

Role of integration host factor in the regulation of the *glnHp2* promoter of *Escherichia coli*

(transcription/ σ^{54} /NR_I/nitrogen regulation)

FELIX CLAVERIE-MARTIN AND BORIS MAGASANIK

Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

Contributed by Boris Magasanik, December 14, 1990

ABSTRACT The *glnHPQ* operon of *Escherichia coli* encodes components of the high-affinity glutamine transport system. One of the two promoters of this operon, *glnHp2*, is responsible for expression of the operon under nitrogen-limiting conditions. The general nitrogen regulatory protein (NR_I) binds to two overlapping sites centered at –109 and –122 from the transcription start site and, when phosphorylated, activates transcription of *glnHp2* by catalyzing isomerization of the closed σ^{54} -RNA polymerase promoter complex to an open complex. The DNA-bending protein integration host factor (IHF) binds to a site immediately upstream of *glnHp2* and enhances the activation of open complex formation by NR_I phosphate. The NR_I-binding sites can be moved several hundred base pairs further upstream without altering the ability of NR_I phosphate to activate open complex formation. However, in this case, IHF diminishes open complex formation. We propose that the IHF-induced bend can facilitate or obstruct the interaction between NR_I phosphate and the closed complex depending on the relative positions of NR_I phosphate and σ^{54} -RNA polymerase on the DNA.

Activation of transcription of nitrogen-regulated operons in *Escherichia coli* and other enteric bacteria requires RNA polymerase σ^{54} ($E\sigma^{54}$) and the *glnG* gene product NR_I in its phosphorylated active form (1, 2). In the well-studied *glnAp2* promoter, NR_I phosphate binds to upstream enhancer-like sequences and catalyzes the isomerization of a preexisting $E\sigma^{54}$ -promoter closed complex to the open form (3, 4). Recent studies have provided evidence that bound NR_I phosphate makes contacts with the $E\sigma^{54}$ -promoter closed complex through the formation of a DNA loop (5, 6).

Transcription of the *nif* operons of *Klebsiella pneumoniae* is also dependent on $E\sigma^{54}$ (7). The activator protein NIFA also binds to upstream sites (8, 9) and is functionally and structurally similar to NR_I (10, 11). In contrast to NR_I, NIFA has not been purified in active form (12, 13). Integration host factor (IHF) binds just upstream from the *nifH* and *nifU* promoters and stimulates NIFA-mediated activation (13–15). IHF is a sequence-specific DNA-bending protein, which is involved in gene expression and other processes in *E. coli* and some of its bacteriophages and plasmids (16).

The *glnHPQ* operon of *E. coli*, which encodes the components of the high-affinity glutamine transport system, is among the operons whose expression is induced under nitrogen-limiting conditions (17, 18). A promoter with homology to the σ^{54} promoters, *glnHp2*, has been identified (19). In this study, we present evidence for the existence of overlapping binding sites for NR_I upstream from the *glnHp2* promoter. We also found that IHF binds between the *glnHp2* promoter and the NR_I binding sites. This system allowed us

to study the role of IHF in the activation of transcription by NR_I by using purified components.

MATERIALS AND METHODS

Proteins, Primers, and Materials. Core RNA polymerase, σ^{54} , NR_I, and NR_{II} were purified as described (1, 20, 21). IHF was a gift from C. Robertson and H. Nash (National Institutes of Health). Primers FC5 (5'-CCACATCATCACA-CAATCG-3'), FC6 (5'-CAGACTTCATAGCATTTCC-3'), and FC7 (5'-GCATCTTCAGGGTATTGCC-3') hybridizing at –217, +50, and –103 (5' position), respectively, and primer FC1 (5'-GCGAGAGATATTCGTGG-3'), which hybridizes to T7 sequences close to the *Hind*III site of plasmid pTE103 (22), were synthesized at the Biopolymers Laboratory, Howard Hughes Medical Institute, Massachusetts Institute of Technology. The following materials were used: DNase I, *Mae* II, alkaline phosphatase, and bovine serum albumin, from Boehringer Mannheim; Klenow and other restriction endonucleases or DNA modifying enzymes, from New England Biolabs; radiolabels and Protosol, from Dupont/NEN; ultrapure solution ribonucleotides and Sephadex G-25, from Pharmacia LKB.

