JOURNAL OF BACTERIOLOGY, July 2008, p. 4677–4686 Vol. 190, No. 13
0021-9193/08/$08.00+0 doi:10.1128/JB.00315-08

Switching Control of Expression of *ptsG* from the Mlc Regulon to the NagC Regulon▽

Samir El Qaidi and Jacqueline Plumbridge*

Institut de Biologie Physico-Chimique, UPR9073-CNRS, 13 rue Pierre et Marie Curie, 75005 Paris, France

Received 2 March 2008/Accepted 25 April 2008

The Mlc and NagC transcriptional repressors bind to similar 23-bp operators. The sequences are weakly palindromic, with just four positions totally conserved. There is no cross regulation observed between the repressors in vivo, but there are no obvious bases which could be responsible for operator site discrimination. To investigate the basis for operator recognition and to try to understand what differentiates NagC sites from Mlc sites, we have undertaken mutagenesis experiments to convert *ptsG* from a gene regulated by Mlc into a gene regulated by NagC. There are two Mlc operators upstream of *ptsG*, and to switch *ptsG* to the NagC regulon, it was necessary to change two different characteristics of both operators. Firstly, we replaced the AT base pair at position +/−11 from the center of symmetry of the operators with a GC base pair. Secondly, we changed the sequence of the CG base pairs in the central region of the operator (positions −4 to +4 around the center of symmetry). Our results show that changes at either of these locations are sufficient to lose regulation by Mlc but that both types of changes in both operators are necessary to convert *ptsG* to a gene regulated by NagC. In addition, these experiments confirmed that two operators are necessary for regulation by NagC. We also show that regulation of *ptsG* by Mlc involves some cooperative binding of Mlc to the two operators.

Mlc and NagC are homologous proteins, and both act as transcriptional repressors in *Escherichia coli*. Mlc represses genes involved in the uptake of glucose, while NagC controls the use of *N*-acetylglucosamine (GlcNAc). Both glucose and GlcNAc are transported into *E. coli* via the phosphotransferase system (PTS). Mlc represses *ptsG*, the gene for the major glucose transporter; the *ptsHI-crr* genes, which encode the soluble components of the PTS; the *manXYZ* genes, which encode an alternative transporter for glucose, as well as other hexoses; and also *malT*, the positive transcriptional regulator of the *mal* regulon (4, 11, 12, 24, 26, 27). NagC represses the divergent *nagE-nagBACD* operons for the uptake and degradation of GlcNAc and the *chb* operon, which contains genes for the transport and degradation of chitobiose (a dimer of GlcNAc) (28, 31). In addition, NagC activates the expression of the *glmUS* operon, which contains genes of the biosynthetic pathway for UDP-GlcNAc (23), and also the expression of the *fimB* recombinase necessary for the off-to-on switching of the *fim* operon for type I fimbriae (39, 40).

The Mlc and NagC proteins are 40% identical, with 70% similarity, and are members of the repressor subgroup of the ROK (repressors, open reading frames, kinases) family (44). Although Mlc and NagC have clearly defined and different functions in *E. coli*, it is surprising that the sequence of the helix-turn-helix (H-T-H) motifs in the N-terminal DNA binding domain are unexpectedly similar (Fig. 1B). The C-terminal domains of Mlc and NagC are also homologous, but the inducing signals which displace Mlc and NagC from their DNA binding sites are very different (14, 18, 28, 37, 43).

The DNA binding sites of Mlc and NagC upstream of all of the regulated genes have been identified by DNase I footprinting and are listed in Fig. 1A. The sites for both proteins are rather similar, but there is no strong consensus sequence for either protein. The protected DNA sequence covers 23 bp and is based on a palindrome with a central "zero" position; however, only positions −6, −5, +5, and +6 are 100% conserved in all of the operators of both repressors (shown in yellow in Fig. 1A). The lack of a strong defining consensus for either repressor is apparent in the sequence logos generated from the binding sites (35), indicating that the information content of bases outside positions +/−5 and 6 is low.

We previously noted certain characteristics of the sites (25). In Mlc operators, positions 7 to 11 are occupied exclusively by A or T but with more T's to the left and A's to the right (shown in green). In NagC sites, positions 7 to 10 are mostly A or T while there is a strong tendency to find a C or G at position +11 or −11 (12/18 positions, shown in red). The *nagE* operator is a notable exception, since it has −11T and +11A. Previous *nagE* operator mutagenesis experiments showed that changing either or both nucleotides at positions +/−11 to C or G strongly increased the affinity for NagC and there was no requirement for a palindrome (30). However, the fact that the *nagE* operator is not regulated by Mlc (25) showed that the identity of the base at +/−11 is not sufficient to distinguish between Mlc and NagC sites.

We also noted that there was a bias in the distribution of the C and G bases in the central region of NagC sites. Most of the CG base pairs in positions −4 to −1 and +1 to +4 (excluding those in the zero position) conform to the pattern −4 CGCGNCGCG +4 (31/36 bp, in magenta in Fig. 1A). On the other hand, in Mlc sites there was more of a tendency for the CG base pairs in the central region to follow the pattern −4 GCGCNGCGC +4 (18/27 bp, in blue in Fig. 1A). The systematic mutagenesis analysis of the *nagE* operator had shown that

* Corresponding author. Mailing address: Institut de Biologie Physico-Chimique, UPR9073-CNRS, 13 rue Pierre et Marie Curie, 75005 Paris, France. Phone: 33 (0)1 58 41 51 52. Fax: 33 (0)1 58 41 50 20. E-mail: Jackie.Plumbridge@ibpc.fr.

▽ Published ahead of print on 9 May 2008.

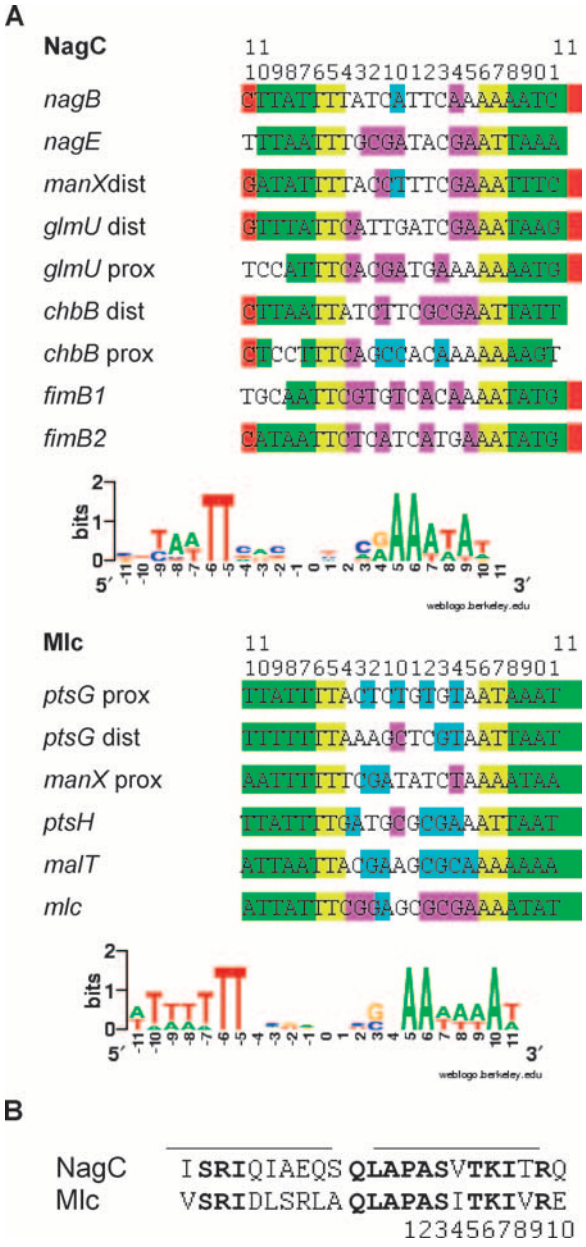


FIG. 1. (A) Operator sequence alignments and logos. The sequences of the known NagC and Mlc operators are shown. The positions are numbered around the central zero position of the imperfect palindrome covering 23 bp. They are colored coded as follows. Totally conserved positions −5T, −6T, +5A, and +6A are in yellow. A or T bases in positions 7 to 10 in NagC and positions 7 to 11 in Mlc are shaded green. C or G bases at positions +/−11 in NagC sites are shaded red. CG bases in the central region from −4 to +4, which make up the NagC-type central CGCG motif, are shaded magenta, whereas bases in the GCGC Mlc-type central pattern are shaded blue. Sequence logos (35) for the NagC and Mlc sites were derived from the http://weblogo.berkeley.edu/ site. (B) Comparison of H-T-H motifs of Mlc and NagC. Conserved bases are in bold. Amino acids of the recognition helix are numbered.

a C at −4 and a G at −3 exhibited the highest affinity for NagC while a G at −4 and a C at −3 showed the lowest (30). Although the majority of the NagC sites have either −4C or +4G, the *nagB* operator, which is a high-affinity NagC site, does not. In fact, the central region of the *nagB* operator has fewer CG base pairs than the other operators. For convenience, when the C and G bases in the central region conform to the pattern mostly found in NagC sites, we will refer to a "central CGCG" pattern (magenta in Fig. 1A) while the reverse pattern, found more frequently in Mlc sites, will be called the "central GCGC" pattern (blue in Fig. 1A). It should be noted that when we refer to a CGCG pattern, it does not imply that every position is occupied by a C or a G but that the majority of C and G bases are found in the positions indicated.

