcription of the *pps* gene of *Escherichia coli*. *Mol. Gen. Genet.*, **218**, 348–352.

Gralla,J.D. (1991) Transcriptional control—lessons from an *E.coli* promoter data base. *Cell*, **66**, 415–418.

Harley,C.B. and Reynolds,R.P. (1987) Analysis of *E.coli* promoter sequences. *Nucleic Acids Res.*, **15**, 2343–2361.

Hawley,D.K. and McClure,W.R. (1983) Compilation and analysis of *Escherichia coli* promoter sequences. *Nucleic Acids Res.*, **11**, 2237–2255.

He,B. and Zalkin,H. (1992) Repression of *Escherichia coli purB* is by a transcriptional roadblock mechanism. *J. Bacteriol.*, **174**, 7121–7127.

Henry,M.F. and Cronan,J.E. (1991a) *Escherichia coli* transcription factor that both activates fatty acid synthesis and represses fatty acid degradation. *J. Mol. Biol.*, **222**, 843–849.

Henry,M.F. and Cronan,J.E. (1991b) A new mechanism of transcriptional regulation: release of an activator triggered by small molecule binding. *Cell*, **70**, 671–679.

Hochschild,A. and Ptashne,M. (1986) Cooperative binding of lambda repressors to sites separated by integral turns of the DNA helix. *Cell*, **44**, 681–687.

Hoover,T.R., Santero,E., Porter,S. and Kustu,S. (1990) The integration host factor stimulates interaction of RNA polymerase with NifA, the transcriptional activator for nitrogen fixation operons. *Cell*, **63**, 11–22.

Ishihama,A. (1993) Protein–protein communication within the transcription apparatus. *J. Bacteriol.*, **175**, 2483–2489.

Joung,J.K., Linh,U.L. and Hochschild,A. (1993) Synergistic activation of transcription by the *Escherichia coli* cAMP receptor protein. *Proc. Natl Acad. Sci. USA*, **90**, 3083–3087.

Kuhnke,G., Theres,C., Fritz,H.-J. and Ehring,R. (1989) RNA polymerase and *gal* repressor bind simultaneously and with DNA bending to the control region of the *Escherichia coli* galactose operon. *EMBO J.*, **8**, 1247–1255.

Kustu,S., North,A.K. and Weiss,D.S. (1991) Prokaryotic transcriptional enhancers and enhancer-binding proteins. *Trends Biochem. Sci.*, **16**, 397–402.

Law,S.M., Bellomy,G.R., Schlax,P.J. and Record,M.T. (1993) *In vivo* thermodynamic analysis of repression with and without looping in *lac* constructs: estimates of free and local *lac* repressor concentrations and of physical properties of a region of supercoiled plasmid DNA *in vivo*. *J. Mol. Biol.*, **230**, 161–173.

Lee,D.-H. and Schleif,R.F. (1989) *In vivo* DNA loops in *araCBAD*: size limits and helical repeat. *Proc. Natl Acad. Sci. USA*, **86**, 476–480.

Mandal,N., Su,W., Haber,R., Adhya,S. and Echols,H. (1990) DNA looping in cellular repression of transcription of the galactose operon. *Genes Dev.*, **4**, 410–418.

Mengin-Lecreulx,D. and van Heijenoort,J. (1993) Identification of the *glmU* gene encoding *N*-acetyl glucosamine-1-phosphate uridyltransferase in *Escherichia coli*. *J. Bacteriol.*, **175**, 6150–6157.

Mengin-Lecreulx,D. and van Heijenoort,J. (1994) Copurification of glucosamine-1-phosphate acetyltransferase and *N*-acetyl-glucosamine-1-phosphate uridyltransferase activities of *Escherichia coli*: characterization of the *glmU* gene product as a bifunctional enzyme catalyzing two subsequent steps in the pathway for UDP-*N*-acetylglucosamine synthesis. *J. Bacteriol.*, **176**, 5788–5795.

Miller,J.H. (1972) *Experiments in Molecular Genetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Neidhardt,F.C., Bloch,P.L. and Smith,D.F. (1974) Culture medium for enterobacteria. *J. Bacteriol.*, **119**, 736–747.