Construction of Plasmids. All transcription templates were derived from plasmid pTE103, which contains the multicloning site from pUC8 placed upstream from a bacteriophage T7 transcriptional terminator (22). Plasmid pFC50 was constructed by inserting the 540-base-pair (bp) *EcoRV/Sac* II fragment from pTN240 (18), into the *Sma* I site of pTE103. The sticky ends of this fragment and of those mentioned below were made blunt by using T4 DNA polymerase. The 540-bp fragment contains the *glnHp2* promoter with the upstream regulatory sequences and 66 nucleotides of the *glnH* coding region. Plasmid pFC54 was constructed by inserting the 180-bp *Mae* II/*Sac* II fragment from pTN240 (see position of the *Mae* II site in Fig. 4), which contains *glnHp2* and the IHF binding site but not the NR_I binding sites, into the *Hinc*II site of pTE103. Plasmid pFC55b was constructed by inserting the 360-bp *EcoRV/Mae* II fragment from pTN240, containing the NR_I binding sites of the *glnHp2* promoter region (see Fig. 4), into the *Sma* I site of pTE103. The orientation of the 360-bp fragment in pFC55b is opposite that in pFC50. Plasmid pFC57 was constructed by ligating the 180-bp *Bam*HI/*Hind*III fragment from pFC54 to plasmid pFC55b cut with *Bam*HI and *Hind*III. The NR_I binding sites in plasmid pFC57 are 273 bp further upstream than in pFC50. Plasmids were purified with a "mini" kit from Qiagen (Studio City, CA). The orientation of the fragments in the recombinant plasmids was confirmed by DNA sequencing with a Sequenase kit (United States Biochemical). Transcription at *glnHp2* on plasmids pFC50, pFC54, and pFC57 would generate transcripts of 418 nucleotides.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: IHF, integration host factor; DMS, dimethyl sulfate.

Transcription Assays. Transcriptions were as described (23). Plasmid DNA was purified by centrifugation in CsCl/EtdBr gradients and DNA concentrations were determined by absorbance at 260 nm. The concentration of DNA templates in all experiments was 5 nM. Core RNA polymerase, σ^{54} , NR_I, and IHF were present at 25, 100, 15, and 50 nM, respectively, or as indicated. NR_I was added at the indicated concentrations. The same radiolabeled nucleotide ($[\alpha\text{-}^{32}\text{P}]\text{UTP}$; 3000 Ci/mmol; 10 mCi/ml; 1 Ci = 37 GBq) was used in all the experiments. Electrophoresis of RNA in urea/acrylamide gels was as described (1). After autoradiography, bands were cut out from the dried gels and placed in vials containing 5 ml of Scinti VerseTM I (Fisher Scientific) and 0.3 ml of Protosol. Vials were incubated overnight at 37°C and assayed in a 1211 Rackbeta liquid scintillation counter (LKB).

Footprinting. DNA footprinting experiments were carried out on supercoiled DNA as described (24). Plasmid pFC50 was incubated with the proteins under the same conditions used for transcription. After 20 min of incubation at 37°C, the DNA was treated with DNase I, dimethyl sulfate (DMS), or potassium permanganate as described (refs. 24 and 25; S. Sasse-Dwight and J. D. Gralla, personal communication). Samples treated with DNase I or potassium permanganate were extracted with phenol/chloroform. DMS-treated DNA was cleaved with 1 M piperidine. All samples were then passed through Sephadex G-25 spin columns equilibrated in water; 5×10^5 cpm of ^{32}P -end-labeled FC5, FC6, or FC7 primer was added per reaction mixture. Each sample was analyzed by primer-extension analysis using alkaline denaturation as described (24, 26). Hybridizations to the FC5, FC6, and FC7 primers were carried out at 54°C, 52°C, and 55°C, respectively. Reaction products recovered by ethanol precipitation were analyzed together with dideoxynucleotide sequencing reaction products on 7% polyacrylamide gels containing 7 M urea.

