JOURNAL OF BACTERIOLOGY, May 1988, p. 2240–2246
0021-9193/88/052240-07$02.00/0

Vol. 170, No. 5

Bidirectional Promoter in the *hut(P)* Region of the Histidine Utilization (*hut*) Operons from *Klebsiella aerogenes*

ANTHONY J. NIEUWKOOP,† SANDRA A. BALDAUF, MICHAEL E. S. HUDSPETH,‡ AND ROBERT A. BENDER*

Department of Biology, University of Michigan, Ann Arbor, Michigan 48109

Received 5 October 1987/Accepted 18 February 1988

The *hut(P)* region (i.e., the region responsible for regulation of *hutUH* expression) of the *Klebsiella aerogenes* histidine utilization (*hut*) operons contains a bidirectional promoter. One transcript from this promoter encodes the *hutUH* operon; the role of the oppositely directed transcript is unknown, although it appears to be involved in regulating *hutUH* expression (A. J. Nieuwkoop, S. A. Boylan, and R. A. Bender, J. Bacteriol. 159:934–939, 1984). A 247-base-pair (bp) fragment containing *hut(P)* carries two RNA-polymerase-binding sites, protecting 48 and 57 bp of DNA from digestion by exonuclease III. The locations of these two binding sites agree with the start sites of the two transcripts produced from *hut(P)* DNA in vitro and in vivo. The binding sites share a 4-bp region, suggesting that occupancy of the regulatory site precludes occupancy of the *hutUH* promoter, and vice versa. In the absence of positive effectors, the binding to the site responsible for *hutUH* transcription is weaker than the binding to the site responsible for regulation. The nucleotide sequence of the 250-bp fragment containing *hut(P)* contains two possible matches to the consensus sequence for *Escherichia coli* promoters, a better and worse match, corresponding in position to the stronger and weaker RNA-polymerase-binding sites, respectively. The sequence also contains a region similar to the consensus sequence for binding of the catabolite gene activator protein of *E. coli*. A sequence similar to the consensus for Ntr-dependent promoters was also found, overlapping both RNA-polymerase-binding sites, but it is not a functional promoter. Finally, an initiation codon preceded by a Shine-Dalgarno consensus sequence and followed by an open reading frame identifies a probable start of the *hutU* gene coding sequences.

Klebsiella aerogenes can degrade the amino acid histidine enzymatically to ammonia, glutamate, and formamide (17). The histidine utilization (*hut*) genes are arranged in two tandem operons in the order *hut(M)IGC(P)UH*, where *I*, *G*, *U*, and *H* encode the four enzymes of the pathway and *hutC* encodes a repressor which regulates both operons (5, 11). The *hut(M)* and *hut(P)* regions represent the sequences controlling expression of the left and right operons, respectively. Therefore, the *hut(P)* region would be expected to include the binding site for RNA polymerase ($hutU_p$), the binding site for repressor ($hutU_o$), and the binding site(s) for positive effectors.

Histidine can serve as the sole carbon or sole nitrogen source for *K. aerogenes*, and high-level expression of the *hut* operons requires not only the inactivation of the repressor but also the presence of a positive activator. This activator can be either the catabolite gene activator protein (CAP) in the presence of cyclic AMP (cAMP), signaling carbon or energy limitation, or an unknown factor(s) signaling nitrogen limitation (4, 23).

We have previously suggested that the positive regulation of *hutUH* expression by CAP-cAMP might actually be a double-negative control (22), in which, under nonactivating conditions, RNA polymerase binds to a site (P_C) near the *hutUH* promoter (P_{UH}). RNA polymerase bound at P_C would block the binding of RNA polymerase to P_{UH}. CAP-cAMP would prevent RNA polymerase from binding to P_C, which in turn would allow access of RNA polymerase to P_{UH}.

In this report we present the complete nucleotide sequence of the wild-type *hut(P)* region and identify the two RNA-polymerase-binding sites, P_C and P_{UH}, within this sequence. In addition we identify consensus sequences suggesting binding sites for carbon and nitrogen effectors of *hutUH* expression (2, 8). Thus, the data presented here support our previous double-negative control model of a bidirectional, mutually antagonistic promoter within *hut(P)*.

MATERIALS AND METHODS

Growth conditions. Plasmid-containing strains were grown in L broth or on L agar (3) with ampicillin at 100 μg/ml (LBamp). Minimal medium consisted of W salts (3) supplemented with 0.4% histidine as the sole carbon and nitrogen source. For S1 nuclease mapping, minimal medium was 2W5 salts (double-strength W salts adjusted to pH 7.5) supplemented with 0.4% glucose and 0.2% histidine for nitrogen-limited growth and 0.4% histidine alone for carbon-limited growth.