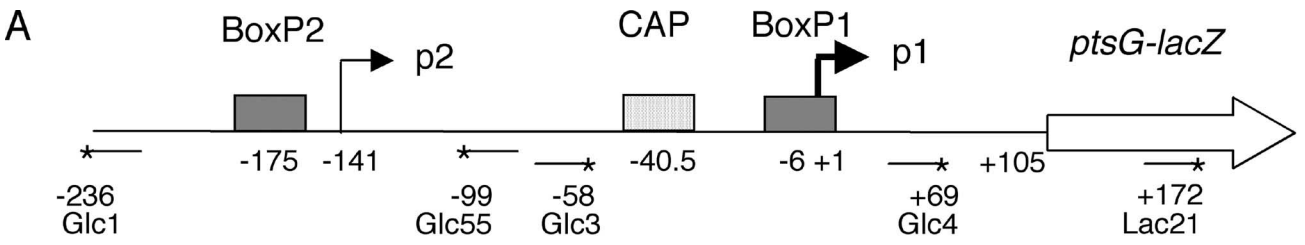
Subsequently, to try to distinguish Mlc sites from NagC sites, binding sites for one or the other protein were selected in vitro (by a SELEX-type experiment). Unexpectedly, this pulled out operator sites for both NagC and Mlc where the central regions were very CG rich and almost all of the sites were in the CGCG pattern (25). This apparent contradiction between native in vivo sites and in vitro-selected results did not clarify what was distinguishing NagC sites from Mlc sites in vivo.

A third difference between Mlc- and NagC-controlled genes is that genes controlled by NagC all possess two operators and regulation by NagC necessitates cooperative binding of NagC to the two sites, resulting in the formation of a DNA loop between repressor-bound operators, as shown for *nagEB*, *glmU*, and *fimB* (23, 29, 40). The center-to-center distance between the two operators varies between 93 bp for *nagE-B* to 212 bp for *fimB*. To try to clarify what is necessary to define an Mlc or a NagC site, we have attempted to change a gene regulated by Mlc into one regulated by NagC. In order to increase our chance of success, we carried out this operator conversion on the *ptsG* gene because it is the only gene in the Mlc regulon to be controlled from two operator sites (separated by 169 bp). We tested both the effects of C/G at positions +/−11 and of a central CGCG-type pattern on regulation by Mlc and NagC to investigate whether either or both were capable of preventing Mlc binding and permitting NagC binding.

MATERIALS AND METHODS

Bacterial methods. The JM-G301 strain carries a *ptsG-lacZ* transcriptional fusion on a λ lysogen (22). β-Galactosidase activities were measured in morpholinepropanesulfonic acid (MOPS) medium with 0.5% Casamino Acids and 0.4% glycerol, 0.2% glucose, or 0.2% GlcNAc at 30°C. Aliquots were removed at several points during exponential growth, and β-galactosidase activities (Miller units) were measured as described previously (16). At least two, and usually more, cultures were tested, and the mean with standard deviations was calculated. *mlc*::tc (24), Δ*mlc*::cat, and Δ*nagC*::tc (22) mutations were introduced into the lysogens by P1*vir* transduction.

Mutagenesis. The EcoRI-BamHI insert of the *ptsG-lacZ* fusion in JM-G301 was inserted into pTZ18R (to give pTZ/Glc1E-Nsi), and oligonucleotide-directed mutagenesis of the *ptsG* operators was initially carried out by the Kunkel method (13). Alternatively, mutations were made by two rounds of PCR with the mutagenic oligonucleotide in a first PCR with a downstream primer (Lac22; GGTTTTCCCAGTCACGACGTTG). The product of the first PCR was purified and used as a primer in a second PCR with an upstream oligonucleotide (Rev22; CACACAGGAAACAGCTATGACC) and with the pTZ/Glc1-Nsi plasmid as the template. *Pwo* (Roche) was used for the PCR. Multiple mutations in the two operators were made with mutated plasmids as the templates. The mutated fragments were recloned back into the pRS415 fusion vector for recombination with λRS45 as previously described (38). JM101 was lysogenized, and monolysogens were tested for by the method of Powell et al. (32). Oligonucleotides corresponding to DNA upstream of the fusion insert in the lysogen (within the *bla* gene, RBP22; CCGAAAAGTGCCACCTGACGTC) and Lac22 (downstream, within *lacZ*) were used to amplify a fragment from the chromosome of the lysogenic bacteria for sequencing (by MWG Biotech, Martinsried, Germany) to verify the presence of the mutations in the lysogens.



B

	Mutation	BoxP2 (-35 p2) 11 11 1098765432101234567890 1	BoxP1 (-10 p1) 11 11 1098765432101234567890 1	*wt*	*mlc*	*nagC*
A	WT	TTTTTTTAAAGC***TCGTAA***TTAAT	TTATTT***TACTCT***GTGTAATAAAT	297 ± 59	2781 ± 274	220 ± 55
B	Glc52	CTTTTTTAAAGC***TCGTAA***TTAAG	TTATTT***TACTCT***GTGTAATAAAT	1392 ± 26	2770 ± 42	1205 ± 10
	Glc51	TTTTTTTAAAGC***TCGTAA***TTAAT	CTATTT***TACTCT***GTGTAATAAAG	1535 ± 192	1827 ± 347	1482 ± 187
	Glc52-51	CTTTTTTAAAGC***TCGTAA***TTAAG	CTATTT***TACTCT***GTGTAATAAAG	1450 ± 185	1556 ± 72	1475 ± 47
C	Glc64	TTTTTTTCGCGC***TCGTAA***TTAAT	TTATTT***TACTCT***GTGTAATAAAT	479 ± 8	2490 ± 72	428 ± 2
	Glc65	TTTTTTTAAAGC***TCGTAA***TTAAT	TTATTT***TAGTGT***CGCGAATAAAT	230 ± 8	260 ± 11	213 ± 2
	Glc64-65	TTTTTTTCGCGC***TCGTAA***TTAAT	TTATTT***TAGTGT***CGCGAATAAAT	182 ± 11	196 ± 3	155 ± 2
D	Glc60	CTTTTTTCGCGC***TCGTAA***TTAAG	TTATTT***TACTCT***GTGTAATAAAT	957 ± 33	2780 ± 42	1197 ± 82
	Glc63	TTTTTTTAAAGC***TCGTAA***TTAAT	CTATTT***TAGTGT***CGCGAATAAAG	196 ± 9	270 ± 31	284 ± 31
E	Glc61-62	CTTTTTTCGCGC***TTCGAA***TTAAG	CTATTT***TACTCT***CGCGAATAAAG	164 ± 5	160 ± 6	351 ± 13
	Glc60-63	CTTTTTTCGCGC***TCGTAA***TTAAG	CTATTT***TAGTGT***CGCGAATAAAG	39 ± 4	40 ± 2	193 ± 33
	Glc61-63	CTTTTTTCGCGC***TTCGAA***TTAAG	CTATTT***TAGTGT***CGCGAATAAAG	12 ± 2	11 ± 1	67 ± 1
F	Glc60-51	CTTTTTTCGCGC***TCGTAA***TTAAG	CTATTT***TACTCT***GTGTAATAAAG	1613 ± 313	1636 ± 134	1507 ± 33
	Glc61-51	CTTTTTTCGCGC***TTCGAA***TTAAG	CTATTT***TACTCT***GTGTAATAAAG	1280 ± 44	1318 ± 11	1262 ± 84
	Glc52-62	CTTTTTTAAAGC***TCGTAA***TTAAG	CTATTT***TACTCT***CGCGAATAAAG	509 ± 77	494 ± 35	588 ± 61
	Glc52-63	CTTTTTTAAAGC***TCGTAA***TTAAG	CTATTT***TAGTGT***CGCGAATAAAG	336 ± 58	314 ± 14	319 ± 35
	Glc60-65	CTTTTTTCGCGC***TCGTAA***TTAAG	TTATTT***TAGTGT***CGCGAATAAAT	184 ± 2	212 ± 1	216 ± 32
	Glc64-63	TTTTTTTCGCGC***TCGTAA***TTAAT	CTATTT***TAGTGT***CGCGAATAAAG	85 ± 6	104 ± 2	145 ± 1