RESULTS

Binding Sites for IHF and NR_I. Examination of the regulatory region of the *glnHPQ* operon (ref. 18; see Fig. 4) shows the presence of two putative IHF binding sites, one immediately upstream from the *glnHP2* promoter (-46 to -34; see Fig. 4) identical to the IHF DNA binding consensus (5'-A/TATCAAN₄TTA/G-3') (16, 27) and another one further upstream (-158 to -146) with one mismatch. DNase I protection patterns of IHF on both strands of the *glnHP2* promoter region were obtained (Fig. 1A, top strand; Fig. 1B, bottom strand). IHF protected the site close to *glnHP2* but not the other site. The protected region is extensive, approximately from -60 to -24 on both strands (Fig. 1A and B), considering the small size of the IHF dimer (M_r 20,000) (16). We also probed specific interactions between IHF and *glnHP2* with DMS, a reagent that does not cleave the DNA backbone. Inhibition or enhancement of methylation can then be detected after chemical cleavage of the modified bases. IHF-dependent protection of the G residues at positions -50, -40, and -34 was observed (Fig. 2B, lanes 4 and 5). Furthermore, in the presence of IHF the reactivity of G residues at positions -57 and -39 and of A residues at -47 and -48 was enhanced.

It has been shown that sequences upstream from *glnHP2* play a role in the inducible expression of the *glnHPQ* operon under nitrogen-limiting conditions (18). By inspection of this region, we found four sequences that fit the consensus NR_I binding site 5'-TGCACCAN₃TGGTGCA-3' (see Fig. 4). Sites 3 and 2 overlap and have only one and two mismatches, respectively. Sites 1 and 4 deviate from consensus by 3 bp, and the spacing in site 4 is four instead of three. Sites 1 and 2 also overlap. We carried out DNase I footprinting experiments to determine which sites are occupied by NR_I. In all

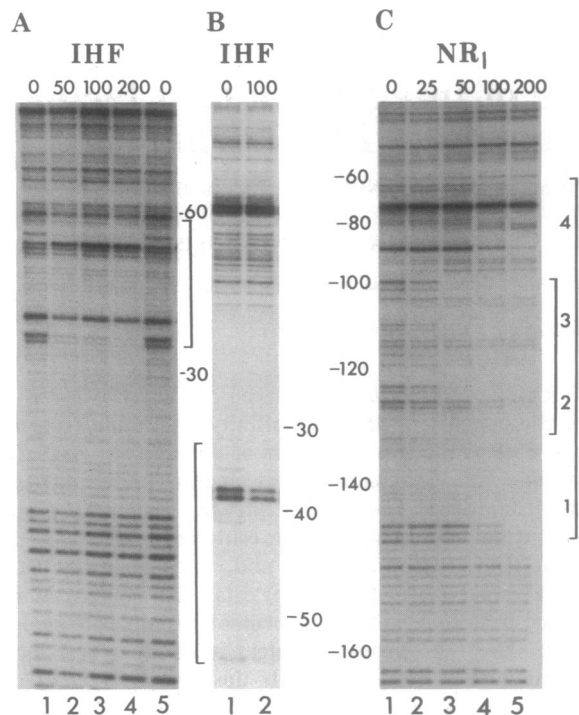


FIG. 1. DNase I footprinting of IHF and NR_I-binding sites in the *glnHP2* promoter regulatory region. Plasmid DNA was incubated with increasing concentrations of IHF (A and B) or NR_I (C), shown above each lane (nM). Primer extensions were carried out with primer FC6 (A) for the top strand and primer FC7 (B) or FC5 (C) for the bottom strand. The extent of protection by IHF or NR_I is indicated by brackets; in the case of NR_I, the small and the large brackets indicate protection at 50 nM and 200 nM, respectively. The sequence and the NR_I-binding sites are numbered as shown in Fig. 4.