Chemicals and enzymes. Restriction enzymes were purchased from Amersham Corp., Bethesda Research Laboratories, Inc., Boehringer Mannheim Biochemicals, and New England BioLabs, Inc., and used in accordance with the directions of the suppliers. Calf intestinal alkaline phosphatase was purchased from Boehringer Mannheim. Polynucleotide kinase was purchased from Amersham or Boehringer Mannheim. The large fragment of *Escherichia coli* DNA polymerase I (Klenow fragment) was obtained from Amersham, New England Nuclear Corp., or New England BioLabs. *E. coli* RNA polymerase holoenzyme and exonuclease III were obtained from New England BioLabs. Nonradioactive deoxynucleotides were purchased from P-L Biochemi-

* Corresponding author.
† Present address: Department of Biology, Hope College, Holland, MI 49423.
‡ Present address: Department of Biological Sciences, Northern Illinois University, DeKalb, IL 60015.

cals, Inc.; radioactive nucleotides ([γ-^{32}P]ATP and [α-^{32}P] dATP) were from Amersham. Acrylamide and bisacrylamide were obtained from Bio-Rad Laboratories; ultrapure urea was obtained from Schwarz/Mann or International Biotechnologies, Inc. Chemicals used for Maxam-Gilbert sequencing reactions were of reagent grade or were purchased in kit form (NEH-010) from New England Nuclear. Agarose and CsCl were from Miles Laboratories, Inc., and Kawecki Berylco Industries, respectively.

Isolation and manipulation of DNA. Large-scale preparations of plasmid DNA were purified as described previously (6). In addition, plasmid DNA to be used for end labeling was passed over a Biogel A 1.5-m agarose (Bio-Rad) column to remove low-molecular-weight nucleic acids and nucleotides (18). Small-scale preparations as cleared lysates were prepared as boiled Triton extracts (14) for plasmid transformations and DNA analysis. Transformation of *E. coli* strains with plasmid DNA was carried out by the $MgCl_2$-$CaCl_2$ method (16). Competent cells were stored in 15% glycerol at −70°C (20). Unless noted, transformants were purified by single-colony isolation on selective medium before further characterization.

Construction of plasmid pCB220. For reasons described below, we wanted to be sure that we were not sequencing a cloning artifact. Therefore, bacteriophage P1*clr100*KM (9) was used to transduce the *hut(P)* region from the chromosome in strain KG2000 onto plasmid pCB103, a deletion derivative of pCB101 lacking the central region of *hut*. Thus, pCB103 lacks part of *hutG*, all of *hutC*, *hut(P)*, and *hutU*, and part of *hutH* (6). Since pCB103 is a deletion plasmid, transductants which acquire the ability to catabolize histidine (Hut$^+$ phenotype) must have replaced at least the entire region corresponding to the deletion with material from the transducing DNA, in this instance with DNA from strain KG2000 (11). Transduction was performed as described by Goldberg et al. (9). Although the transductions were between different genera, *K. aerogenes* KG2000 and *E. coli* EB752 (RH202 [1] containing plasmid pCB103), the ratio of phage to cells was kept at 1 PFU/cell (28). This was done to increase the likelihood of transducing the homologous (*K. aerogenes*) plasmid DNA rather than the heterologous (*E. coli*) chromosomal DNA to Hut$^+$ (growth with histidine as the sole carbon source). Hut$^+$ transductants appeared at a frequency of ca. 10^{-9} after 10 to 12 days of incubation. Four transductants were purified by single-colony isolation on nonselective (LBamp) plates, and DNA was isolated from these purified transductants as described above.

This DNA was used to transform *E. coli* RH202 to ampicillin resistance. Twelve of twelve transformants tested were able to use histidine as the sole carbon source. Plasmid DNA was isolated from four of these purified transformants and digested with the restriction endonuclease *Sal*I. The newly constructed plasmids, digested with *Sal*I, yielded four bands; pCB101 (intact *hut*) digested with *Sal*I yielded three bands; the deletion plasmid pCB103 digested with *Sal*I gave two bands. The newly constructed plasmids appeared to contain the fragments of both pCB101 and pCB103. Since the plasmids examined had been transformed at low DNA concentrations, it appeared unlikely the host strain had been independently transformed with both a pCB101-like plasmid and pCB103. However, by necessity the transductions had been performed in a recombination-proficient strain. Therefore, it was a distinct possibility that the plasmid isolated was in fact dimeric. The restriction pattern of the plasmid resulting from transduction could be explained if these plasmids were composed of a copy of pCB103 joined head to tail with a copy of pCB101-like DNA. To test this possibility, we attempted to resolve the dimer into monomers by digestion with *Eco*RI. This enzyme cuts both pCB103 and pCB101 only once and not within the cloned *hut* sequences. The *Eco*RI-digested mixture was diluted, treated with DNA ligase, and used to transform strain RH202 to ampicillin resistance. Of 95 transformants tested, 59 exhibited the Hut$^+$ phenotype and 36 were Hut$^-$. Plasmid DNA was isolated from six Hut$^+$ colonies and two Hut$^-$ colonies and used to transform strain RH202 to ampicillin resistance. In every instance, all the transformants tested (12 of 12) exhibited the same Hut phenotype as the strain from which the transforming DNA had been isolated.

