JOURNAL OF BACTERIOLOGY, May 1990, p. 2728–2735
0021-9193/90/052728-08$02.00/0

Vol. 172, No. 5

Induction of the *nag* Regulon of *Escherichia coli by N*-Acetylglucosamine and Glucosamine: Role of the Cyclic AMP-Catabolite Activator Protein Complex in Expression of the Regulon

JACQUELINE A. PLUMBRIDGE

Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, 75005 Paris, France

Received 18 October 1989/Accepted 30 January 1990

The divergent *nag* regulon located at 15.5 min on the *Escherichia coli* map encodes genes necessary for growth on *N*-acetylglucosamine and glucosamine. Full induction of the regulon requires both the presence of *N*-acetylglucosamine and a functional cyclic AMP (cAMP)-catabolite activator protein (CAP) complex. Glucosamine produces a lower level of induction of the regulon. A nearly symmetric consensus CAP-binding site is located in the intergenic region between *nagE* (encoding EII^{Nag}) and *nagB* (encoding glucosamine-6-phosphate deaminase). Expression of both *nagE* and *nagB* genes is stimulated by cAMP-CAP, but the effect is more pronounced for *nagE*. In fact, very little expression of *nagE* is observed in the absence of cAMP-CAP, whereas 50% maximum expression of *nagB* is observed with *N*-acetylglucosamine in the absence of cAMP-CAP. Two mRNA 5′ ends separated by about 100 nucleotides were located before *nagB*, and both seem to be similarly subject to *N*-acetylglucosamine induction and cAMP-CAP stimulation. To induce the regulon, *N*-acetylglucosamine or glucosamine must enter the cell, but the particular transport mechanism used is not important.

Escherichia coli is capable of growth on a variety of sugars as a source of carbon. Several of these sugars, notably glucose and related hexoses, hexitols, and hexosamines, are part of a common uptake pathway, the phosphotransferase system (PTS), in which a phosphate group originating from phosphoenolpyruvate is transferred by a series of soluble and membrane proteins to the sugar, which is concomitantly phosphorylated as it enters the cell (for reviews, see references 32 and 33).

Both *N*-acetylglucosamine (GlcNAc) and glucosamine (GlcN) are PTS carbohydrates in *E. coli* (4, 49). Genes for both the uptake (14) and metabolism (13, 48) of GlcNAc have been located at 15.5 min on the *E. coli* chromosome (1). So far, five genes of the *nag* regulon have been identified in this region (30). The genes identified include *nagE* encoding the GlcNAc-specific EII of the PTS (EII^{Nag}), which is transcribed in the clockwise direction (28, 36), and four other genes, *nagBACD*, which are transcribed in the counterclockwise direction. The *nag* genes are thus distributed in the two arms of a divergent regulon. *nagB* encodes GlcN-6-phosphate deaminase, and *nagA* encodes GlcNAc-6-phosphate deacetylase (48). The functions of *nagC* and *nagD* are not well characterized but are included in the *nag* regulon because they are expressed from the same mRNA which encodes *nagB* and *nagA*. *nagC* probably encodes a repressor for the *nag* regulon (30, 44).

DNA sequence analysis and S1 mapping showed that *nagB* and *nagE* are transcribed from opposite strands of DNA and that their translational start points are separated by about 330 nucleotides (36). Near the middle of this region, a sequence very similar to the consensus catabolite activator protein (CAP)-binding site (6) was noted.

The presence of this consensus CAP sequence suggested that the *nag* regulon is controlled by the cyclic AMP (cAMP)-CAP complex and that expression of the regulon is sensitive to the presence of glucose in the medium. Other components of the PTS have been shown to be sensitive to cAMP-CAP control (35, 38). Consensus CAP-binding sites have been identified in both the mannitol (5) and glucitol (51) operons and also in the regulatory region of the *ptsHI crr* operon (7).

The effect of glucose on the utilization of other sugars can be conveniently divided into two categories, catabolite repression and inducer exclusion (43). Catabolite repression is the result of a reduction in the concentration of the cAMP-CAP complex. It is manifested by a level of expression of catabolite-sensitive genes which is higher during growth on glycerol than on glucose. On the other hand, inducer exclusion, at least in the cases studied so far (22, 25, 26, 39), has been shown to be due to the inhibitory action of the nonphosphorylated form of $EIII^{Glc}$ (product of the *crr* gene) on the transport or synthesis of the intracellular inducer.

The organization of the genes for uptake and metabolism of GlcNAc in two divergent operons is reminiscent of non-PTS sugar regulons like maltose and arabinose where functionally related genes occur in separate operons; e.g., the *malPQ* operon is necessary for maltose degradation and the *malK lamB* operon is involved in maltose uptake. For the *malA* region, a single CAP-binding site exists in the intergenic region between *malT* and *malPQ*. Although cAMP-CAP affects the expression of both genes, the effect on *malPQ* is indirect via the level of the MalT protein, the positive regulator (3, 12). In the *malB* and *ara* regions, both arms of the divergent operons seem to be directly affected by cAMP-CAP but in a complex fashion. Raibaud et al. (34) have examined the binding of both CAP and MalT to the 270-base-pair intergenic region of *malB* and produced a model in which the DNA bound in phase around a core of several CAP and MalT proteins results in activation of both divergently arranged promoters. For the *ara* regulon, although footprinting suggested two CAP-binding sites (27), in vivo experiments suggest that just one site is involved in stimulating *araC* and *araBAD* expression (9, 16, 17, 23, 42).

In this report, the effect of the cAMP-CAP complex on the expression of the divergent *nag* regulon is studied. Protein fusions between *nagE-lacZ* and *nagB-lacZ* were used to