C

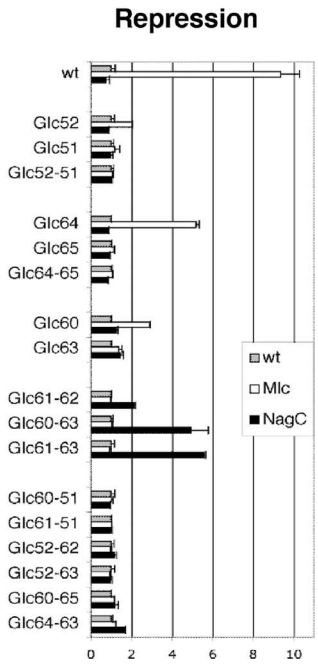


FIG. 2. Organization of the *ptsG* regulatory region and effects of mutations in the Mlc operators on the expression of *ptsG-lacZ*. (A) Positions of the centers of the Mlc operators, BoxP1 and BoxP2, and the CAP site compared to the p1 transcription start site, labeled +1. The *ptsG* p1 start site was previously mapped as a series of bands centered at the A which corresponds to position +6 of the BoxP1 operator (26). The minor p2 promoter starts at −141. Oligonucleotides are indicated by a line with a star at the 5′ end, and the numbers underneath indicate the positions of the 5′ ends. (B) The sequences of the WT *ptsG* operators and the different mutations studied are shown. BoxP2 overlaps the −35 sequence of p2 (bold characters), and BoxP1 overlaps the −10 sequence of p1 (bold characters). The mutations are named according to the oligonucleotide used to make the mutation. When two numbers are given, the first corresponds to the mutation in BoxP2 and the second to that in BoxP1. Changes to C or G at positions +/−11 are shown as white characters on black. Changes in the central region, positions −4 to +4, which conform to the CGCG pattern characteristic of NagC sites are shown by gray shading. The −1 position in BoxP2 is already a G in the WT BoxP2 operator and so is shown shaded throughout. The effects of the *mlc* and *nagC* mutations on the expression of the WT and mutated *ptsG-lacZ* fusions are shown in Miller units. Bacteria were grown in MOPS-Casamino Acids medium with 0.4% glycerol at 30°C. (C) Repression by Mlc and NagC was calculated as the ratio of expression in the *mlc* or *nagC* mutant strain compared to the WT background.

DNase I footprinting. PCR fragments Glc1-3, Glc4-55, and Glc1-4 (Fig. 2A), covering BoxP1, BoxP2, or both of the *ptsG* operators, were made by PCR with the Glc1 or Glc4 oligonucleotide labeled with [γ-^{32}P]ATP and polynucleotide kinase. Mlc with a C-terminal His tag was the kind gift of Sabine Seitz (University of Konstanz, Konstanz, Germany) or was prepared without a His tag by Olivier Pellegrini on the basis of the method described previously (11). NagC with a C-terminal His tag was from Charles Bell and Mitchell Lewis (University of Pennsylvania, Philadelphia). DNAs and various dilutions of proteins were mixed in 25 mM HEPES–100 mM K-glutamate buffer (pH 8.0)–0.5 mg ml^{-1} bovine serum albumin for 10 min at room temperature and treated with DNase I (0.1 μg ml^{-1} for 1 min). The reaction was stopped with 100 μl phenol saturated with Tris-EDTA (pH 8.0), and 200 μl 0.4 M Na acetate–2.5 mM EDTA containing 10 μg ml^{-1} sonicated herring sperm DNA was added. After phenol extraction and ethanol precipitation, the footprints were analyzed on 6 or 8% denaturing acrylamide gels, dried, and subjected to phosphorimaging. Quantitative analysis of Mlc and NagC binding to single mutated operators was measured by DNase I footprinting with twofold serial dilutions of Mlc or NagC. Protection was quantified by the ImageQuant TL software. Several DNase I-cleaved bands were measured both within and outside the protected region. The intensities of the bands within each lane were normalized to two bands outside the protected region as standards, and then the levels of protection of three bands within the protected site at different concentrations of the proteins were calculated by comparison to the intensities of the same bands in the absence of protein. Percent protection was plotted to calculate the concentration giving 50% protection, which was taken as the apparent dissociation constant (K_d). The DNA concentrations were routinely 3 to 10 nM, so that the values of tight binding sites are probably overestimated by this method. Protein concentrations are expressed for the monomer. Both Mlc and NagC presumably bind DNA at least as dimers (like other H-T-H proteins binding to palindromic operators). At higher concentrations (>100 nM), NagC has a tendency to aggregate. This meant that it was impossible to measure NagC binding affinities by a band shift assay because at these concentrations the DNA formed an aggregate with the protein which stayed in the wells. This same tendency to self-aggregate means that in the footprints DNA adjacent to the operator site is also protected at higher concentrations (see Fig. 5, lanes 2 and 10).

S1 mapping. RNA was prepared by the hot-phenol method from exponential-phase cultures of the different lysogenic bacteria. The DNA probe used was the *ptsG-lacZ* fusion-specific Glc1-Lac21 PCR fragment labeled at Lac21 (ACTGG CGGCTGTGGGATTAAC) (Fig. 2A). To map mRNAs from lysogens with mutations in BoxP1 (Glc51, -63, -65, and -44), the appropriate mutant fusion plasmid DNA was used as the template to synthesize the labeled probe. S1 analysis was carried out as described previously (24). Thirty micrograms of total RNA or tRNA after denaturation with the probe was hybridized overnight at 54°C in a final volume of 50 μl containing 40 mM piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) PIPES (pH 6.4), 1 mM EDTA, 0.4 M NaCl, and 80% formamide and then treated with S1 (Roche; 100 U in 0.4 ml 30 mM Na acetate [pH 5.0]–0.25 M NaCl–1 mM $ZnCl_2$–5% glycerol for 30 min at 37°C), precipitated with ethanol, analyzed on a 6% denaturing acrylamide gel, dried, and subjected to phosphorimaging.

RESULTS

Mutations in the *ptsG* operators. The organization of the *ptsG* promoter is shown in Fig. 2A. The *ptsG* gene is expressed from two promoters, major promoter p1 and minor promoter p2, which starts 141 bp upstream of p1 and accounts for less than 10% of the total transcription. There are two Mlc operators, BoxP1, which overlaps the −10 promoter sequence and +1 transcription start site of p1, and BoxP2, which overlaps the −35 region of the p2 promoter (12, 26). Two types of mutations were created in both of the operators, changing either the +/−11 positions from AT to CG or the pattern of bases within the central −4 to +4 positions from the central GCGC pattern characteristic of Mlc operators to the CGCG pattern characteristic of NagC operators. All of these mutations were inserted into the *ptsG-lacZ* fusion on a λ lysogen. The mutations are named according to the number of the Glc oligonucleotide used to create the mutation in either operator, so that mutations in both operators have double numbers, as in Glc60-63. All of the mutations created are shown in Fig. 2B. Since the BoxP1 operator overlaps the −10 sequence of p1 and the BoxP2 operator overlaps the −35 sequence of p2, the operator mutations also affect RNA polymerase efficiency and thus the maximum β-galactosidase activities of the fusions vary. Figure 2B gives the β-galactosidase activities during growth on glycerol (noninducing conditions) of the original wild-type (WT) fusion and the fusions carrying the mutant operators in the presence and absence of *mlc* and *nagC* mutations. Repression by Mlc or NagC for each operator mutant (Fig. 2C) was calculated as the ratio of expression in the WT background, compared to that in the *mlc* or *nagC* mutant strain. For the WT *ptsG-lacZ* fusion, the presence of the *mlc* mutation produces a 10-fold increase in β-galactosidase expression during growth on glycerol, while the *nagC* mutation has no effect (Fig. 2B, section A).