Plasmid DNA isolated from six Hut$^+$ transformants and two Hut$^-$ transformants was digested with *Sal*I. All of the plasmids from the Hut$^+$ strains yielded digestion patterns identical to that of pCB101; all of the plasmids from the Hut$^-$ strains yielded digestion patterns identical to that of the original deletion plasmid, pCB103. Thus, the dimeric plasmid constructed by P1 transduction had been successfully resolved into the original deletion plasmid, pCB103, and a new plasmid, pCB220, analogous to pCB101 but containing the *hut(P)* region from KG2000.

Sequencing procedures. Plasmid pCB101, pCB111 (6) or pCB220 (see above) was digested with *Sal*I or *Pvu*II and subsequently 5′ end labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase or 3′ end labeled at the *Sal*I site by filling in the ends with α-^{32}P-labeled nucleotides. These single-end-labeled restriction fragments were separated on 1% agarose or 4% polyacrylamide gels, and the DNA fragment of interest was recovered by elution onto Whatman filter paper (18). The eluted DNA fragments were precipitated with ethanol, rinsed with 70% ethanol, and suspended in deionized water. The nucleotide sequences of the fragments were determined by the chemical modification methods of Maxam and Gilbert, essentially as described (19). The DNA cleavage products were separated on 6 and 11.8% (wt/vol) polyacrylamide gels containing 7 M urea and TBE (0.2 M Tris-borate [pH 8.3], 4 mM EDTA) and autoradiographed by exposure to X-ray film (Kodak XAR-5) at −70°C.

DNA protection experiments. The DNA protection experiment procedure was essentially that of Shalloway et al. (25), modified by M. Peterson and W. Reznikoff (personal communication). All reactions were carried out in 1.5-ml Eppendorf tubes in a 50-μl reaction volume. About 0.1 pmol of 5′-end-labeled fragment was combined with 20 mM Tris (pH 7.4), 10 mM 2-mercaptoethanol, 0.1 mM EDTA, 6 mM $MgCl_2$, 100 mM KCl, and 15% glycerol. This mixture was equilibrated at 37°C for ca. 5 min. RNA polymerase (25 pmol) was then added and allowed to bind for 30 min at 37°C. Heparin (1 μl, 1 mg/ml) was added to the reaction mixture, and the DNA was digested for 15 min at 37°C with 100 U of exonuclease III. The reaction was terminated by the addition of 25 μl of 50 mM EDTA. Protein was extracted from this mixture with 100 μl of phenol-Sevag (phenol-chloroform-isoamyl alcohol, 25:24:1), followed by extraction of the phenol with 100 μl of chloroform. Samples were made 0.3 M in ammonium acetate and precipitated with 2 volumes of ethanol at −70°C for 15 min. The precipitate was collected by centrifugation, dried under vacuum for 30 min, and suspended in 2 μl of 0.1% xylene cyanol–0.1% bromphenol blue–1 mM EDTA–90% deionized formamide. The digestion products were separated in polyacrylamide gels, and an autoradiograph was made of the gels by exposure to X-ray film (Kodak XAR-5) at −70°C.

S1 nuclease mapping. RNA was extracted from *K. aero-*

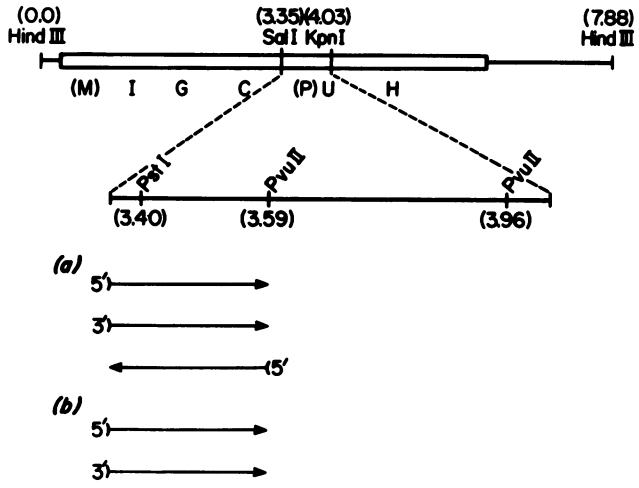

FIG. 1. Partial restriction map of the *hut(P)* control region and fragments used for DNA nucleotide analysis. Locations (map units) are from the restriction enzyme cleavage map of Boylan et al. (6). The tail of the arrow represents the ^{32}P-labeled end of the fragment, and the arrowhead represents the direction in and extent to which each fragment was sequenced. The DNA ends labeled with ^{32}P are indicated by 5′ and 3′. (a) Fragments sequenced from plasmid pCB101; (b) fragments sequenced from plasmid pCB220.

genes by the method of Salser et al. (24) except that the hot-phenol extraction was performed three times with phenol saturated with 0.1 M sodium acetate at pH 5.2. The RNA solution was extracted twice more with phenol-Sevag and twice with Sevag. S1 mapping was done as described by Maniatis et al. (18). Denaturation was at 85°C for 15 min, hybridization was at 40°C for 3 h, and digestion was at 37°C for 45 min with 100 U of S1 nuclease (Sigma Chemical Co.). The resulting species were separated by electrophoresis in 8% polyacrylamide gels, using *Hin*fI-digested pBR322 as size standards.