TABLE 1. Bacterial strains

Strain	Relevant genotype	Reference or origin
IBPC 5321	F⁻ *thi-1 argG6 argE3 his-4 mtl-1 xyl-5 tsx-29 rpsL ΔlacX74*	31
IBPC 461	F⁻ *galK2 Δ(his-gnd)79 mgl-50 ptsF10 thyA111 galP64 ptsM8 zdj-225*::Tn*10*	29
IBPC 463	F⁻ *thr-1 leuB6 tonA21 lacY1 supE44 thi-1 hsdR514 nagE467 ptsM8 zdj-225*::Tn*10*	29
IBPC 463C	F⁻ *thr-1 leuB6 tonA21 lacY1 supE44 thi-1 hsdR514 nagE467 ptsM8*	IBPC 463 cured of Tn*10*
IBPC 466	IBPC 463C *zbf-507*::Tn*10*	P1 (405) × IBPC 463C
405	F⁻ *ara-14 lacY1 tsx-57 supE44 glnS1 galK2 rpsL71 metC56 xyl-5* or *xyl-7 thi-1 tfr-5 zbf-507*::Tn*10*	31
IBPC 566	IBPC 5321 *ptsM8 zdj-225*::Tn*10*	P1 (IBPC 461) × IBPC 5321
IBPC 567	IBPC 5321 *ptsM8*	IBPC 566 cured of Tn*10*
IBPC 569	IBPC 5321 *pstM8 nagE467 zbf-507*::Tn*10*	P1 (IBPC 466) × IBPC 567
IBPC 579	IBPC 5321 *ptsM8 nagE467*	IBPC 569 cured of Tn*10*
DC369	*zdj-225*::Tn*10 fadR16 butD12 mel1 supF58*	29
IBPC 581	IBPC 5321 *nagE467 zdj-225*::Tn*10*	P1 (DC369) × IBPC 579
TP2506	F⁻ *lacZ82* or *lacZ827 glk-7 rha-4 rpsL223 ptsG22 zcf-229*::Tn*10*	7
IBPC 522	IBPC 5321 *ptsM8 nagE467 ptsG22 zcf-229*::Tn*10*	P1 (TP2506) × IBPC 579
TP2006	F⁻ *ΔlacX74 Δcya xyl glp-8306*	37
TP2100	F⁻ *ΔlacX74 xyl argH1 ilvA*	37

quantitate the induction of the regulon by GlcNAc and GlcN and the effect of cAMP-CAP. The results with the fusions are correlated with the mRNA levels measured by S1 mapping of the *nagE* and *nagB* transcripts.

MATERIALS AND METHODS

Bacteriological methods and strain constructions. The minimal MOPS (morpholinopropanesulfonic acid) medium of Neidhardt et al. (24) was used to culture bacteria for measurement of β-galactosidase activity. It was supplemented with amino acids (50 μg/ml) as required. β-Galactosidase activities were measured as described by Miller (21). A set of isogenic Δ*lac* strains carrying the different *nag* mutations was constructed by P1 transduction with *zbf-507*::Tn*10*, which is about 40% cotransducible with the *nag* region. GlcNAc enters the *E. coli* cell not only by EII^{Nag} encoded by *nagE* but also by the more general hexose transporter $EIIM,P/III^{Man}$ encoded by the genes *manXYZ* (10, 40) located at 40 min on the *E. coli* chromosome (1). This locus was previously called *ptsM*, and this term will be used for convenience in this report. The $EIIM,P/III^{Man}$ complex transports several sugars including mannose, glucose, 2-deoxyglucose, and GlcNAc and is the principal transporter for GlcN (4, 14). The *ptsM8* allele was introduced into IBPC 5321 with *zdj-225*::Tn*10* from IBPC 461 (29), and screening was performed for those colonies showing reduced growth on mannose as the sole carbon source. Various candidate bacteria were cured of the Tn*10* (by using chlortetracycline and fusaric acid [19]), and the *nagE467* mutation was introduced from JM1179 via *zbf-507*::Tn*10*. The Nag^- colonies resulting from this cross were presumed to contain both *ptsM* and *nagE* alleles. The original *ptsM* strain was called IBPC 566, whereas the double *ptsM nagE* mutant strain was called IBPC 569. After the Tn*10* was cured from IBPC 569 to give IBPC 579, a strain carrying just the *nagE* mutation (IBPC 581) was constructed by using P1 grown on a *zdj-225*::Tn*10* strain to select for Tc^r bacteria which were now Nag^+ (i.e., $ptsM^+$). *ptsG* encodes EII^{Glc}, which transports glucose and also mannose (4). The *ptsG22* allele was introduced via *zcf-229*::Tn*10* from TP2506 into the *ptsM* strain, screening for those colonies which were now really Man^-. The genotypes of these and other strains used are given in Table 1. The Δ*cya* strain, TP2006, was grown in LB medium containing maltose, Mg^{2+}, and cAMP (1 mM) to permit induction of the λ receptor before lysogenization with the λ bacteriophage carrying the *nag-lacZ* fusions. The TP2006 strain carries the *glp-8306* mutation, previously called *glp**, which maps at 88 min and allows growth on glycerol in the absence of cAMP (15, 37). The strain grows on glucose because *ptsG*, the major glucose transporter, is only partly dependent on cAMP-CAP.

Construction of *nagE* and *nagB* protein fusions with *lacZ*. To construct a *nagB-lacZ* fusion, the 1.2-kilobase *Pvu*II fragment of pB31-1 (29) was cloned into the *Sma*I site of pMC1403 (2). This fragment contains the 3′ end of *nagE*, all the intercistronic region, and about three-fourths of *nagB*. The *nagB* and *lacZ* protein-coding frames are out of phase at the *Pvu*II-*Sma*I junction, and correct plasmids were found by restriction mapping of the colonies that were very pale blue after 2 to 3 days. Inspection of the DNA sequence showed that a +1 phase change was needed to align the *nagB-lacZ* reading frames, and this was achieved by filling in the *Bam*HI site of pMC1403 to give pMC/PvB1.

An equivalent *nagE-lacZ* fusion was created by cloning the *Cla*I-*Hpa*I fragment of pB31-1 (as an *Eco*RI-*Hpa*I fragment of pB35-25 [36]) into pMC1403. Again, the *nagE-lacZ* fusion was out of phase and corrected by filling in the *Bam*HI site to give pMC/EHpB14. This plasmid carries the entire *nagB* gene. The fusions created are shown in Fig. 1. DNA manipulations were performed by standard procedures (20).

Both fusions were transferred to the bacteriophage λ *b2* region as *Eco*RI-to-*Sac*II fragments to give λ *c*I857 thermosensitive phages. They were used to lysogenize the different bacterial strains at a low multiplicity of infection as described previously (31). Several independent lysogens were tested for β-galactosidase activity levels to identify monolysogens. Multiple lysogens could be detected since they exhibited β-galactosidase levels two or three times the monolysogen value.

mRNA preparation and S1 mapping. mRNA was prepared from exponentially growing cultures by the hot-phenol method. Cultures were grown in minimal MOPS medium supplemented with auxotrophic amino acids (50 μg/ml). Carbon sources were at 0.2% except glycerol, which was 0.4%.