(i) Replacement of A/T at positions +/−11 with C/G is sufficient to lose regulation by Mlc. Initially, we tested the effect of replacing A/T at positions +/−11 with C or G in BoxP1 and BoxP2 (Fig. 2B, section B). The mutation Glc51 (+/−11CG in BoxP1) resulted in complete loss of regulation by Mlc, while the mutation Glc52 in BoxP2 was regulated only twofold by Mlc. The double mutation (Glc51-52) was also completely derepressed, so that β-galactosidase activities were identical in the WT and *mlc* and *nagC* mutant strains. Thus, +/−11CG in the two operators results in loss of regulation by Mlc but does not allow regulation by NagC. The Glc52 mutation (and the other mutations in BoxP2, Glc64 and Glc60 [see below]) produced partial derepression of *ptsG* expression because regulation of *ptsG* involves cooperative binding of Mlc to the two operators and the residual regulation of BoxP2 operator mutants is due to Mlc regulation of p1 via BoxP1, as discussed below.

(ii) Replacement of the Mlc-type central GCGC pattern with the NagC-type CGCG pattern also provokes derepression of *ptsG*. The mutation Glc65 in BoxP1 replaced 6 bp so that instead of having two C residues conforming to the GCGC pattern typical of Mlc operators, there were 6 bp conforming to the NagC-type central CGCG pattern. The Glc64 mutation has 4 bp in BoxP2 in the CGCG pattern. There was no effect of either an *mlc* or an *nagC* mutation on the Glc65 or Glc64-65 fusion; both were completely derepressed, although the maximum activity was much lower because the Glc65 mutation has affected the p1 promoter (Fig. 2B, section C). As in the case of the Glc52 mutation, the Glc64 mutation was partially derepressed. Since the double mutation Glc64-65 gave the same activity in the WT and *mlc* and *nagC* mutant strains, the presence of the NagC-type CGCG central patterns in both operators is not sufficient to allow regulation by NagC. (The effects of the mutations on promoter activity are discussed below.)

(iii) Combining the +/−11CG change with an NagC-type central CGCG pattern in one operator does not allow regulation by NagC. Since either +/−11CG or a central CGCG pattern in the BoxP1 operator is sufficient to completely inhibit Mlc regulation without allowing any regulation by NagC, we tested whether the two types of mutation together in one operator would allow regulation by NagC (Fig. 2B, section D). The expression of the fusion carrying the mutation Glc63 in BoxP1 was slightly increased (1.3-fold) by either an *mlc* or an *nagC* mutation, implying that there is low-level regulation by both Mlc and NagC. However, the presence of the *nagC* and *mlc* mutations simultaneously did not produce any further increase in *ptsG* expression than did that of either mutation individually (data not shown). (We have previously observed that the presence of both the *nagC* and *mlc* mutations in a strain has a detrimental effect on the level of expression from a *nagE-lacZ* reporter [25] and so cannot exclude the possibility that any small increase in Glc63 expression, due to the simultaneous loss of Mlc and NagC binding, is offset by the detrimental effect of the two mutations together on cell physiology.) The fusion carrying the Glc60 mutation with +/−11CG and 4 bp in the NagC-type central CGCG pattern in BoxP2 is derepressed threefold by the *mlc* mutation because p1 is still regulated by Mlc and only slightly (1.25-fold) increased by the *nagC* mutation. Thus, neither Glc60 nor Glc63 alone allows significant regulation by NagC.

(iv) The presence of +/−11CG and a NagC-type CGCG pattern in both operators allows regulation by NagC. Combining the Glc60 mutation in BoxP2 with the Glc63 mutation in BoxP1 (Glc60-63) produced a fusion with very low expression in the WT strain (40 U) but which was increased nearly fivefold by the *nagC* mutation and was unaffected by the *mlc* mutation, showing that the fusion is now under the control of NagC (Fig. 2B, section E, and panel C). Two other combinations of BoxP1 and BoxP2 operators with +/−11CG and a CGCG central pattern were tested, Glc61-62 and Glc61-63. Both were also unaffected by the *mlc* mutation, and both were induced by the *nagC* mutation, showing that the presence of +/−11CG and a

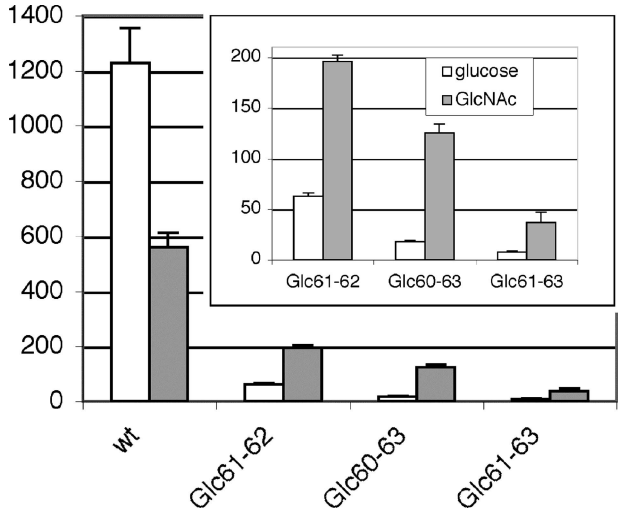


FIG. 3. Effect of growth on glucose or GlcNAc on expression from operator mutations. Bacteria carrying the WT *ptsG-lacZ* lysogen or the mutants indicated were grown in MOPS-Casamino Acids medium with either 0.2% glucose or 0.2% GlcNAc at 30°C. β-Galactosidase activities are expressed in Miller units. The insert shows the regulation of the mutations on an expanded scale. The expression of the WT fusion is partly induced by growth on GlcNAc compared to growth on glycerol (Fig. 2B). This is due to the partial dephosphorylation of PtsG during the uptake of GlcNAc. Transport of GlcNAc by the PTS drains phosphates from the other PTS proteins in the cell, notably, $EIIA^{Glc}$ and hence $EIIB^{Glc}$ (27).

central CGCG pattern in both operators allows NagC to bind and to regulate *ptsG* expression.

Regulation of these operator mutants by NagC was confirmed by studying the induction of *ptsG-lacZ* by growth on GlcNAc compared to glucose. All three mutant fusions were better expressed in GlcNAc than in glucose, whereas expression of the WT fusion was higher in glucose than in GlcNAc (Fig. 3). The relatively high level of expression of the WT fusion on GlcNAc compared to glycerol (Fig. 2A) is due to the fact that the transport of GlcNAc, a PTS sugar, also produces some dephosphorylation of other PTS transporters, including PtsG. Since the inducing signal for Mlc is dephosphorylated PtsG, growth on GlcNAc produces partial derepression of Mlc-regulated genes (27).

(v) The presence of +/−11CG and the NagC-type CGCG pattern in both operators are necessary for regulation by NagC. We tested a series of constructs where three of the four types of changes in the two operators were made; i.e., either both operators have +/−11CG but only one has the NagC-type central CGCG pattern or both operators have the NagC-type central CGCG pattern and only one has +/−11CG (Fig. 2B, section F). In no case was there any significant regulation by either *mlc* or *nagC*. The possible exception is Glc64-63, which is induced about 1.7-fold by the *nagC* mutation; but this should be compared to Glc60-63, with the four changes, which is regulated 5-fold. We conclude that in the context of the *ptsG* promoter, both the presence of a C or G at +/−11 and a predominance of bases in the CGCG pattern in the central region in both operators are necessary for the regulation of *ptsG* by NagC.

Effects of the mutations on the *ptsG* p1 and p2 promoters. The BoxP1 and BoxP2 operators overlap the transcription start site and −10 sequence of p1 and the −35 sequence of p2, respectively. As can be seen from the maximum values of the β-galactosidase activities (which give the sum of the two promoters) in Fig. 2B, the mutations created often have severe effects on the promoters. We have used S1 analysis to distinguish the contributions of the two promoters. Both transcripts are very weak during growth on glycerol, but the presence of the *mlc* mutation produced a strong transcript for p1 and a weak but detectable transcript for p2 (Fig. 4A, lanes 1 and 2).