RESULTS

Nucleotide sequence of the *hut(P)* region. The *hut(P)* region from *K. aerogenes* contains a bidirectional promoter (22), which is wholly contained within the restriction fragment extending from the *Pst*I site at 3.4 map units on the standard *hut* restriction map (5, 22) to the *Pvu*II site at 3.59 map units (Fig. 1). The nucleotide sequence of the DNA from this region was determined (by the chemical cleavage method of Maxam and Gilbert [19]) from three independent analyses (Fig. 1). The nucleotide sequence of one strand was determined by using DNA which had been 5′ end labeled at the *Sal*I site, with the sequence extending rightward (Fig. 1a). The sequence of the other strand was obtained both from DNA which was 3′ end labeled at the *Sal*I site, with the sequence extending rightward, and from DNA which was 5′ end labeled at the *Pvu*II site, with the sequence extending leftward. By comparison of the three independent sequences thus obtained, an unambiguous DNA sequence of the *hut(P)* region was deduced (Fig. 2).

Confirmation of the DNA sequence of *hut(P)*. The DNA sequence shown in Fig. 2 was determined by using plasmid pCB101 as the source of DNA (6). The *hut* sequences on this plasmid had been maintained outside *K. aerogenes* for a considerable length of time, first as a λ d*hut* transducing phage (28) and then as a plasmid (6). We have also shown that the regulation of histidase synthesis in *E. coli* strains carrying pCB101 is different from that in *E. coli* strains carrying the *K. aerogenes hut* genes on the chromosomes (5, 6; D. Hart and R. A. Bender, manuscript in preparation). To eliminate the possibility that the *hut(P)* regulatory region had been altered during the isolation or maintenance of pCB101, a new plasmid, pCB220, was constructed with *hut(P)* DNA derived directly from the chromosome of a *hut*$^+$ *K. aerogenes* strain, KG2000 (11), as described in Materials and Methods.

The nucleotide sequence of the *hut(P)* region of pCB220 was determined from two independent sequence analyses. As before, the nucleotide sequence of one strand was determined by using DNA which was 5′ end labeled at the *Sal*I site, with the sequence extending rightward. The sequence of the other strand was determined by using DNA which was 3′ end labeled at the *Sal*I site and extending rightward (Fig. 1b). The nucleotide sequence obtained from these analyses was identical to that determined for pCB101 (Fig. 2). Thus, the derived nucleotide sequence shown in

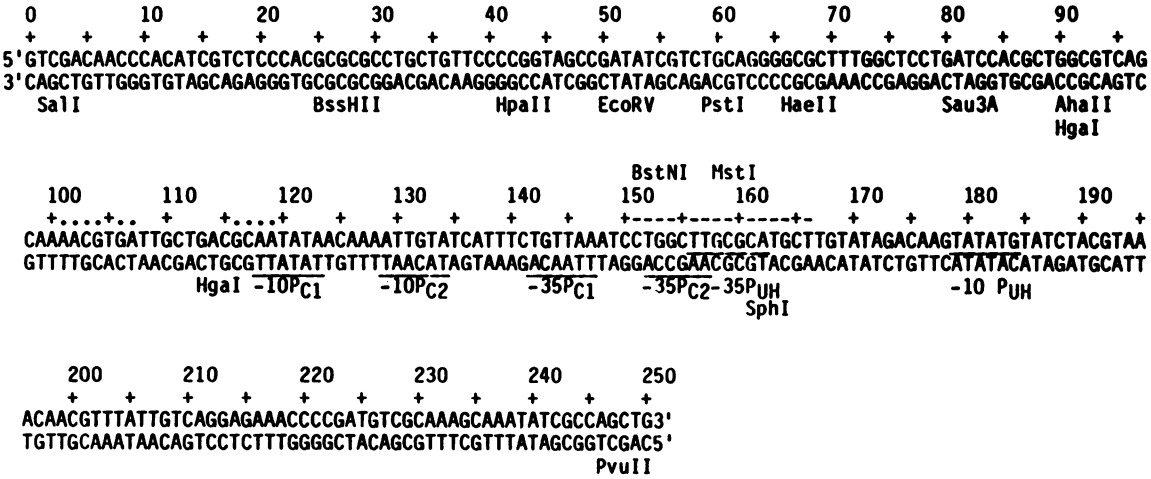

FIG. 2. Complete nucleotide sequence of *hut(P)*. Homologies to known consensus sequences (see Fig. 6 for details) are indicated as follows: ——, RNA polymerase binding; . . . . . ., CAP-cAMP binding; ----------, nitrogen control sequence. Nucleotide 1 represents the end labeled at the *Sal*I site.