the experiments in which NR_I was present, the kinase NR_{II} and ATP were added to allow phosphorylation of NR_I. As expected, sites 2 and 3 were protected at the low NR_I concentrations (25–50 nM), while sites 1 and 4 became occupied only at 200 nM NR_I (Fig. 1C). The interaction of NR_I with its binding sites was also studied on intact plasmid DNA by DMS footprinting. The pattern of methylation at both strands is shown in Fig. 2A and B. In the presence of 25–50 nM NR_I, protection of G residues in sites 2 and 3 only was observed; -102, -112 to -115, -125, and -126 in the bottom strand (Fig. 2A, lanes 4 and 5), and -105, -106, -118, and -119 in the top strand (Fig. 2B, lane 2). G residues at -128 in the bottom strand (Fig. 2A, lane 5) and at -103 in the top strand (Fig. 2B, lane 2) were hypermethylated. At a high concentration of NR_I (200 nM), sites 1 and 4 also became occupied; positions -72, -76, and -139 on the bottom strand (Fig. 2A, lane 6) and -87, -131, and -142 on the top strand (Fig. 2B, lane 3) were protected from DMS attack while positions -84, -86, -141, and especially -130 on the bottom strand, and -73, -75, and -144 on the top strand were hyperreactive. The hyperreactive bands may reflect both a bending of the DNA induced by the bound NR_I and a distortion of the DNA between two occupied half sites (28).

Closed and Open Complexes at the *glnHP2* Promoter. The *glnHP2* promoter has a T residue at position -14 instead of the G in the GC doublet found in all previously studied σ^{54} -dependent promoters (7, 29). We used DMS footprinting to determine whether or not closed complexes are formed at this unusual promoter. Fig. 2C shows the analysis of the bottom strand of *glnHP2*. The interaction of $E\sigma^{54}$ with the promoter was observed in the absence of NR_I phosphate (compare lanes 1 and 2); G residues at -28, -24, -22, and -13 were protected from methylation, indicating that $E\sigma^{54}$