The mutation Glc51 (+/−11CG in BoxP1) reduces the maximum activity by about one-third, and S1 analysis shows that this decrease is associated with loss of the shorter p1 transcripts which overlap the +11G mutation (Fig. 4A, mutation Glc52-51, lane 3). Other mutations in the central region of BoxP1 have even stronger effects on p1 promoter activity. For example, Glc63 (6 bp in the central CGCG pattern and +/−11CG) and Glc65 (6 bp in the central CGCG pattern) have lost nearly all of the p1 transcript (Fig. 4A, lanes 4, 5, 8, and 9). The 6-bp replacement changes 2 bp within the −10 sequence (TACTCT to TAGTGT) and also changes the 4 bp immediately downstream of the −10 sequence. These changes to the region involved in open complex formation have effectively silenced the p1 promoter.

The weak p2 transcript was still detected in mRNAs prepared from strains with Glc52, Glc60-63 *nagC*, Glc64, and Glc65 mutations but not from the Glc60-63 mutation in the absence of *nagC* (Fig. 4A, lanes 3 to 9). The derepression of the p2 transcript by the *nagC* mutation confirmed that the Glc60-63 mutant is now part of the NagC regulon. Although the p1 promoter was essentially inactive, the BoxP1 operator is essential for the regulation of the p2 promoter by NagC. Analysis of mRNA from the strain carrying the Glc65 mutation (6 bp in the central CGCG pattern in BoxP1) showed that the p2 transcript is detectable at similar levels in the presence or absence of the *mlc* mutation (Fig. 4A, lanes 8 and 9), which is consistent with similar β-galactosidase activities in the three Glc65 strains, i.e., the WT and *mlc* and *nagC* mutant strains (Fig. 2B, section C). The p2 transcript is thus derepressed by the presence of the Glc65 mutation and probably accounts for most of the β-galactosidase activity detected.

NagC and Mlc binding to mutated operators. DNase I footprinting was used to measure the effects of the mutations in the BoxP1 and BoxP2 operators on DNA binding in vitro. The WT operators bound Mlc with high affinity (Fig. 5, lanes 5 to 8) but were only weakly protected by NagC (lanes 2 to 4). On the other hand, the fragment carrying the Glc60-63 mutations was no longer capable of binding Mlc (Fig. 5, lanes 14 to 16) but bound NagC with higher affinity (Fig. 5, lanes 9 to 13).

The effects of the two types of mutations (CGCG in the central region and +/−11C/G) on the specificity of Mlc and NagC binding were examined by measuring the binding of Mlc and NagC to the individual mutated operators by quantitative DNase I footprinting (Table 1). Mutations in either the central region (Glc64 or Glc65) or C/G at positions +/−11 (Glc52) decreased the binding of Mlc but had hardly any effect on NagC binding. The combination of +/−11CG with the central CGCG pattern (operators Glc60 and Glc63) produced strong NagC binding sites and eliminated Mlc binding (Table 1).

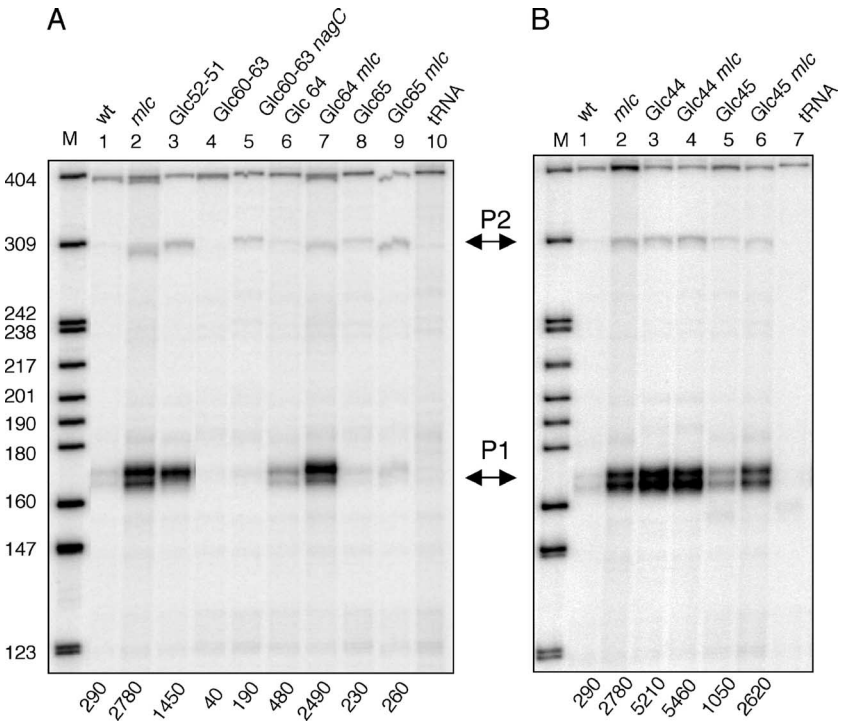


FIG. 4. S1 analysis of mRNA from lysogenic bacteria carrying different operator mutations. (A and B) The probe used is the Glc1-Lac21 fragment specific to the *ptsG-lacZ* fusion (Fig. 2A). Total RNAs (30 μg) prepared from the different WT and mutant strains indicated at the top of each lane were hybridized to the probes, and S1-resistant transcripts were analyzed on a denaturing gel. The p1 and p2 transcripts are indicated. The values below the lanes give the corresponding β-galactosidase activities (Miller units) of the lysogens. The mutations Glc52-51, Glc60-63, Glc64, and Glc65 are listed in Fig. 2. The mutation Glc44 is −7G and −6C in BoxP1, and the mutation Glc45 is −6C and −5C in BoxP2. Lanes M contain markers (pBR322 digested with MspI) whose molecular sizes are shown in base pairs on the left.

NagC binding to Glc60 and Glc63 operators is comparable to the binding of NagC to the WT *nagE-B* operator-carrying fragments, BoxE and BoxB. Thus, as shown by the β-galactosidase assays, mutations either at positions +/−11 or in the central region GCGC pattern reduce Mlc binding but without allowing NagC binding and the two types of mutations, +/−11CG and a central CGCG pattern, are necessary to allow strong NagC binding.

Cooperative binding of Mlc to the *ptsG* operators. The WT p2 promoter contributes no more than 10% to the total *ptsG* expression (26). The fact that some of the mutations in BoxP2 produced substantial increases in *ptsG-lacZ* expression, corresponding to more than 10% of the total *ptsG* expression (e.g., Glc52 and Glc60), strongly suggested that loss of Mlc binding to the operator BoxP2 leads to at least partial derepression of p1, which in turn implies that Mlc is capable of regulating expression from the two promoters cooperatively. Even the Glc64 mutation (4 bp in the CGCG pattern in the central region of BoxP2), which produces just a twofold increase in *ptsG-lacZ* expression, enhanced expression from p1, as shown by S1 analysis (Fig. 4A, lane 6).

To examine the cooperative binding by Mlc in the absence of any possible effects of NagC, we analyzed the effects of mutations in the totally conserved TT at positions −5 and −6 of the two operators. Mutations at these positions prevent the binding of either Mlc or NagC. S1 analysis of mRNAs prepared from the fusion strains was used to assess the relative levels of expression from p1 and p2. The mutation Glc44 (−7G and −6C in BoxP1) produced complete derepression of the p1 and p2 promoters since the transcript levels, as well as the β-galactosidase activities, are unchanged by the presence of the *mlc* mutation (Fig. 4B, lanes 3 and 4). This mutation actually enhances the p1 promoter activity since it produces an extended "−10 promoter" sequence (17) (*TGCTACTCT* [the extended −10 sequence is in italics, and the −7 and −6 positions of the BoxP1 operator changed by mutagenesis are underlined]). The additional contact between the promoter and $E\sigma^{70}$ afforded by TG 2 and 3 bp upstream of the −10 consensus should account for the nearly twofold higher β-galactosidase activities produced by the Glc44 *ptsG-lacZ* fusion (Fig. 4B). Significantly, the p2 transcript from the Glc44 mutant is identical in intensity to the p2 transcript from the WT fusion in the *mlc* mutant strain (lane 2), showing that preventing Mlc binding to BoxP1 fully derepresses p2. The Glc45 mutation (−6C and −5C in BoxP2) produces partial derepression of the p1 promoter but complete derepression of the p2 promoter, as shown by comparison with the same fusion in the presence of the *mlc* mutation (Fig. 4B, lanes 5 and 6), although p2 expression from Glc45 actually looks slightly weaker than from the WT fusion in the *mlc* mutant strain.