Fig. 2 does in fact represent the wild-type *hut(P)* region and no alterations had occurred during isolation or maintenance of the cloned *hut* genes.

Accuracy of the DNA sequence determined for *hut(P)*. The *hut(P)* region is presumed to be the target site for the action of a large number of regulatory proteins. The region is expected to contain binding sites for at least RNA polymerase, a repressor (*hutC* product), CAP, nitrogen activation factor(s), and translation initiation, probably for the oxygen regulatory factor(s) described by Goldberg and Hanau (10), and perhaps for terminating transcription of the adjacent operon. Because of this density of regulatory sites, the accuracy of this sequence is particularly important. Therefore, the sequence was confirmed by the dideoxy method (A. Schwacha, R. Osuna, and R. A. Bender, unpublished data). The existence of restriction sites offers an independent confirmation of at least part of the sequence. An analysis of the nucleotide sequence shown in Fig. 2 predicts cleavage sites for the following enzymes: *Sal*I, *Bss*HI, *Hpa*I, *Eco*RV, *Pst*I, *Hae*II, *Sau*3A, *Aha*II, *Hga*I (present twice), *Bst*NI, *Mst*I, *Sph*I, and *Pvu*II. The existence and precise location of each of these sites have been confirmed by digestion of the *Sal*I-*Pvu*II DNA fragment, confirming the identity of 70 of the 247 nucleotides in this fragment. The absence of restriction sites also provides some slight confirmation of the sequence. Direct analysis has determined that this region is not cleaved by *Bam*HI, *Bgl*II, *Eco*RI, *Hae*III, *Hin*dIII, *Hin*fI, *Sma*I, *Sst*I, *Xba*I, or *Xho*I (6), consistent with the absence of these sites in the nucleotide sequence (Fig. 2).

RNA-polymerase-binding sites within *hut(P)*. Since the *hut(P)* region contains a bidirectional promoter (22), two RNA-polymerase-binding sites should lie in this region. To confirm the location of these two sites, DNA protection experiments were performed as described in Materials and Methods. The 247-base-pair (bp) *hut(P)* DNA fragment was isolated and 5′ end labeled at the *Sal*I site. This fragment was incubated with RNA polymerase and then digested with exonuclease III. Exonuclease III digests DNA in a 3′-to-5′ direction until encountering a bound RNA polymerase molecule. The size of the resulting fragment represents the distance from the *Sal*I site to the 3′ boundary of the bound RNA polymerase molecule. The digestion products were resolved on polyacrylamide gels and detected by autoradiography. Two major bands were identified (Fig. 3, lane 2). The more intense band corresponded (after correcting for the mobility differences between Maxam-Gilbert cleavage products and exonuclease III cleavage products) to a length of 154 bases. The other band (representing 207 bases) is actually much less intense than it appears in Fig. 3 since the entire lower band of the doublet and about half of the upper band resulted from digestion artifacts and were seen even in the absence of added RNA polymerase. These lengths correspond to the distance from the labeled *Sal*I site to the rightmost boundaries of the RNA-polymerase-binding sites