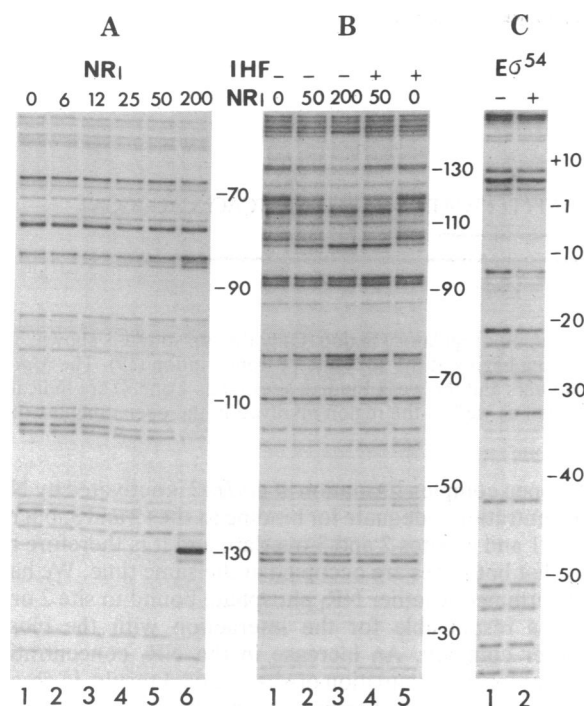


FIG. 2. DMS footprints of NR₁, IHF, and E σ^{54} interactions with the *glnHp2* promoter region. (A) Methylation pattern at NR₁-binding sites on the bottom strand. (B) Methylation pattern at NR₁ (lanes 1–4) and IHF (lanes 4 and 5) binding sites on the top strand. Numbers above each lane indicate concentration of NR₁ (nM). (C) Methylation pattern of the closed complex on the bottom strand. The presence (+) or absence (–) of IHF or E σ^{54} in each incubation mixture is indicated. Concentrations were as follows: E, 25 nM; σ^{54} , 100 nM; IHF, 50 nM; NR₁, as indicated. The sequence is numbered as shown in Fig. 4. Extensions were carried out with primers FC5 (A), FC6 (B), and FC7 (C).

forms a closed promoter complex with *glnHp2*. Protection of these residues was unchanged when NR₁ and NR₁ plus IHF were also present (data not shown). A number of residues in the region between –6 and +3 became reactive to DMS in the presence of E σ^{54} and NR₁, and they became even more reactive when IHF was present (data not shown). The reactivity of residues flanking the transcription start site in the presence of NR₁ phosphate may reflect formation of the open promoter complex.

Open complexes between E σ^{54} and the *glnAp2* promoter have been visualized with potassium permanganate (28), which oxidizes preferentially T and C residues in single-stranded DNA. We therefore used this reagent to probe *glnHp2* (Fig. 3A, bottom strand; Fig. 3B, top strand). In the presence of E σ^{54} and NR₁ phosphate, T residues at –12, –10, –8, –3, –2, and +8 and a C residue at –4 on the bottom strand (Fig. 3A, lanes 4–6), and T residues at –9, –7, –5, –1, and +2 on the top strand (Fig. 3B, lanes 4–7) became hyperreactive. These residues were not reactive when NR₁ was absent (Fig. 3, lanes 2). This region therefore represents the DNA melted in an open complex (Fig. 4). When NR₁ was present at a high concentration (200 nM), there was less open complex formation (compare lanes 4 and 6 in Fig. 3A and B). On the other hand, the same residues became even more reactive when IHF was present (compare lanes 4 and 5 in Fig. 3A and lanes 4 and 5 and lanes 6 and 7 in Fig. 3B) but did not react when only IHF or IHF and E σ^{54} were present, indicating that IHF stimulates the NR₁-dependent open complex formation. The results of the footprinting experiments are summarized in Fig. 4.

IHF Affects Activation of *glnHp2* Transcription by NR₁. All the DNA templates used in these experiments carried the

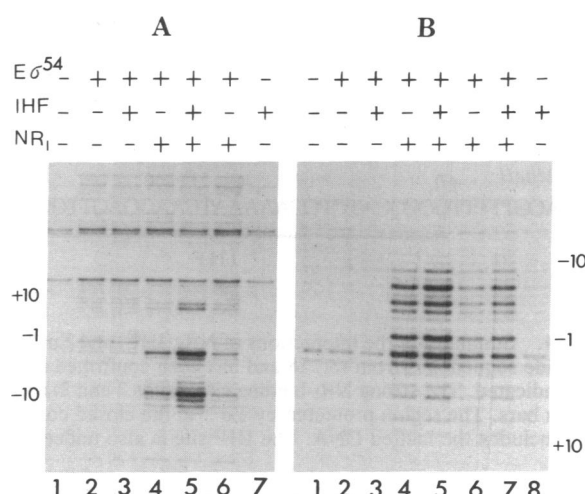


FIG. 3. Detection of the open complex at *glnHp2* with potassium permanganate. Supercoiled DNA was treated with potassium permanganate after incubation in the presence (+) or absence (–) of proteins as indicated above each lane. Concentrations of proteins when present were as follows: E, 25 nM; σ^{54} , 100 nM; IHF, 50 nM; NR₁, 50 nM (lanes 4 and 5) and 200 nM (lanes 6 and 7B). The sequence is numbered as shown in Fig. 4. Primer extensions were carried out with primer FC7 for the bottom strand (A) and primer FC6 for the top strand (B).

glnHp2 promoter and the IHF-binding site and included the following: pFC50, which contains the NR₁-binding sites in the wild-type position—that is, the center of sites 2 and 3 are at –122 and –109, respectively; pFC54, which has NR₁-binding sites removed; and pFC57, which contains the NR₁-binding sites further upstream in inverted orientation—that is, the center of sites 2 and 3 are at positions –382 and –395, respectively.