Pittard,A.J. and Davidson,B.E. (1991) TyrR protein of *Escherichia coli* and its role as repressor and activator. *Mol. Microbiol.*, **5**, 1585–1592.

Plumbridge,J.A. (1990) Induction of the *nag* regulon of *Escherichia coli* by *N*-acetylglucosamine and glucosamine: role of the cAMP–catabolite activator protein complex in expression of the regulon. *J. Bacteriol.*, **172**, 2728–2735.

Plumbridge,J.A. (1991) Repression and induction of the *nag* regulon of *Escherichia coli* K12: the roles of *nagC* and *nagA* in maintenance of the uninduced state. *Mol. Microbiol.*, **5**, 2053–2062.

Plumbridge,J.A. and Kolb,A. (1991) CAP and Nag repressor binding to the regulatory regions of the *nagE–B* and *manX* genes of *E.coli*. *J. Mol. Biol.*, **217**, 661–679.

Plumbridge,J.A. and Kolb,A. (1993) DNA loop formation between Nag repressor molecules bound to its two operator sites is necessary for repression of the *nag* regulon of *Escherichia coli in vivo*. *Mol. Microbiol.*, **10**, 973–981.

Plumbridge,J.A. and Kolb,A. (1995) Nag repressor–operator interactions: protein–DNA contacts cover more than two turns of the DNA helix. *J. Mol. Biol.*, **249**, 890–902.

Plumbridge,J.A., Cochet,O., Souza,J.M., Altamirano,M.M., Calcagno, M.L. and Badet,B. (1993) Coordinated regulation of amino sugar-synthesizing and degrading enzymes in *Escherichia coli* K-12. *J. Bacteriol.*, **175**, 4951–4956.

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

Rabin,R.S., Collins,L.A. and Stewart,V. (1992) *In vivo* requirement of integration host factor for *nar* (nitrate reductase) operon expression in *Escherichia coli* K-12. *Proc. Natl Acad. Sci. USA*, **89**, 8701–8705.

Ramseier,T.M., Nègre,D., Cortay,J.-C., Scarabel,M., Cozzone,A.J. and Saier,M.H. (1993) *In vitro* binding of the pleiotropic transcriptional regulatory protein, FruR to the *fru, pps, ace, pts* and *icd* operons of *Escherichia coli* and *Salmonella typhimurium*. *J. Mol. Biol.*, **234**, 28–44.

Reeder,T. and Schleif,R. (1993) AraC can activate transcription from only one position and when pointed in only one direction. *J. Mol. Biol.*, **231**, 205–218.

Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schroder,I., Darie,S. and Gunsalas,R.P. (1993) Activation of the *Escherichia coli* nitrate reductase (*narGHJI*) operon by NarL and Fnr requires integration host factor. *J. Biol. Chem.*, **268**, 771–774.

Simons,R.W., Houman,F. and Kleckner,N. (1987) Improved single and multicopy *lac* based cloning vectors for protein and operon fusions. *Gene*, **53**, 85–96.

Stewart,V. (1993) Nitrate regulation of anaerobic respiratory gene expression in *Escherichia coli*. *Mol. Microbiol.*, **9**, 425–434.

Titgemeyer,F., Reizer,J., Reizer,A. and Saier,M.H. (1994) Evolutionary relationships between sugar kinases and transcriptional repressors in bacteria. *Microbiology*, **140**, 2349–2354.

Vogler,A.P. and Lengeler,J.W. (1989) Analysis of the *nag* regulon from *Escherichia coli* K12 and *Klebsiella pneumoniae* and of its regulation. *Mol. Gen. Genet.*, **219**, 97–105.

Walker,J.E., Gay,N.J., Saraste,M. and Eberle,A.N. (1984) DNA sequence around the *Escherichia coli unc* operon. Completion of the sequence of a 17 kilobase segment containing *asnA, oriC, unc, glmS* and *phoS*. *Biochem. J.*, **224**, 799–815.

White,R.J. (1968) Control of aminosugar metabolism in *Escherichia coli* and isolation of mutants unable to degrade amino sugars. *Biochem. J.*, **106**, 847–858.

Received on March 16, 1995; revised on May 25, 1995