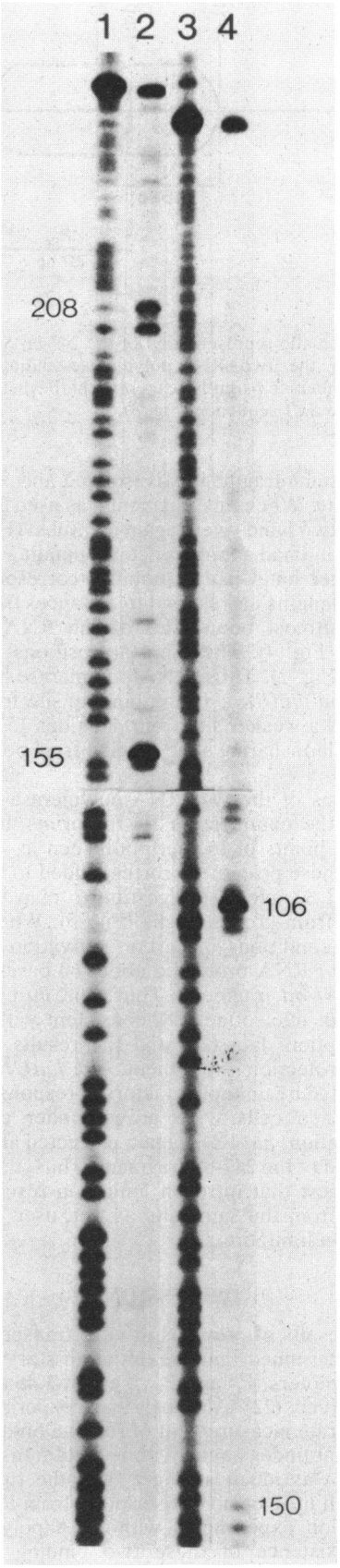

FIG. 3. Protection of *hut(P)* DNA by RNA polymerase against exonuclease III digestion. Lanes 2 and 4, Approximately 0.1 pmol of DNA fragment was incubated with 25 pmol of *E. coli* RNA polymerase holoenzyme before being digested with exonuclease III (100 U/ml) as described in the text; lane 1, Maxam and Gilbert A+G reaction of 247-bp *Sal*I-*Pvu*II fragment 5′ end labeled at the *Sal*I site; lane 2, 247-bp fragment 5′ end labeled at the *Sal*I site; lane 3, Maxam and Gilbert A+G reaction of 247-bp *Sal*I-*Pvu*II fragment 5′ end labeled at the *Pvu*I site; lane 4, 247-bp fragment 5′ end labeled at the *Pvu*I site. The numbers in the margins are lengths in bases.

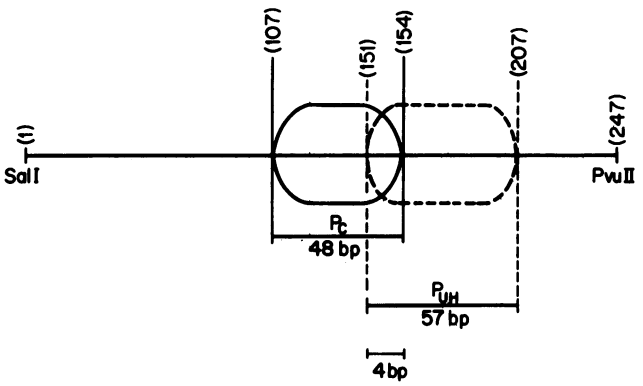

FIG. 4. Schematic representation of the 247-bp *Sal*I-*Pvu*II fragment depicting the two RNA-polymerase-binding sites within *hut(P)*. The numbers in parentheses represent the distances from the beginning of the *Sal*I sequence (Fig. 3).

(Fig. 4). The same fragment was isolated and 5′ end labeled at the *Pvu*II site. When this fragment was used for protection experiments, two bands were again identified (Fig. 3, lane 4). The intense band had a mobility corresponding to 141 bases, and the weaker band had a mobility corresponding to 97 bases. These lengths correspond to distances from the *Pvu*II site to the leftmost boundaries of the RNA-polymerase-binding sites (Fig. 4). These are at positions 107 and 151, respectively (Fig. 2). Thus, RNA polymerase has two binding sites within *hut(P)*: a strong binding site from positions 107 to 154 and a weaker site from positions 151 to 207. The two sites overlap, sharing a 4-bp region from positions 151 to 154 (Fig. 4).

The positions of the two RNA-polymerase-binding sites identified by the exonuclease III footprints are consistent with the start points of transcription seen in vitro (22). To test whether these promoters corresponded to those used in vivo, we used S1 nuclease digestion to map the 5′ end of *hutUH* RNA from growing cells (Fig. 5). When cells were grown under conditions of carbon activation (CAP-cAMP activation), *hut* RNA protected about 60 bases of the right end of the 247-bp fragment. Thus, the *hutUH* transcript begins at about nucleotide 188, consistent with both the in vitro transcription data (22) and the results of the RNA polymerase protection experiment. The *hutUH* operon can also be activated by unknown factors in response to nitrogen limitation. When cells were grown under conditions of nitrogen limitation, *hut* RNA again protected about 60 bases of the right end of the 247-bp fragment. Thus, the data shown in Fig. 5 suggest that nitrogen limitation results in *hutUH* transcription from the same site as that used under conditions of carbon limitation.

DISCUSSION

From the results of previous in vitro transcription experiments we determined that the initiation start sites for two divergent promoters, P_C and P_{UH}, are at 3.46 and 3.53 map units, respectively (22). Although these experiments did not allow an accurate measurement of relative binding affinities, it appeared that under nonactivating conditions the leftward P_C promoter was much stronger than the rightward P_{UH} promoter, with little or no transcription from P_{UH} detected. DNA protection experiments with RNA polymerase confirmed the existence of these two binding sites. These experiments also confirmed the P_C binding site as being stronger than the P_{UH} site. The bands which indicate the P_C boundaries were more intense than those for P_{UH} (Fig. 3). In fact, the bands identified with P_{UH} were barely detectable. This would be expected if much more end-labeled DNA had RNA polymerase bound to P_C. The less intense bands denoting the P_{UH} boundaries would be the result of fewer end-labeled DNA fragments binding RNA polymerase at P_{UH} than at P_C. Therefore, upon encountering *hut(P)* DNA in the absence of positive effectors, RNA polymerase would preferentially bind at the P_C site.

The bands for the fragments used to determine the boundaries for P_{UH} were faint. They have been confirmed as representing P_{UH} by using DNA from *hut* mutants which, in vitro, express only the P_{UH} transcript (22). Protection experiments using mutant *hut(P)* DNA protected by RNA polymerase yielded only the bands at the positions seen for P_{UH} with wild-type DNA (A. J. Nieuwkoop and R. A. Bender, manuscript in preparation).