Using the wild-type promoter, we first showed that NR₁ phosphate was able to activate initiation of transcription at *glnHp2* (Fig. 5, pFC50, lanes 1–6). The concentration of NR₁ needed for activation was \approx 5 times higher than that needed for activation of *glnAp2* (23). When the NR₁ concentration was higher than 50 nM, the activation was decreased (Fig. 5, pFC50, lanes 5 and 6). In the presence of IHF, the activation by NR₁, at concentrations of 10–200 nM, was stimulated (Fig. 5, pFC50, lanes 7–12). IHF by itself did not activate transcription (Fig. 5, lane 7). The stimulation by IHF was greater at lower concentrations of NR₁ (\approx 5-fold at 10 nM NR₁) than at higher concentrations (2-fold at 50 nM NR₁) (Fig. 5, pFC50). On the other hand, the stimulation by IHF was the same over a wide range of E σ^{54} concentrations (data not shown). The stimulatory effect could be observed at an IHF concentration of 15 nM but was maximal at 50 nM (data not shown). IHF stimulation was not observed in the NR₁-mediated activation of *glnAp2* (data not shown), a promoter that lacks an IHF-binding site.

When the NR₁-binding sites were removed, NR₁ was able to activate transcription but only at the higher concentration (Fig. 5, pFC54). In this case, IHF inhibited the activation of *glnHp2* (Fig. 5, pFC54, lanes 7–12). When the NR₁-binding sites were moved further upstream, NR₁ was still able to activate transcription at a low concentration (Fig. 5, pFC57, lanes 1–6) as has been shown for *glnAp2* (3, 30); however, in this case too, IHF inhibited the activation (Fig. 5, pFC57, lanes 7–12).

DISCUSSION

The *glnHp2* promoter differs from all previously identified σ^{54} -dependent promoters by having the dinucleotide TC rather

nifH promoter by NIFA. These results may also explain the observation that in intact cells the NR_I binding sites of the IHF-independent *glnAp2* promoter can be moved more than 1 kilobase in either direction and remain functional, while in the case of *nifH*, where activation is stimulated by IHF, a corresponding change reduces transcription 10-fold (8). Our results suggest that a DNA-bending protein could stimulate or diminish activation of gene expression without interacting directly with the activator or with the RNA polymerase. This mechanism may play a role in the regulation of gene expression in eukaryotic cells where some transcription factors have been shown to bend DNA (35–37).

We are grateful to Howard Nash and Carol Robertson for the generous gift of purified IHF, to Tsutomu Nohno for kindly supplying plasmid pTN240, to Maria R. Diaz-Torres for reading the manuscript and helping in the preparation of the figures, and to Hilda Harris-Ransom for help in typing the manuscript. We thank Selina Sasse-Dwight and Jay D. Gralla for providing footprinting protocols prior to publication. This work was supported by a grant from the National Science Foundation (DMB-8817091).