The probe used to map *nagE* and *nagB* transcripts was the 1.2-kilobase *Pvu*II fragment covering the intergenic region as

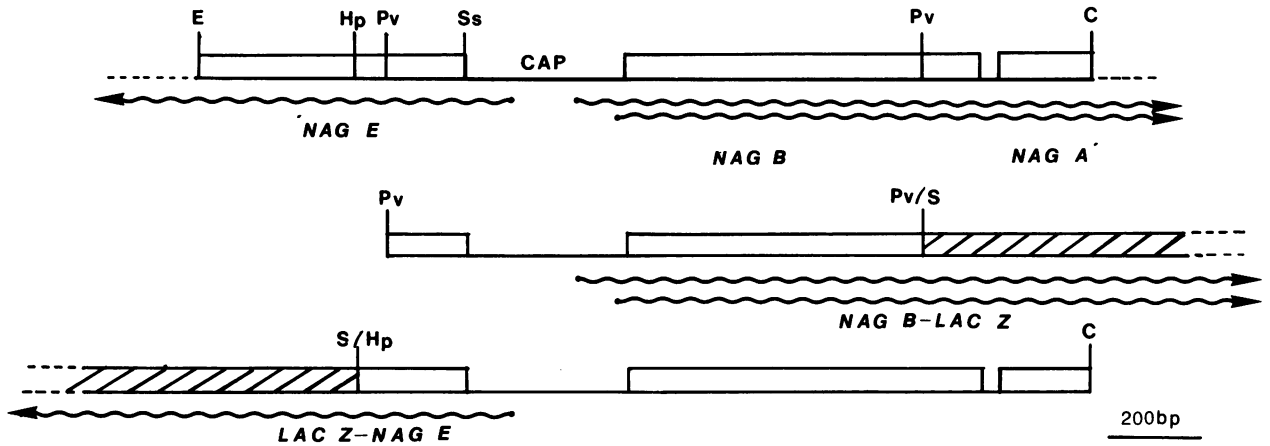

FIG. 1. Structure of the *nagE-nagB* regulatory region and construction of *nagE-lacZ* and *nagB-lacZ* protein fusions. A restriction map of a 2-kilobase region covering the 5′ end of the *nagE* gene, the *nagB* gene, and the 5′ end of the *nagA* gene is shown. The positions of the mapped transcripts are shown by wavy lines. The location of the consensus CAP-binding site sequence is marked. The structures of the *nagB-lacZ* and *nagE-lacZ* fusions are shown underneath. *nag* gene open reading frames are shown by open boxes, while *lacZ* DNA is indicated by diagonal shading. Restriction sites are identified as follows: E, *Eco*RI; Hp, *Hpa*I; Pv, *Pvu*I; Ss, *Ssp*I; C, *Cla*I; Pv/S, hybrid *Pvu*II-*Sma*I junction; S/Hp, hybrid *Sma*I-*Hpa*I junction. bp, Base pairs.

used previously (36). S1 mapping was performed as described previously (36).

RESULTS

Expression of *nagE* and *nagB* is subject to catabolite repression and is induced by growth on GlcNAc or GlcN. DNA fragments carrying the *nagE-nagB* intergenic region and part of the *nagE* and *nagB* genes were inserted before *lacZ* to produce protein fusions *nagE-lacZ* and *nagB-lacZ* as described in Materials and Methods and shown in Fig. 1. The two protein fusions were transferred from plasmids to bacteriophage λ and used to lysogenize a Δ*lac* strain wild type for *nag* (IBPC 5321). The level of expression of the two fusions on various carbon sources (glucose, glycerol, gluconate, GlcN, and GlcNAc) was measured (Table 2, top).

TABLE 2. Effect of growth on different carbon sources on induction of the *nagE-lacZ* and *nagB-lacZ* fusions

Carbon source[a]	β-Galactosidase activity[b]	
	nagE-lacZ	*nagB-lacZ*
GlcNAc	471 (760)	1,190 (1,470)
GlcN	118 (190)	294 (363)
Glycerol	62 (100)	81 (100)
Gluconate	54 (87)	67 (53)
Glucose	20 (32)	55 (68)
GlcNAc + glycerol	458 (739)	1,090 (1,346)
GlcNAc + gluconate	169 (272)	567 (700)
GlcNAc + glucose	136 (219)	629 (777)
GlcN + glycerol	104 (167)	215 (265)
GlcN + gluconate	65 (105)	145 (179)
GlcN + glucose	28 (45)	61 (75)

[a] Strains were grown in minimal MOPS medium containing 50 μg of arginine per ml and 50 μg of histidine per ml at 30°C with the carbon sources indicated. All carbon sources were at 0.2% except glycerol, which was 0.4%.

[b] β-Galactosidase activities (in Miller units [21]) were measured in IBPC 5321 lysogenized with either λ *nagE-lacZ* or λ *nagB-lacZ*. Several independent lysogens were tested to identify monolysogens. The numbers in parentheses are the percent value of the activity compared with the value in glycerol.

Considering first *nagE*, the lowest level of expression observed was during growth on glucose (20 U), whereas with glycerol, three times more expression was detected. Growth on GlcN or GlcNAc produced even higher levels, two and seven times the glycerol value (Table 2, top). Qualitatively similar results were seen with the *nagB* fusion, but the difference between glucose and glycerol was less. Only a 1.5-fold enhancement in β-galactosidase activity was observed for glycerol compared with glucose. For both *nagE* and *nagB* fusions, the maximum induction for glucose compared with GlcNAc was 20-fold. GlcN produced about a fivefold induction compared with glucose (Table 2, top). These results show that both arms of the regulon are induced in parallel by GlcNAc or GlcN and that both directions are subject to catabolite repression, although the effect is weaker for *nagB* than for *nagE*.

To further investigate the effect of glucose on the *nag* regulon, expression of the fusions was measured on binary mixtures of sugars (Table 2, bottom). For both *nagB* and *nagE*, the presence of glycerol did not affect induction by either GlcNAc or GlcN, but the presence of gluconate or glucose reduced the level of induction. Glucose effectively eliminated all induction by GlcN but allowed appreciable induction by GlcNAc (30 to 50% of the values observed with GlcNAc alone). Gluconate, which also exerts an effect of catabolite repression, produced an intermediate response. Mixed with GlcNAc, the result was activity levels similar to those with glucose plus GlcNAc. However mixed with GlcN, gluconate was not like glucose but allowed an intermediate level of induction by GlcN. These results show that there are fundamental differences in the induction of the regulon by GlcNAc and GlcN and in their sensitivity to catabolite repression (see Discussion).

S1 mapping experiments of the *nagE* and *nagB* transcripts present in mRNA of cells grown on the different carbon sources show that the mRNA levels increase with the β-galactosidase activities (36; data not shown). Quantitation of the S1 gels and Northern (RNA) blots (30) suggests that the induction factor for mRNA is actually greater than the 20-fold observed here, which implies some posttranscriptional regulation.