A comparison of several sequences within *hut(P)* with sequences involved in regulating gene expression is shown in Fig. 6. RNA-polymerase-binding sites from *E. coli* have two areas of sequence homology, located 35 and 10 bp upstream of the site of transcription initiation (12, 27). The nucleotide sequences of P_C and P_{UH}, shown to be protected by RNA polymerase, were scanned for −10 and −35 consensus sequences. For the strong (P_C) binding site, two potential matches to the consensus sequence were found, one (P_{C1}) in the region from nucleotide 147 reading leftward on the bottom strand to nucleotide 118 and the other (P_{C2}) from nucleotides 157 to 129. According to the rules of Mulligan et al. (21), P_{C1} would be expected to be stronger than both P_{C2}

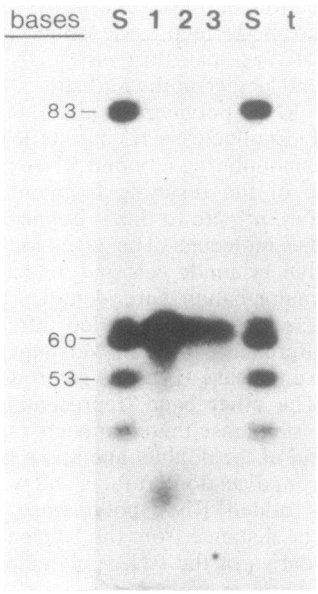

FIG. 5. S1 nuclease mapping of *hut*-specific transcripts. The 247-bp *hut(P)* fragment was 5′ end labeled at the *Pvu*II site (lane 1) or the *Pvu*II and *Sal*I sites (lanes 2, 3, and t), and S1 mapping was performed as described in the text. Lanes: 1, 300 μg of RNA from nitrogen-limited cells (grown in glucose-histidine minimal medium [3]) and 400 ng of DNA; 2, 300 μg of RNA from nitrogen-limited cells and 200 ng of DNA; 3, 300 μg of RNA from carbon-limited cells (grown in histidine minimal medium) and 200 μg of DNA; t, yeast tRNA, instead of *K. aerogenes* RNA, and 200 μg of DNA; S, size standards (*Hin*fI digest of pBR322). Only the lower one-fourth of the autoradiogram is shown. No bands were observed when 300 μg of RNA extracted from *hut*-repressed cells (grown in glucose-ammonia minimal medium) was used (data not shown).

Promoter:		
Consensus	TTGACA--N_{17}--TATAAT	
hutP_{UH}	TTGCGC--N_{17}--TATATG	(#156→184)
hutP_{C1}	TTAACA--N_{18}--TATATT	(#147→118)
hutP_{C2}	AAGCCA--N_{17}--TACAAT	(#157→129)
CAP-cAMP Binding		
Consensus	AANTGTGA--N_7--CACANT	
	AANCGTGA--N_7--CGCAAT	(#100→120)
Ntr Regulation:		
Consensus	CTGGYAYR--N_4--TTGCA	
	CTGGCTTG--N_3--ATGCT	(#151→166)

FIG. 6. Comparison of consensus sequences with sequences found within *hut(P)*. The numbers at the right correspond to the positions of the sequences as shown in Fig. 2.

and P_{UH}, whereas P_{C2} would be about equal in strength to P_{UH}. Since the results of both the in vitro transcription experiments (22) and the footprint analysis (Fig. 3) suggest that P_C is stronger than P_{UH}, P_{C1} is the more likely candidate to be the main promoter for leftward transcription. If so, then the RNA polymerase footprint represents protection from −44 to +7 relative to the predicted start of transcription. If P_{C2} is the main promoter for leftward transcription, the RNA polymerase protection would have included the region from −34 to +18. Although failure to protect the −35 hexamer is not unprecedented, it is not common. In vitro transcription gave three leftward transcripts from P_C to the *Pst*I site whose sizes (determined by comparison with the mobilities of DNA restriction fragments) were approximately 54, 59, and 63 bases; P_{C1} and P_{C2} predict transcripts of about 50 and 60 bases, in good agreement. The significance of the third transcript is unclear. Since only P_{C1} would be strong enough to outcompete P_{UH} and since P_{C1} as the main leftward promoter is consistent with RNA polymerase protection of the −35 hexamer, we tentatively suggest that P_{C1} is the main leftward promoter. S1 nuclease mapping studies of in vivo transcripts and mutational analysis will eventually resolve the question.