The fact that the BoxP1 mutation produces complete derepression of p2 but the BoxP2 mutation only produces partial derepression of p1 implies that the two operators are not equal. The BoxP1 operator can thus function by itself and bind Mlc, producing a regulation factor of about 2.5-fold. This value is comparable in magnitude to the regulation observed for

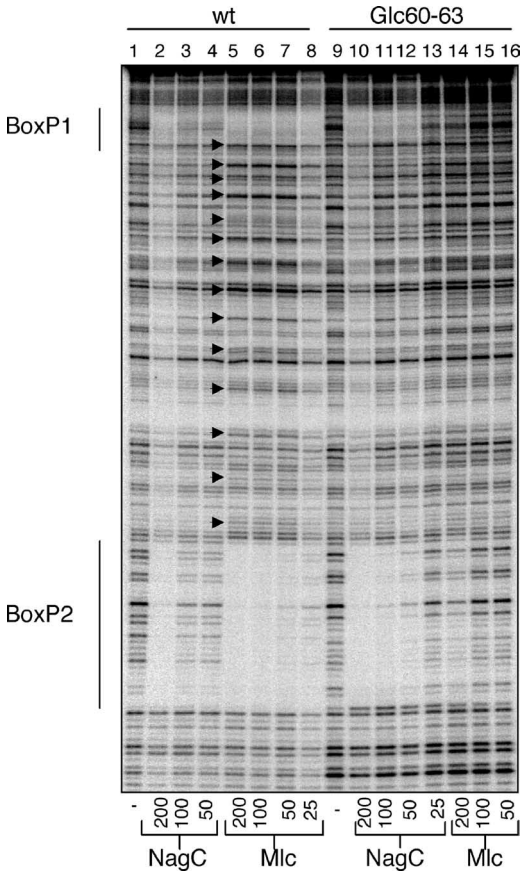


FIG. 5. Mlc and NagC binding to mutated operators. The Glc1-4 PCR fragment (Fig. 2A) carrying either WT or Glc60-63 operators labeled at Glc1 was incubated with various dilutions of Mlc or NagC, as indicated, at room temperature for 10 min before digestion with DNase I. The products were analyzed on a 6% denaturing acrylamide gel. The locations of BoxP1 and BoxP2 are indicated. Note the pattern of hypersensitive cleavages (arrowheads) separated by about 10 bp between BoxP1 and BoxP2 on the WT fragment with Mlc.

manXYZ, *malT*, and *mlc*, which are regulated by Mlc from one operator (4, 24). On the other hand, mutations in BoxP1 (e.g., Glc44, -51, or -65) produce complete derepression of p2. The different effects of mutations in the two operators is reflected in the binding affinities of the two operators for Mlc; BoxP1 exhibits a higher affinity for Mlc than does BoxP2 (Table 1). The presence of the second BoxP2 operator increases Mlc binding to both sites so that regulation occurs over a greater range (10-fold). This is similar to the situation with the *lac* operon, where the presence of two subsidiary operators, *O2* and *O3*, enhances regulation from the primary operator, *O1* (19). Cooperative binding of Mlc to the two operators is expected to enhance binding to both sites, but this is not detectable on the linear fragments of DNA used for footprinting. However, interestingly, binding of Mlc to the WT *ptsG* fragment produced some indication of enhanced DNase I cleavages in the DNA between the two operators (indicated by arrowheads in Fig. 5), which are indicative of DNA bending. This is reminiscent of the pattern of strongly hypersensitive DNase I cleavages detected after NagC binding to the linear *nagE-B* operator-carrying DNA, which are diagnostic of DNA bending and loop formation between operator-bound tetramers of NagC (29). Footprinting on fragments carrying a WT BoxP2 operator and a mutated BoxP1 operator showed that Mlc still bound to BoxP2 but the hypersensitive cleavages were lost (data not shown), confirming that they are the result of a deformation of the DNA due to Mlc binding to both operators. Both Mlc and NagC are tetramers in solution (18, 37), which is consistent with their binding cooperatively to two sites. We can note that there is no evidence of the formation of heteromeric proteins. The construct with one high-affinity NagC site at BoxP1 and the WT operator at BoxP2 (Glc63) was completely derepressed.

DISCUSSION

The results reported here show that, despite the relative similarity in sequence between the Mlc and NagC operators, there is a high level of specificity to each operator and extensive changes are required to convert the Mlc operators upstream of the *ptsG* gene to operators recognized by NagC. Moreover, loss of Mlc binding was rather easily achieved, again implying a high level of specificity. Two potential identifying characteristics of Mlc and NagC sites have been investigated, i.e., the identity of the bases at +/−11 from the center of symmetry and the distribution of CG base pairs in the central region. The experiments described here show that both characteristics are crucial for Mlc or NagC binding and, moreover, that to observe regulation of *ptsG* by NagC in vivo, it is necessary to have both operators changed to NagC-specific sites. This latter result is consistent with the facts that all of the genes controlled by NagC possess two operators and mutations in one operator are sufficient to lose regulation (*nagE-B*, *glmU*, *fimB*) (23, 29, 40).

Three types of DNA-protein contacts can be envisaged between Mlc or NagC and its DNA target. Firstly, there should be base-specific contacts between amino acids of the recognition helix and the major groove of the palindromic operators plus two other types of interaction which allow recognition of the base pair at +/−11 and the sequence of CG base pairs within the central region. For both proteins, the totally conserved TT/AA bases at positions −5, −6, +5, and +6 are essential for binding (23, 29, 40) and are likely to be the direct

TABLE 1. Relative affinities of Mlc and NagC for binding to WT and mutant *ptsG* operators[a]

Mutation(s)	Operator	K_d		Operator	K_d	
		Mlc	NagC		Mlc	NagC
None	WT BoxP2	14	100	WT BoxP1	6	100
+/−11CG	Glc52	300	100			
Central CGCG	Glc64	65	80	Glc65	80	100
+/−11CG, central CGCG	Glc60	400	20	Glc63	>800	13
None	BoxE	30	18	BoxB	>200	10

[a] Estimations of the dissociation constants of Mlc or NagC from the WT and mutant *ptsG* operators were made by quantitative DNase I footprinting (see Materials and Methods). The values are the concentrations (nanomolar monomer) giving 50% protection. Values are the mean of at least two gels, and errors in reproducibility are on the order of 30%. The lower values are probably overestimates since they are comparable to the concentration of DNA in the assay (3 to 10 nM). BoxE and BoxB are the NagC operators of the *nagE-BACD* operons.

targets of specific amino acid recognition. Mlc and NagC possess H-T-H motifs in their N-terminal domains (34) which are almost identical in the two proteins. In particular, the amino acids at positions 1, 2, and 6 in the recognition helix are identical in Mlc and NagC (Pro1, Ala2, and Lys6, Fig. 1B), These are the amino acids which, by comparison with the genetic, nuclear magnetic resonance, and crystallographic data for other H-T-H proteins like LacI, PurR, and CAP, are likely to be involved in specific DNA contacts in the major groove of the DNA (1, 15, 21, 33, 36, 41).

The other two types of protein-operator interaction investigated in this work are not identical in the two proteins. Previous results implicated a C or a G at positions +/−11 in NagC binding (30). In the present work, replacement of the A/T at positions +/−11 of the *ptsG* operators is sufficient to eliminate regulation by Mlc, but it is not sufficient to produce strong binding by NagC. Positions +/−11 are one turn of B-form DNA away from the center of symmetry of the operator. Assuming that Mlc and NagC bind to DNA in a way similar to that of CRP or LacI, with the minor groove at the center of symmetry of the operator, facing the twofold axis of the dimeric repressor protein, then the minor groove at positions +/−11 should also be facing the protein. This is consistent with the fact that a CG or GC base pair at +/−11 is equally effective at enhancing the affinity of NagC (30) because CG base pairs are essentially indistinguishable from GC base pairs in the minor groove but are distinguishable from AT base pairs because of the presence of the exocyclic N_2 of guanine (42). Mutations in the AT-rich sequences around positions 9 to 12 from the center of symmetry of the CAP operator produce large changes in CAP binding and bending, even though they lie outside the CAP canonical consensus sequence (3, 7). This is the region of the secondary kink in the CAP-DNA cocrystals, and several amino acids contact the phosphate backbone in this region (21). The presence of a CG base pair should change both the kink and backbone contacts.