TABLE 3. cAMP-CAP control of *nagE* and *nagB* expression

Medium[a]	β-Galactosidase activity[b]			
	nagE-lacZ		*nagB-lacZ*	
	Δ*cya*	Wild type	Δ*cya*	Wild type
Glucose	18 (17)	16 (15)	104 (59)	80 (45)
Glycerol	13 (12)	104 (100)	85 (48)	176 (100)
Glycerol + cAMP	52 (50)	104 (100)	115 (65)	192 (110)
Glycerol + GlcNAc	25 (24)	377 (360)	340 (190)	1,187 (670)
Glycerol + cAMP + GlcNAc	344 (330)	637 (610)	940 (530)	1,400 (800)

[a] Strains were grown in minimal MOPS medium supplemented with arginine, isoleucine, and valine (50 μg/ml) at 30°C with the carbon sources indicated. cAMP when added was 2 mM.

[b] β-Galactosidase activities (in Miller units [21]) were measured in isogenic wild-type and Δ*cya* strains carrying the λ *nagE-lacZ* or λ *nagB-lacZ* phage. The numbers in parentheses are the percent value of the activity compared with that of the wild-type strain grown in glycerol. The absolute values of the fusions are different from Table 2, presumably owing to differences in genetic background.

Induction of deaminase (*nagB*) and deacetylase (*nagA*) enzyme levels after growth on GlcNAc and GlcN have been measured by White (48) and recently by Vögler and Lengeler (44). The values of White (48), 7- to 8-fold induction by GlcNAc, are somewhat lower than the results reported here for the *nagB-lacZ* fusion, while those of Vögler and Lengeler (44), a 50-fold induction, are somewhat higher. Both found GlcN to be less efficient. Vögler and Lengeler (44) also measured GlcNAc transport; they found a high basal level in glycerol-grown cells which increased just threefold after induction with GlcNAc, compared with the sevenfold found here for the *nagE-lacZ* fusion.

Expression of *nagE-lacZ* is strongly dependent on the cAMP-CAP complex. To confirm that the different levels of expression of the fusions in glucose and glycerol medium are due to an effect of the cAMP-CAP complex, a Δ*cya* strain was used. The λ bacteriophage carrying the *nagE-lacZ* and *nagB-lacZ* fusions were used to lysogenize the Δ*cya*, Δ*lac* strain TP2006, and the isogenic parent TP2100 (*cya*$^+$ Δ*lac*). In the wild-type strain TP2100, glucose repressed the activity of both fusions compared with that observed during growth on glycerol.

In the Δ*cya* strain, levels of expression of both *nagE-lacZ* and *nagB-lacZ* fusions had the same low value in glucose and glycerol, and this value was identical to that observed for the wild-type strain grown in glucose. For *nagE-lacZ*, this value was only about 15% that of the wild-type strain grown in glycerol, whereas for *nagB-lacZ*, it was 50%. Addition of cAMP to the Δ*cya nagE-lacZ* fusion strain produced a fourfold increase in β-galactosidase activity but only a small increase in the *nagB-lacZ* fusion activity (Table 3).

Induction by GlcNAc of the two *nag-lacZ* fusions was studied in the Δ*cya* background and was found to be different for the two fusions (Table 3). In the absence of cAMP, GlcNAc produced a significant increase in *nagB-lacZ* expression (two times the level of the wild-type strain in glycerol). GlcNAc produced a small increase in *nagE-lacZ* expression in the absence of cAMP, but it corresponded to only 25% the level of expression of the wild-type strain in glycerol. The simultaneous addition of cAMP and GlcNAc produced a large induction for *nagE-lacZ* (15 times the value for GlcNAc alone). The simultaneous addition of GlcNAc and cAMP also induced the *nagB-lacZ* fusion but by only a factor of 2.5 compared with GlcNAc alone.

Quantitation of this experiment was rather difficult. Externally added cAMP actually increased expression of both *nagB-lacZ* and *nagE-lacZ* fusions in the wild-type strain in the presence of GlcNAc. (This could be interpreted as suggesting that GlcNAc actually reduces the level of cAMP-CAP in a manner analogous to glucose or that exogenously added cAMP increases the internal cAMP concentration in wild-type cells.) The maximum induction level observed with the Δ*cya* strain was lower than that observed with the wild-type background.

Qualitatively, the following conclusions can be drawn. (i) cAMP-CAP stimulates expression of both *nagE* and *nagB*. (ii) Expression of the *nagE-lacZ* fusion is very low in the absence of cAMP. The simultaneous addition of cAMP and GlcNAc produced a 20-fold increase in expression, whereas each substance individually produced only a 2- to 4-fold enhancement. (iii) Expression of *nagB-lacZ* is less dependent on cAMP, with significant induction being observed with GlcNAc alone, but maximal induction requires both GlcNAc and cAMP. (iv) Although induction of the *nagE-lacZ* fusion in the Δ*cya* strain is absolutely dependent on cAMP, growth of the wild-type strain on glucose plus GlcNAc still allows appreciable induction (219% [Table 2, bottom]). The *nagE* gene apparently can make very efficient use of the basal level of cAMP available in the presence of glucose.

mRNA prepared from glycerol-grown TP2006 (Δ*cya*) cells contained no detectable *nagE* or *nagB* mRNA (at normal exposures of the autoradiograms) (Fig. 2, lanes 1 to 3). Addition of cAMP to the growth medium produced a visible increase in *nagE* mRNA and not much change in *nagB* transcripts (lanes 4 to 6). The presence of GlcNAc in the growth medium produced some increase in *nagE* transcripts (although less than cAMP) and a considerable increase in both long and short *nagB* transcripts (lanes 7 to 9). The combined presence of cAMP and GlcNAc produced a small increase in both *nagB* transcripts compared with GlcNAc alone and a much greater increase in *nagE* transcripts (lanes 10 to 12). Thus, the two *nagB* transcripts are not differentially activated by cAMP-CAP and GlcNAc.

These results correlate with the *lacZ* fusion studies described above and show in addition that there was a direct effect of GlcNAc on the presence of the two *nagB* transcripts in the absence of cAMP. This result eliminates the otherwise attractive hypothesis to explain the partial dependence of *nagB* expression on cAMP-CAP—that one of the 5′ ends located in front of *nagB* is induced by GlcNAc in the absence of the cAMP-CAP, while the other is dependent on the cAMP-CAP complex for its expression. (At the moment, there is no evidence that the shorter *nagB* transcript is the result of a new transcription initiation event and not the result of processing of the longer transcript.)