1. Hunt, T. P. & Magasanik, B. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 8453–8457.
2. Ninfa, A. J. & Magasanik, B. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 5909–5913.
3. Reitzer, L. & Magasanik, B. (1986) *Cell* **45**, 785–792.
4. Ninfa, A. J., Brodsky, E. & Magasanik, B. (1989) in *DNA-Protein Interactions in Transcription*, UCLA Symposia on Molecular and Cellular Biology, ed. Gralla, J. (Liss, New York), pp. 43–52.
5. Wedel, A., Weiss, D. S., Popham, D., Droge, P. & Kustu, S. (1990) *Science* **248**, 486–490.
6. Su, W., Porter, S., Kustu, S. & Echols, H. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 5504–5508.
7. Gussin, G. N., Ronson, C. W. & Ausubel, F. M. (1986) *Annu. Rev. Genet.* **20**, 567–591.
8. Buck, M., Miller, S., Drummond, M. & Dixon, R. (1986) *Nature (London)* **320**, 374–378.
9. Morett, E. & Buck, M. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 9401–9405.
10. Drummond, M., Whitty, P. & Wootton, J. (1986) *EMBO J.* **5**, 441–447.
11. Morett, E. & Buck, M. (1989) *J. Mol. Biol.* **210**, 65–77.
12. Tuli, R. & Merrick, M. J. (1988) *J. Gen. Microbiol.* **134**, 425–432.
13. Santero, E., Hoover, T., Keener, J. & Kustu, S. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 7346–7350.
14. Cannon, W. V., Kreutzer, R., Kent, H. M., Morett, E. & Buck, M. (1990) *Nucleic Acids Res.* **18**, 1693–1701.
15. Hoover, T. R., Santero, E., Porter, S. & Kustu, S. (1990) *Cell* **63**, 11–22.
16. Friedman, D. I. (1988) *Cell* **55**, 545–554.
17. Willis, R. C., Iwata, K. K. & Furlong, C. E. (1975) *J. Bacteriol.* **122**, 1032–1037.
18. Nohno, T., Saito, T. & Hong, J.-S. (1986) *Mol. Gen. Genet.* **205**, 260–269.
19. Nohno, T. & Saito, T. (1987) *Nucleic Acids Res.* **15**, 2777.
20. Ninfa, A. J., Ueno-Nishio, S., Hunt, T. P., Robustell, B. & Magasanik, B. (1986) *J. Bacteriol.* **168**, 1002–1004.
21. Reitzer, L. J. & Magasanik, B. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 5554–5558.
22. Elliot, S. & Geiduschek, E. P. (1984) *Cell* **36**, 211–219.
23. Ray, L., Claverie-Martin, F., Weglenski, P. & Magasanik, B. (1990) *J. Bacteriol.* **172**, 818–823.
24. Gralla, J. D. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 3078–3081.
25. Sasse-Dwight, S. & Gralla, J. D. (1989) *J. Biol. Chem.* **264**, 8074–8081.
26. Borowiec, J. A., Zhang, L., Sasse-Dwight, S. & Gralla, J. D. (1987) *J. Mol. Biol.* **196**, 101–111.
27. Yang, C.-C. & Nash, H. A. (1989) *Cell* **57**, 869–880.
28. Sasse-Dwight, S. & Gralla, J. D. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8934–8938.
29. Kustu, S., Santero, E., Keener, J., Popham, D. & Weiss, D. (1989) *Microbiol. Rev.* **53**, 367–376.
30. Ninfa, A. J., Reitzer, L. J. & Magasanik, B. (1987) *Cell* **50**, 1039–1046.
31. Robertson, C. A. & Nash, H. A. (1988) *J. Biol. Chem.* **263**, 3554–3557.
32. Thompson, J. F. & Landy, A. (1988) *Nucleic Acids Res.* **16**, 9687–9705.
33. Popham, D. L., Szeto, D., Keener, J. & Kustu, S. (1989) *Science* **243**, 629–635.
34. Drummond, M. H., Contreras, A. & Mitchenall, L. A. (1990) *Mol. Microbiol.* **4**, 29–37.
35. Shuey, D. J. & Parker, C. S. (1986) *Nature (London)* **323**, 459–461.
36. Viganis, M.-L. & Sentenac, A. (1989) *J. Biol. Chem.* **264**, 8463–8466.
37. Moskaluk, C. & Bastia, D. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 1826–1830.