The hypothesis that NagC prefers to bind to a site with a central CGCG pattern, as suggested by the sequence alignments of known operators (Fig. 1A), is supported by the present results, although which positions are the most important has not been clarified. Although mostly outside the consensus −35 sequence of p2 and the −10 sequence of p1, the mutations have considerably reduced *ptsG* promoter activity. This effect on promoter strength, together with the fact that the overall factors of regulation of the *ptsG* gene by Mlc is maximally 10-fold, has limited our ability to more precisely analyze the individual contributions of positions 1 to 4 (as was done in the seminal work by the Müller-Hill laboratory on the *lac* operator, where factors of regulation are greater than 200-fold [15, 33]).

The role of this central region (positions −4 to +4) in Mlc or NagC recognition is difficult to assess. It could be that certain positions are implicated in direct readout interactions with amino acids of the H-T-H domain of Mlc or NagC, but this is not usually observed for the central base pair positions in other protein-DNA complexes. Rather, the centers of palindromic operators are often more or less bent, thus enabling the recognition helix of the H-T-H to slip or dip more precisely into the major groove (see reference 47 for a recent example). It is a common observation that proteins produce substantially stronger curvature by exploiting inherent bends present in the free DNA (2). For example, a strong correlation between DNA flexibility and the binding affinity of papillomavirus E2 proteins to consensus sites with various central linker sequences has been demonstrated both experimentally and by molecular dynamic simulations (5, 49). Both CG and GC base pairs produce flexible dinucleotide steps, as calculated from nuclear magnetic resonance structures in solution or derived from analysis of X-ray structures (9, 20), and thus are easily deformed upon protein binding. However, the distribution of CG base pairs in the central region of Mlc and NagC targets is different, implying changes, possibly subtle, in DNA structure which the corresponding proteins could exploit to optimize alternative protein-DNA interfaces. We have recently shown, by the use of chimeric proteins, that the H-T-H regions of Mlc and NagC, by themselves, are not sufficient for normal specific Mlc or NagC binding and that the rest of the native protein is necessary for full DNA binding specificity. These studies suggested that the linker between the H-T-H domain and the rest of the protein could be implicated in operator specificity (22). By analogy with the structures of DNA bound to LacI or PurR, we could propose that binding of Mlc or NagC involves amino acid contacts from outside the H-T-H domain, possibly from this unstructured linker, which could form the equivalent of the hinge helix of the LacI- or PurR-operator complex. Another possible analogy is with the homeobox-containing Hox family proteins Scr and Exd. They make essentially identical contacts in the major groove, but specific binding of Scr and Exd to the *fkh* operator is achieved by two basic amino acids from Scr contacting an unusually narrow minor groove. The narrow minor groove is dictated by the DNA sequence, but specific interaction with Scr involves recognition of the DNA structure rather than a DNA sequence (10). In the same way, the pattern of a CG or GC base pair could influence the structure of the DNA at the center of the operator and hence the preference for Mlc or NagC by such an indirect readout mechanism.

The work described here has not resolved the paradox of the in vitro-selected Mlc sites, which almost exclusively conformed to the NagC preferred central CGCG pattern and in some cases had a C or a G at positions +/−11 (25). The apparently different protein binding properties of these SELEX sequences compared to the NagC binding sites created in the present work need to be compared under identical conditions, i.e., within the context of the fixed 100-bp sequence of the SELEX fragments or by replacing the BoxP1 and BoxP2 operators within the *ptsG* regulatory region with these selected operators. It is possible that the sequence context of the operators also affects their ability to bind and regulate.

Several members of the same protein family are frequently found within the same organism, but normally the orthologues have evolved with specific functions; e.g., there are at least a dozen members of the LacI-GalR family in *E. coli* and they are mostly repressors for the use of different carbon sources (45). Although there are similarities in both of the consensus operator sites for these proteins and their DNA binding H-T-H motifs, certain precise changes are present especially in positions 1 and 2 of their recognition helix, which are the positions most strongly implicated in DNA recognition. In particular, base pair changes within the operators are capable of switching the recognition specificity from, e.g., LacI- to GalR-type bind-

ing (15). GalR and GalS, on the other hand, are isorepressors of the Gal operon and are expected to bind to very similar sites, and their H-T-H motifs are almost identical (8, 46). Even in this case, the reported affinities of the two proteins for the different *gal* target operons vary (8). The GlnR and TnrA transcription factors, necessary for nitrogen regulation in *Bacillus subtilis*, also have almost identical H-T-H DNA binding motifs and bind to targets with the same consensus sequence. However, their C-terminal domains, involved in signal transduction, are quite different, so that the proteins are active under different conditions, and it is thought that amino acids of the β-loop wing, adjacent to the H-T-H motif, contribute to the DNA binding site specificity of these two transcription factors (6, 48). Mlc and NagC seem to resemble a diverging isorepressor pair, where operator specificity has been achieved without altering the canonical H-T-H interaction, and specific inducing signal recognition occurs, even though the C-terminal domains are still homologous (22). It is interesting that Mlc and NagC control the expression of two PTS transporters, PtsG and NagE, which are themselves homologues. It is tempting to imagine that the two hexose transporters and their respective transcriptional regulators have been derived by gene duplication of an ancestral pair of genes for a transporter and its transcriptional repressor. The *nagC* gene is still associated with the *nagE-BACD* operon on the chromosome of *E. coli*, whereas the *mlc* and *ptsG* genes are now dispersed and both regulators have acquired new targets and specificities.

ACKNOWLEDGMENTS

We thank Sabine Seitz, Olivier Pellgrini, and Mitchell Lewis for gifts of Mlc and NagC proteins; Annie Kolb, Brigitte Hartmann, and Harald Putzer for useful discussions and critical comments on the manuscript; and Carole Pennetier for help in the initial stages of this work.

This work was supported by grants from the CNRS and Université Paris 7-Denis Diderot to UPR 9073.