Induction of the *nag* regulon requires internalization of GlcNAc or GlcN. Mutations affecting the transport of GlcN and GlcNAc were introduced into the IBPC 5321 background, and the strains were lysogenized with λ *nagE-lacZ* and λ *nagB-lacZ*. GlcNAc can enter via either the transport system encoded by *nagE* or the *ptsM* locus encoding the nonspecific hexose transporter EIIM,P/III^{Man}. The presence of either single mutation did not affect induction, but the double mutation eliminated growth on GlcNAc and prevented induction when the strain was grown on a mixture of GlcNAc and glucose or glycerol (Table 4).

The *ptsM* mutation severely reduced growth on GlcN (doubling time approximately 300 min compared with 100 min for the wild-type strain), but there was still induction of

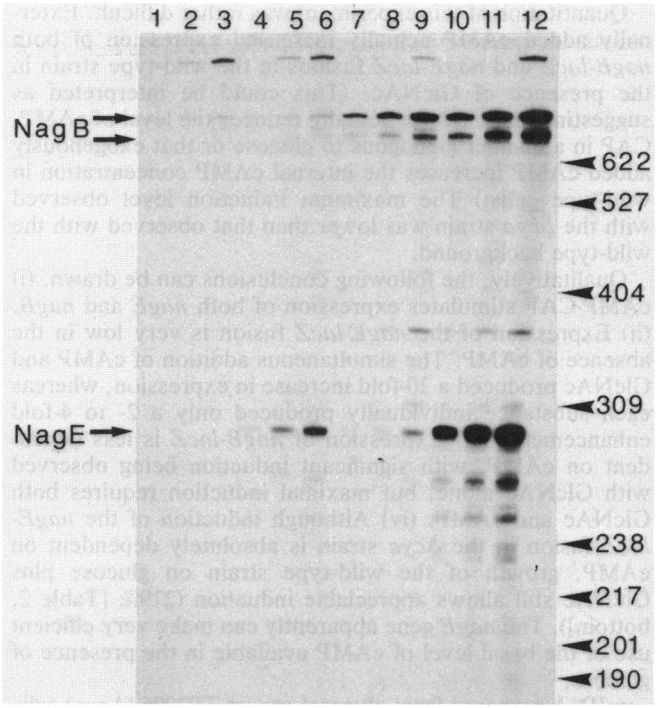

FIG. 2. Effect of cAMP and GlcNAc on induction of *nagE* and *nagB* transcripts in the Δ*cya* strain TP2006. Total RNA was prepared from exponentially growing cultures of TP2006 grown in minimal MOPS medium with 50 μg of arginine, isoleucine, and valine per ml, 0.4% glycerol, and the supplements stated below. The probe used was the 1.2-kilobase *Pvu*II fragment shown in Fig. 1. The probe and RNA (plus tRNA to produce a constant amount of RNA in each reaction) were hybridized overnight at 53.5°C, and the S1-resistant hybrids were analyzed on a 1-mm-thick 5% denaturing polyacrylamide gel. The two *nagB* transcripts and single *nagE* transcript as identified previously (36) are indicated. Lanes: 1 to 3, 10, 20, and 40 μg of RNA, respectively (no supplement); 4 to 6, 10, 20, and 40 μg of RNA, respectively, from a culture grown with 2 mM cAMP; 7 to 9, 5, 10, and 20 μg of RNA, respectively, from a culture grown with 0.2% GlcNAc; 10 to 12, 5, 10, and 20 μg of RNA, respectively, from a culture grown with 0.2% GlcNAc and 2 mM cAMP. Numbers on right show size in nucleotides.

both operons. The introduction of a *ptsG* mutation (in EII^{Glc}) eliminated the residual growth on GlcN and also any induction of the two operons during growth on a mixture of GlcN and glycerol (Table 4). These results show that preventing transport of the sugar substrate GlcNAc or GlcN eliminates induction of both operons and also that induction does not depend on the transporter used. GlcNAc entering via $EIIM,P/III^{Man}$ or EII^{Nag} induced the *nag* regulon about equally, at least at the high concentrations of GlcNAc employed here.

TABLE 5. Induction of the *nag* regulon by GlcN-6-phosphate

Strain	Genotype	β-Galactosidase activity[a]			
		nagE-lacZ		*nagB-lacZ*	
		GlcN-6-phosphate	Glc-6-phosphate	GlcN-6-phosphate	Glc-6-phosphate
IBPC 5321	Wild type	46.4	8.1	380	60
IBPC 522	*ptsM nagE ptsG*	23	5.5	212	41

[a] β-Galactosidase activities (in Miller units [21]) of λ *nagE-lacZ* and λ *nagB-lacZ* lysogens of the strains indicated were measured. Bacteria were grown at 30°C in minimal MOPS medium supplemented with 50 μg of arginine and histidine per ml and 0.2% GlcN-6-phosphate or glucose (Glc)-6-phosphate. Doubling times were 90 to 120 min with GlcN-6-phosphate and 60 to 65 min with glucose-6-phosphate.

Since induction of the *nag* regulon requires productive transport of GlcNAc or GlcN but is not dependent on any particular transporter, it implies that the inducing signal is detected internally. The common intermediate in the metabolism of GlcN and GlcNAc is GlcN-6-phosphate. The effect of exogenously added GlcN-6-phosphate was tested as an inducer of the *nag* regulon. Hexose phosphates (primarily glucose-6-phosphate) are taken up by the *uhp* system, a non-PTS sugar permease (11 and references therein). GlcN-6-phosphate is also a substrate for the permease (8). The triple mutant *ptsM nagE ptsG* (IBPC 522), which cannot grow on GlcNAc or GlcN, was used to eliminate any effect by contaminating GlcN. Growth of this strain on GlcN-6-phosphate (with or without low concentrations of glucose-6-phosphate which should induce the *uhp* operon [8]) produced growth rates comparable to those on GlcN and a level of induction also comparable to growth on GlcN (Table 5).

These results show that intracellular GlcN-6-phosphate is sufficient to induce the *nag* regulon. However, the magnitude of the induction observed was quite low and much less than that observed by growth on GlcNAc, although this could be due to GlcN-6-phosphate being a poor substrate for the *uhp* transporter or to externally added GlcNAc-6-phosphate being rapidly degraded intracellularly.

TABLE 4. Effect of mutations affecting GlcN and GlcNAc uptake on induction of the *nagE-lacZ* and *nagB-lacZ* fusions

Strain	Genotype	β-Galactosidase activity[a]									
		nagE-lacZ					*nagB-lacZ*				
		GlcNAc	GlcNAc and glycerol	GlcN	GlcN and glycerol	Glycerol	GlcNAc and glycerol	GlcNAc	GlcN	GlcN and glycerol	Glycerol
IBPC 5321	Wild type	471	458	118	103	62	1,187	1,080	294	215	81
IBPC 581	*nagE*	696	610	87	61	30	736	547	130	80	64
IBPC 566	*ptsM*	434	410	141	51	35	830	750	132	86	72
IBPC 569	*ptsM nagE*	NG[b]	42	114	45	30	NG	72	150	67	64
IBPC 522	*ptsM nagE ptsG*	NG	26	NG	28	28	NG	62	NG	51	75

[a] β-Galactosidase activities (in Miller units [21]) of the λ *nagE-lacZ* and λ *nagB-lacZ* lysogens of the strains indicated were measured. Several independent lysogens were tested to determine monolysogens. Bacteria were grown at 30°C in MOPS minimal medium supplemented with 50 μg of arginine and histidine per ml and the sugars indicated.

[b] NG, No growth.

```
CAP SITE 1
nagE  TTTGTGAGTTTTGTCACCAAA  nagB
      AAACACTCAAAACAGTGGTTT
CAP SITE 2
      AATTTAATTCGTATCGCAAAT
      TTAAATTAAGCATAGCGTTTA
consensus
      ANTGTGANNTNNNNCANATTN
```

FIG. 3. *nag* regulon CAP-binding sites compared with the consensus CAP-binding site. CAP site 1 corresponds to nucleotides 1108 to 1208, and CAP site 2 corresponds to nucleotides 1222 to 1242 of the sequence of Rogers et al. (36). The consensus site sequence is taken from de Crombrugghe et al. (6).

DISCUSSION

The experiments described in this report investigated the conditions required for induction of the divergent *nag* regulon of *E. coli*. Two basic criteria were defined: the presence of an inducing substrate which can enter the cell and a functional cAMP-CAP complex.

A sequence showing strong homology with the consensus CAP site (6) located near the center of the intergenic region is the presumed site of action of the cAMP-CAP complex. This CAP site is almost symmetric; the highly conserved TGTGA motif is present on the antisense strand proximal to *nagE*, while the sequence TGGTGA (i.e., the consensus with one extra G) is found on the *nagB* side (Fig. 3). De Crombrugghe et al. (6) have noted that a semipalindromic character is found in other CAP-binding sites and that the orientation of the binding site relative to the transcription start point is variable; e.g., for *lac*, the TGTGA motif is found on the sense strand, whereas for *galE* it is on the antisense strand.

The experiments described here show that there is an effect of cAMP-CAP on the expression of both arms of the regulon but that the two arms are not equally affected. Fully induced expression of the *nagE-lacZ* fusion in the Δ*cya* strain required the simultaneous presence of cAMP and GlcNAc. Each affector alone produced only a slight enhancement, whereas together they caused a 20-fold enhancement in β-galactosidase activity. For *nagB-lacZ*, there was much less dependence on cAMP. GlcNAc alone could produce nearly 50% maximum expression. Analysis of the 5′ ends of the *nagB* transcripts present during growth on GlcNAc in the absence of cAMP showed that both the long and short *nagB* transcripts, differing in length by about 100 nucleotides, were equally induced. Remember that the possibility exists that the shorter *nagB* transcript is derived from the larger transcript by, e.g., nucleolytic processing. The distance between the first *nagB* and the *nagE* transcription start points is about 130 base pairs, with the CAP box located near the middle. The strong effect observed on the *nagE* gene might be a result of the absolutely conserved consensus TGTGA motif on its side of the CAP box, whereas the weaker effect observed for *nagB* expression could be the result of the less conserved consensus TGGTGA motif. In addition, a second less conserved CAP consensus sequence can be identified in the DNA sequence on the *nagE* side of the primary site (Fig. 3), and both sites have been shown to bind cAMP-CAP in vitro (J. Plumbridge and A. Kolb, manuscript in preparation). These motifs are the obvious target for site-directed mutagenesis to try to define the molecular origin of the asymmetric induction. CAP binds to DNA as a dimer, but the two subunits do not have the same conformation when complexed with cAMP in a crystal (46). Binding of CAP to DNA has been shown to bend the DNA (18, 50), and model-building studies suggest that cAMP-CAP dimers severely distort the DNA, producing a molecule bent through 100 to 150° (45). In the *lac* system (18), the bending produced by CAP binding was greater on one side than the other, showing that the binding and, by implication, the activation can be asymmetric.

The two sugars GlcNAc and GlcN have been shown to induce both arms of the regulon. NagE encodes EII^{Nag}, the GlcNAc-specific transporter of the PTS. GlcN is not a substrate for this transporter but is primarily transported by the complex $EIIM,P/III^{Man}$ encoded by the genes of the *ptsM* locus (14). Transport of both sugars is accompanied by phosphorylation at the 6 position of the sugar ring. GlcN is subsequently metabolized by GlcN-6-phosphate deaminase, the product of the *nagB* gene. Thus, it is necessary for GlcN to induce the *nagB* gene but not the *nagE* gene. This seemingly unnecessary induction of *nagE* could be explained if induction of both arms can only occur in parallel by the action of a common inducer.

GlcNAc, after entering the cell as GlcNAc-6-phosphate, is deacetylated by the action of the *nagA* gene product, GlcNAc-6-phosphate deacetylase, to give GlcN-6-phosphate. This compound is thus a common intermediate in GlcNAc and GlcN metabolism. The next product, fructose-6-phosphate, is not specific to this pathway. Results reported here show that GlcN-6-phosphate, transported by a non-PTS system in a Nag^- $GlcN^-$ strain, can induce the *nag* regulon. However, the effect was small and much less than that observed with GlcNAc as the carbon source.

A problem with the common inducer hypothesis is that the growth and induction characteristics of GlcNAc and GlcN are very different. GlcNAc is a good carbon source producing growth rates comparable to those of glucose and a high level of induction of both *nagE* and *nagB* genes. GlcN, on the other hand, allows much slower growth rates and induction of the regulon is much less. One reason why GlcN is a worse carbon source than GlcNAc could be that it is a worse substrate for the transporter and cannot accumulate to a significantly high level within the cell. Alternatively, it might indicate that GlcNAc produces a stronger and/or different inducing signal than GlcN. GlcNAc-6-phosphate is the obvious candidate for a different signal. Unfortunately, GlcNAc-6-phosphate added to the growth medium failed to induce the regulon, presumably because it is not taken up by the cells.

Although the identity of the true inducer of the regulon is still in doubt, the experiments described here showed that induction does not depend on the activity of a particular transport system; an active *nagE* or *ptsM* for GlcNAc or *ptsM* or even *ptsG* for GlcN is sufficient. This, together with the observation that GlcN-6-phosphate transported by the *uhp* system is an inducer, strongly suggests that the inducing signal is sensed inside the cell.