REFERENCES

1. **Bell, C., and M. Lewis.** 2001. Crystallographic analysis of Lac repressor bound to natural operator *O1*. J. Mol. Biol. **312:**921–926.
2. **Crothers, D.** 1998. DNA curvature and deformation in protein-DNA complexes: a step in the right direction. Proc. Natl. Acad. Sci. USA **95:**15163–15165.
3. **Dalma-Weiszhausz, D., M. Gartenberg, and D. Crothers.** 1991. Sequence-dependent contribution of distal binding domains to CAP protein-DNA binding affinity. Nucleic Acids Res. **19:**611–616.
4. **Decker, K., J. Plumbridge, and W. Boos.** 1998. Negative transcriptional regulation of a positive regulator: the expression of *malT*, encoding the transcriptional activator of the maltose regulon of *Escherichia coli*, is negatively controlled by Mlc. Mol. Microbiol. **27:**381–390.
5. **Djuranovic, D., and B. Hartmann.** 2005. Molecular dynamics studies on free and bound targets of the bovine papillomavirus type I E2 protein: the protein binding effect on DNA and the recognition mechanism. Biophys. J. **89:**2542–2551.
6. **Fisher, S.** 1999. Regulation of nitrogen metabolism in *Bacillus subtilis*: vive la différence. Mol. Microbiol. **32:**223–232.
7. **Gartenberg, M. R., and D. M. Crothers.** 1988. DNA sequence determinants of CAP-induced bending and protein affinity. Nature **333:**824–829.
8. **Geanacopoulos, M., and S. Adhya.** 1997. Functional characterization of roles of GalR and GalS as regulators of the *gal* regulon. J. Bacteriol. **179:**228–234.
9. **Heddi, B., N. Foloppe, N. Bouchemal, E. Hantz, and B. Hartmann.** 2006. Quantitation of DNA BI/BII backbone states in solution. Implications for DNA overall structure and recognition. J. Am. Chem. Soc. **128:**9170–9177.
10. **Joshi, R., J. Passner, R. Rohs, R. Jain, A. Sosinsky, M. Crickmore, V. Jacob, A. Aggarwal, B. Honig, and R. Mann.** 2007. Functional specificity of a Hox protein mediated by the recognition of minor groove structure. Cell **131:**530–543.
11. **Kim, S.-Y., T.-W. Nam, D. Shin, B.-M. Koo, Y.-J. Seok, and S. Ryu.** 1999. Purification of Mlc and analysis of its effects on the *pts* expression in *Escherichia coli*. J. Biol. Chem. **274:**25398–25402.
12. **Kimata, K., T. Inada, H. Tagami, and H. Aiba.** 1998. A global repressor (Mlc) is involved in glucose induction of the *ptsG* gene encoding major glucose transporter in *Escherichia coli*. Mol. Microbiol. **29:**1509–1519.
13. **Kunkel, T. A.** 1985. Rapid and efficient site-specific mutagenesis without phenotypic selection. Proc. Natl. Ac. Sci. USA **82:**488–492.
14. **Lee, S.-J., W. Boos, J.-P. Bouché, and J. Plumbridge.** 2000. Signal transduction between a membrane-bound transporter, PtsG, and a soluble transcription factor, Mlc, of *Escherichia coli*. EMBO J. **19:**5353–5361.
15. **Lehming, N., J. Sartorius, B. Kisters-Woike, B. von Wilcken-Bergmann, and B. Müller-Hill.** 1990. Mutant *lac* repressors with new specificities hint at rules for protein-DNA recognition. EMBO J. **9:**615–621.
16. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
17. **Mitchell, J., D. Zheng, S. Busby, and S. Minchen.** 2003. Identification and analysis of "extended-10" promoters in *Escherichia coli*. Nucleic Acids Res. **31:**4689–4695.
18. **Nam, T.-W., S.-H. Cho, D. Shin, J.-H. Kim, J.-Y. Jeong, J.-H. Lee, J.-H. Roe, A. Peterkofsky, S.-O. Kang, S. Ryu, and Y.-J. Seok.** 2001. The *Escherichia coli* glucose transporter enzyme $IICB^{Glc}$ recruits the global repressor Mlc. EMBO J. **20:**491–498.
19. **Oehler, S., E. R. Eismann, H. Krämer, and B. Müller-Hill.** 1990. The three operators of the *lac* operon cooperate in repression. EMBO J. **9:**973–979.
20. **Packer, M., M. Dauncey, and C. Hunter.** 2000. Sequence-dependent DNA structure: tetranucleotide conformational maps. J. Mol. Biol. **295:**85–103.
21. **Parkinson, G., C. Wilson, A. Gunasekera, Y. W. Ebright, R. E. Ebright, and H. M. Berman.** 1996. Structure of the CAP-DNA complex at 2.5Å resolution: a complete picture of the protein-DNA interface. J. Mol. Biol. **260:**395–408.
22. **Pennetier, C., L. Dominguez-Ramirez, and J. Plumbridge.** 2008. Different regions of Mlc and NagC, homologous transcriptional repressors controlling expression of the glucose and *N*-acetylglucosamine phosphotransferase systems in *Escherichia coli*, are required for inducer signal recognition. Mol. Microbiol. **67:**364–377.
23. **Plumbridge, J.** 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. EMBO J. **14:**3958–3965.
24. **Plumbridge, J.** 1998. Control of the expression of the *manXYZ* operon in *Escherichia coli*: Mlc is a negative regulator of the mannose PTS. Mol. Microbiol. **27:**369–381.
25. **Plumbridge, J.** 2001. DNA binding sites for the Mlc and NagC proteins: regulation of *nagE*, encoding the *N*-acetylglucosamine-specific transporter in *Escherichia coli*. Nucleic Acids Res. **29:**506–514.
26. **Plumbridge, J.** 1998. Expression of *ptsG*, the gene for the major glucose PTS transporter in *Escherichia coli*, is repressed by Mlc and induced by growth on glucose. Mol. Microbiol. **29:**1053–1063.
27. **Plumbridge, J.** 1999. Expression of the phosphotransferase system (PTS) both mediates and is mediated by Mlc regulation in *Escherichia coli*. Mol. Microbiol. **33:**260–273.
28. **Plumbridge, J.** 1991. Repression and induction of the *nag* regulon of *Escherichia coli* K-12: the roles of *nagC* and *nagA* in maintenance of the uninduced state. Mol. Microbiol. **5:**2053–2062.
29. **Plumbridge, J., and A. Kolb.** 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.
30. **Plumbridge, J., and A. Kolb.** 1995. Nag repressor-operator interactions: protein-DNA contacts cover more than two turns of the DNA helix. J. Mol. Biol. **249:**890–902. (Erratum: J. Mol. Biol. **253:**219–220.)
31. **Plumbridge, J., and O. Pellegrini.** 2004. Expression of the chitobiose operon of *Escherichia coli* is regulated by three transcription factors: NagC, ChbR and CAP. Mol. Microbiol. **52:**437–449.
32. **Powell, B., D. Court, Y. Nakamura, M. Rivas, and C. Turnbough.** 1994. Rapid confirmation of single copy lambda prophage integration by PCR. Nucleic Acids Res. **22:**5765–5766.
33. **Sartorius, J., N. Lehming, B. Kisters, B. von Wilcken-Bergmann, and B. Müller-Hill.** 1989. *lac* repressor mutants with double or triple exchanges in the recognition helix bind specifically to *lac* operator variants with multiple exchanges. EMBO J. **8:**1265–1270.
34. **Schiefner, A., K. Gerber, S. Seitz, W. Welte, K. Diederichs, and W. Boos.** 2005. The crystal structure of Mlc, a global regulator of sugar metabolism in *Escherichia coli*. J. Biol. Chem. **280:**29073–29079.
35. **Schneider, T., and R. Stephens.** 1990. Sequence logos: a new way to display consensus sequences. Nucleic Acids Res. **18:**6097–6100.
36. **Schumacher, M. A., K. Y. Choi, H. Zalkin, and R. G. Brennan.** 1994. Crystal structure of LacI member, PurR, bound to DNA: minor groove binding by α helices. Science **266:**763–770.
37. **Seitz, S., S.-J. Lee, C. Pennetier, W. Boos, and J. Plumbridge.** 2003. Analysis of the interaction between the global regulator Mlc and $EIIB^{Glc}$ of the glucose-specific phosphotransferase system in *Escherichia coli*. J. Biol. Chem. **278:**10744–10751.
38. **Simons, R. W., F. Houman, and N. Kleckner.** 1987. Improved single and

multicopy *lac*-based cloning vectors for protein and operon fusions. Gene **53:**85–96.
39. **Sohanpal, B., S. El-Labany, M. Lahooti, J. Plumbridge, and I. Blomfield.** 2004. Integrated regulatory responses of *fimB* to *N*-acetylneuraminic (sialic) acid and GlcNAc in *Escherichia coli* K-12. Proc. Natl. Acad. Sci. USA **101:**16322–16327.
40. **Sohanpal, B. K., S. Friar, J. Roobol, J. A. Plumbridge, and I. C. Blomfield.** 2007. Multiple co-regulatory elements and IHF are necessary for the control of fimB expression in response to sialic acid and *N*-acetylglucosamine in *E. coli* K-12. Mol. Microbiol. **63:**1223–1236.
41. **Spronk, C. A., A. M. Bonvin, P. K. Radha, G. Melacini, R. Boelens, and R. Kaptein.** 1999. The solution structure of Lac repressor headpiece 62 complexed to a symmetrical lac operator. Structure **7:**1483–1492.
42. **Steitz, T. A.** 1990. Structural studies of protein-nucleic acid interaction: the sources of sequence specific binding. Q. Rev. Biophys. **23:**205–280.
43. **Tanaka, Y., K. Kimata, and H. Aiba.** 2000. A novel regulatory role of glucose transporter of *Escherichia coli*: membrane sequestration of a global repressor Mlc. EMBO J. **19:**5344–5352.
44. **Titgemeyer, F., J. Reizer, A. Reizer, and M. H. Saier.** 1994. Evolutionary relationships between sugar kinases and transcriptional repressors in bacteria. Microbiology **140:**2349–2354.
45. **Weickert, M. J., and S. Adhya.** 1992. A family of bacterial regulators homologous to Gal and Lac repressors. J. Biol. Chem. **267:**15869–15874.
46. **Weickert, M. J., and S. Adhya.** 1993. The galactose regulon of *Escherichia coli*. Mol. Microbiol. **10:**245–251.
47. **White, C., and S. Winans.** 2007. The quorum-sensing transcription factor TraR decodes its DNA binding site by direct contacts with DNA bases and by detection of DNA flexibility. Mol. Microbiol. **64:**245–256.
48. **Zalieckas, J. M., L. V. Wray, and S. H. Fisher.** 2006. Cross-regulation of the *Bacillus subtilis glnRA* and *tnrA* genes provides evidence for DNA binding site discrimination by GlnR and TnrA. J. Bacteriol. **188:**2578–2585.
49. **Zhang, Y., Z. Xi, R. Hegde, Z. Shakked, and D. Crothers.** 2004. Predicting indirect readout effects in protein-DNA interactions. Proc. Natl. Acad. Sci. USA **101:**8337–8341.