This result is in contrast to that observed for two other sugar transport operons, those for glucose-6-phosphate uptake (41, 47) and for glucose induction of the *ptsHI crr* operon (7), where the presence of extracellular glucose-6-phosphate or glucose, respectively, has been shown to be sufficient to induce expression of the operons. In these

cases, a signal must be generated at the sugar receptor which crosses the membrane to activate transcription.

As GlcNAc and GlcN produce the same end product, fructose-6-phosphate, they might have been expected to produce equivalent growth rates, but this was clearly not the case. Possibly, the rate of uptake of GlcN is much slower than that of GlcNAc, even when using the same transporter (that encoded by the *ptsM* locus) and is not sufficient to maintain an adequate concentration of either inducer or substrate. This difference between GlcNAc and GlcN is emphasized by their variant sensitivities to glucose and gluconate. Both glucose and gluconate exert an effect of catabolite repression; however, it is possible that gluconate does not exert an effect of inducer exclusion (Antoine Danchin, personal communication). If GlcN does indeed generate a different inducing signal from that generated by GlcNAc, it is quite conceivable that it is subject to inducer exclusion, unlike the signal generated by GlcNAc. Alternatively, the differences observed could also be explained if glucose produces a higher level of catabolite repression than gluconate.

In conclusion, the two arms of the *nag* regulon are induced in parallel by GlcNAc and GlcN, but their relative levels of expression are controlled by the activity of the cAMP-CAP complex. The cAMP-CAP requirement for *nagE* expression is greater than for *nagB*. Quantitation of the induction observed with GlcNAc and GlcN under different conditions suggests that induction is not mediated identically by the two sugars. At the moment, it is not possible to decide whether this is due to a quantitative effect on the amount of the inducer produced or to the existence of different inducing substances.

ACKNOWLEDGMENTS

Constructive discussions with Hilde de Reuse and Philippe Régnier are gratefully acknowledged. I thank Josette Yaniv for criticism of the manuscript and Laurence Paineau for preparation of the text. This work was performed in the laboratory of Marianne Grunberg-Manago, whose interest is appreciated.

The work was supported by grants from the CNRS (URA 1139), from the Association pour la Recherche sur le Cancer, from the I.N.S.E.R.M. (C.R.E. 891017), from the E.E.C. (Contract SC1*/0194-C(AM)), from la Fondation pour la Recherche Médicale, and from E.I. Du Pont de Nemours.

LITERATURE CITED

1. **Bachmann, B. J.** 1983. Linkage map of *Escherichia coli* K-12, edition 7. Microbiol. Rev. **47**:180–230.
2. **Casadaban, M. J., J. Chou, and S. N. Cohen.** 1980. In vitro gene fusions that join an enzymatically active β-galactosidase segment to amino-terminal fragments of exogenous proteins: *Escherichia coli* plasmid vectors for the detection and cloning of translational initiation signals. J. Bacteriol. **143**:971–980.
3. **Chapon, C.** 1972. Role of the catabolite activator protein in the maltose regulon of *Escherichia coli*. J. Bacteriol. **150**:722–729.
4. **Curtis, S. J., and W. Epstein.** 1975. Phosphorylation of D-glucose in *Escherichia coli* mutants defective in glucose phosphotransferase, mannose phosphotransferase, and glucokinase. J. Bacteriol. **122**:1189–1199.
5. **Davis, T., M. Yamada, M. Elgort, and M. H. Saier.** 1988. Nucleotide sequence of the mannitol (*mtl*) operon in *E. coli*. Mol. Microbiol. **2**:405–412.
6. **de Crombrugghe, B., S. Busby, and H. Buc.** 1984. Cyclic AMP receptor protein: role in transcription activation. Science **224**: 831–838.
7. **de Reuse, H., and A. Danchin.** 1988. The *ptsH*, *ptsI*, and *crr* genes of the *Escherichia coli* phosphoenolpyruvate-dependent phosphotransferase system: a complex operon with several modes of transcription. J. Bacteriol. **170**:3827–3837.
8. **Dietz, G. W.** 1976. The hexose phosphate transport system of *E. coli*. Adv. Enzymol. **44**:237–259.
9. **Dunn, T. M., and R. Schleif.** 1984. Deletion analysis of the *E. coli* $araP_C$ and P_{BAD} promoters. J. Mol. Biol. **180**:201–204.
10. **Erni, B., B. Zanolari, and H. P. Kocher.** 1987. The mannose permease of *E. coli* consists of three different proteins. J. Biol. Chem. **262**:5238–5247.
11. **Friedrich, M. J., and R. J. Kadner.** 1987. Nucleotide sequence of the *uhp* region of *Escherichia coli*. J. Bacteriol. **169**:3556–3563.
12. **Gutierrez, C., C. Chapon, and M. Schwartz.** 1985. Indirect effects of the 3′-5′ cAMP binding protein (CAP) on the transcription of the *malPQ* operon in *E. coli*. Biochimie **67**:145–148.
13. **Holmes, R. P., and R. R. B. Russell.** 1972. Mutations affecting amino sugar metabolism in *Escherichia coli* K-12. J. Bacteriol. **111**:290–291.
14. **Jones-Mortimer, M. C., and H. L. Kornberg.** 1980. Amino sugar transport systems of *E. coli* K12. J. Gen. Microbiol. **117**: 369–376.
15. **Joseph, E., A. Danchin, and A. Ullmann.** 1981. Regulation of galactose operon expression: glucose effects and role of cyclic AMP. J. Bacteriol. **146**:149–154.
16. **Lee, N. L., W. O. Gielow, and R. G. Wallace.** 1981. Mechanism of *araC* autoregulation and the domains of two overlapping promoters P_c and P_{BAD}, in the L-arabinose regulatory region of *E. coli*. Proc. Natl. Acad. Sci. USA **78**:752–756.
17. **Lichenstein, H. S., E. P. Hamilton, and N. Lee.** 1987. Repression and catabolite gene activation in the *araBAD* operon. J. Bacteriol. **169**:811–822.
18. **Liu-Johnson, H.-N., M. R. Gatenberg, and D. M. Crothers.** 1986. The DNA binding domain and bending angle of *E. coli* CAP protein. Cell **47**:995–1005.
19. **Maloy, S. R., and W. D. Nunn.** 1981. Selection for loss of tetracycline resistance by *Escherichia coli*. J. Bacteriol. **145**: 1110–1112.
20. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
22. **Mitchell, W. J., D. W. Saffen, and S. Roseman.** 1987. Sugar transport by the bacterial phosphotransferase system. *In vivo* regulation of lactose transport in *E. coli* by III^{Glc} of the PEP-glycose phosphotransferase system. J. Biol. Chem. **262**: 16254–16260.
23. **Miyada, C. G., L. Stoltzfus, and G. Wilcox.** 1984. Regulation of the *araC* gene of *E. coli*: catabolite repression, autoregulation, and effect on *araBAD* expression. Proc. Natl. Acad. Sci. USA **81**:4120–4124.
24. **Neidhardt, F. C., P. L. Bloch, and D. F. Smith.** 1974. Culture medium for enterobacteria. J. Bacteriol. **119**:736–747.
25. **Nelson, S. O., J. K. Wright, and P. W. Postma.** 1983. The mechanism of inducer exclusion. Direct interaction between purified enzyme III^{Glc} of the phosphoenolpyruvate: sugar phosphotransferase system and the lactose carrier of *E. coli*. EMBO J. **2**:715–720.
26. **Novotny, M. J., W. L. Frederickson, E. B. Waygood, and M. H. Saier.** 1985. Allosteric regulation of glycerol kinase by enzyme III^{Glc} of the phosphotransferase system in *Escherichia coli* and *Salmonella typhimurium*. J. Bacteriol. **162**:810–816.
27. **Ogden, S., D. Haggerty, C. M. Stoner, D. Kolodrubetz, and R. Schleif.** 1980. The *E. coli* L-arabinose operon: binding sites of the regulatory proteins and a mechanism of positive and negative regulation. Proc. Natl. Acad. Sci. USA **77**:3346–3350.
28. **Peri, K. G., and E. B. Waygood.** 1988. Sequence of cloned enzyme II^{Nag} of the phosphoenolpyruvate: *N*-acetylglucosamine phosphotransferase system of *E. coli*. Biochemistry **27**:6054–6061.
29. **Plumbridge, J. A.** 1987. Organization of the *E. coli* chromosome between *glnS* and *glnU,V*. Mol. Gen. Genet. **209**:618–620.
30. **Plumbridge, J. A.** 1989. Sequence of the *nagBACD* operon in *E. coli* and pattern of expression within the operon. Mol. Microbiol. **23**:505–515.

31. **Plumbridge, J. A., and D. Soll.** 1989. Characterization of cis-acting mutations which increase expression of a *glnS-lacZ* fusion in *E. coli*. Mol. Gen. Genet. **216**:113–119.
32. **Postma, P.** 1987. Phosphotransferase system for glucose and other sugars, p. 127–141. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington D.C.
33. **Postma, P. W., and J. Lengeler.** 1985. Phosphoenolpyruvate carbohydrate phosphotransferase system of bacteria. Microbiol. Rev. **49**:232–269.
34. **Raibaud, O., D. Vidal-Ingigliardi, and E. Richet.** 1989. A complex nucleoprotein structure involved in activation of transcription of two divergent *E. coli* promoters. J. Mol. Biol. **205**: 471–485.
35. **Raphaell, A. W., and M. H. Saier.** 1980. Regulation of genes coding for enzyme constituents of the bacterial phosphotransferase system. J. Bacteriol. **141**:658–663.
36. **Rogers, M. J., T. Ohgi, J. Plumbridge, and D. Söll.** 1988. Nucleotide sequences of *E. coli nagE* and *nagB* genes: the structural genes for the N-acetyl glucosamine transport protein of the bacterial phosphoenolpyruvate:sugar phosphotransferase system and for glucosamine-6-phosphate deaminase. Gene **62**: 197–207.
37. **Roy, A., C. Haziza, and A. Danchin.** 1983. Regulation of adenylate cyclase synthesis in *E. coli*: nucleotide sequence of the control region. EMBO J. **2**:791–797.
38. **Saier, M. H.** 1989. Protein phosphorylation and allosteric control of inducer exclusion and catabolite repression by the bacterial phospho-enol-pyruvate:sugar phosphotransferase system. Microbiol. Rev. **53**:109–120.
39. **Saier, M. H., M. J. Novotny, D. Comeau-Fuhrman, T. Osumi, and J. D. Desai.** 1983. Cooperative binding of the sugar substrates and allosteric regulatory protein (enzyme III^{Glc} of the phosphotransferase system) to the lactose and melibiose permeases in *Escherichia coli* and *Salmonella typhimurium*. J. Bacteriol. **155**:1351–1357.
40. **Saris, P. E. J., and T. Palva.** 1987. The *ptsL*, *pel/ptsM* (*manXYZ*) locus consists of three genes involved in mannose uptake in *E. coli* K12. FEMS Microbiol. Lett. **44**:371–376.
41. **Shattuck-Eidens, D. M., and R. J. Kadner.** 1981. Exogenous induction of the *Escherichia coli* hexose phosphate transport system defined by *uhp-lac* operon fusions. J. Bacteriol. **148**: 203–209.
42. **Stoltzfus, L., and G. Wilcox.** 1989. Effect of mutations in the cyclic AMP receptor protein binding site on *araBAD* and *araC* expression. J. Bacteriol. **171**:1178–1184.
43. **Ullman, A., and A. Danchin.** 1983. Role of cyclic AMP in bacteria. Adv. Cyclic Nucleotide Res. **15**:1–53.
44. **Vögler, A. P., and J. W. Lengeler.** 1989. Analysis of the *nag* regulon from *E. coli* K12 and *Klebsiella pneumoniae* and its regulation. Mol. Gen. Genet. **219**:97–105.
45. **Warwicker, J., B. P. Engelman, and T. A. Steitz.** 1987. Electrostatic calculations and model building suggest that DNA bound to CAP is sharply bent. Proteins Struct. Funct. Genet. **2**: 283–289.
46. **Weber, I. T., and T. A. Steitz.** 1987. Structure of a complex of catabolite gene activator protein and cAMP refined at 2.5 Å resolution. J. Mol. Biol. **198**:311–326.
47. **Weston, L. A., and R. J. Kadner.** 1987. Identification of Uhp polypeptides and evidence for their role in exogenous induction of the sugar phosphate transport system of *Escherichia coli* K-12. J. Bacteriol. **169**:3546–3555.
48. **White, R. J.** 1968. Control of amino sugar metabolism in *E. coli* and isolation of mutants unable to degrade amino sugars. Biochem. J. **106**:847–858.
49. **White, R. J.** 1970. The role of the phosphoenolpyruvate phosphotransferase system in the transport of *N*-acetylglucosamine by *E. coli*. Biochem. J. **118**:89–92.
50. **Wu, H.-M., and D. M. Crothers.** 1984. The locus of sequence-directed and protein-induced DNA bending. Nature (London) **308**:509–513.
51. **Yamada, M., and M. H. Saier.** 1987. Glucitol-specific enzymes of the phosphotransferase system in *E. coli*. Nucleotide sequence of the *gut* operon. J. Biol. Chem. **262**:5455–5463.