The rightward P_{UH} binding site showed protection by RNA polymerase from bases 152 to 207 (Fig. 4). The nucleotides from positions 179 to 184 (TATATG) match four of the six consensus nucleotides for a standard −10 region (TATAAT). This alignment violates the rule of a final T in the hexamer, a violation seen in only 4 of 112 *E. coli* promoters (12). However, if we demand a final T, then the next best −10 region hexamer, at positions 178 to 183 (GTATAT), matching only two of the consensus nucleotides, would be quite unlikely based on the predictions of Mulligan et al. (21). A −35 consensus region, TTGCGC, could also be identified for P_{UH}, lying between bases 156 and 161. Like the P_C −35 region, this sequence matches the consensus sequence in only three of six bases; however, the more important triplet TTG is present. These sequences would place the site of P_{UH} transcription initiation at about base 191. This start site is also consistent with previous in vitro transcription data (22), as well as with the preliminary S1 mapping data for in vivo transcription (Fig. 5). By using *Pvu*II-digested template DNA, runoff transcripts of ca. 60 bases were produced from P_{UH}. A transcript initiating at base 191 and continuing through the *Pvu*II site would produce a transcript of 57 bases. In addition, the location of the P_{UH} consensus sequence was confirmed by analysis of *hut(P)* mutants which increase *hutUH* expression in the absence of positive activators and contain alterations in this region, which is identified here as the P_{UH} −10 consensus (Nieuwkoop and Bender, manuscript in preparation).

From the protection experiments and identification of consensuslike sequences, it is quite obvious that the two RNA-polymerase-binding sites, P_C and P_{UH}, overlap. An overlap of 4 bp was determined, from bases 151 to 154 (Fig. 4). According to the model of RNA polymerase binding described by Sienbenlist et al. (27), it should be impossible for RNA polymerase to occupy these two overlapping sites at the same time. This supports our model of double-negative control, whereby RNA polymerase binding to P_C would inhibit RNA polymerase binding to P_{UH} (22).

Our model also proposed that the positive effects of CAP-cAMP on *hutUH* expression were in large part due to the ability of CAP to prevent RNA polymerase from binding to P_C. The sequence between bases 100 and 120 matches 10 of 12 bases identified as a CAP-binding consensus sequence by deCrombrugghe et al. (8). This site overlaps with the RNA-polymerase-binding site of P_C, in fact covering the putative start site for P_C transcription. A CAP molecule bound here would prevent the binding of RNA polymerase to P_C.

Since the *hutUH* operon is also activated under conditions of limiting nitrogen, it seemed probable that a site recognized by the nitrogen factor(s) would also be present within *hut(P)*. A site closely resembling the consensus sequence for nitrogen-regulated genes was identified (2). This sequence, located between bases 151 and 166, has 10 of 13 bases matching the consensus, except that the spacing of the two blocks of the consensus is 1 bp shorter in *hut(P)* than in the consensus. We can be quite sure of the spacing in *hut(P)* since the combination of *Mst*I and *Sph*I restriction sites spans the spacer and includes part of both blocks (totaling 9 of the 16 bp in the region). The *Bst*NI site confirms another 4 bp of the consensus, thus guaranteeing 13 of 16 bp, as well as the spacing.

This "Ntr consensus sequence" has recently been suggested to be an Ntr-specific promoter recognized by an Ntr-specific RNA polymerase whose sigma subunit has been replaced by the product of the *glnF* (*ntrA*) gene (2, 13, 15). However, it is unlikely that the sequence from bases 151 to 166 plays this role, for two reasons: (i) the results of our preliminary S1 mapping experiments suggest that Ntr-activated *hutUH* transcription begins about 10 bp downstream from the −10 region of the classical (Pribnow) promoter, not 12 bp downstream from the −12 region of the Ntr-specific promoter; and (ii) Buck (7) has recently shown that reduction of the spacing by 1 bp converts an active Ntr promoter into an inactive site.

A possible translation initiation site was identified 33 bases from the putative start site of the P_{UH} message (base 224). Between bases 211 and 215 a Shine-Dalgarno ribosome-binding sequence, AGGAG (26), followed by an ATG start codon and an open reading frame, was identified. This is probably the N terminus of the *hutU* (urocanase) gene product. However, the amino acid sequence for the protein has not been determined, so the amino acid sequence predicted from this nucleotide sequence cannot be confirmed. No Shine-Dalgarno sequence could be identified for the leftward P_C transcript. In fact, even when an additional 200 bp nucleotide sequence to the left of the *Sal*I site (3.35 to 3.15 map units) was examined, no translational start codon followed by an open reading frame was identified (unpub-

lished observation). Thus, it appears that the P_C transcript, although functional in vivo (22), may not encode a protein.

Finally, two features of the sequence whose significance is unclear deserve comment. The first is the extreme asymmetry in G+C content in this region. The region to the left of the proposed CAP-binding site (bases 1 through 99) is transcribed from P_C and is more than 65% G+C. The region from bases 99 through 188, including all the known regulatory sites, is less than 35% G+C, and if the Ntr consensus sequence was removed from the calculation, the region would be less than 20% G+C. The second feature is an inverted repeat located in the right half of the proposed CAP sequence (GCAAT from bases 116 to 120 on the top strand and 111 to 107 on the bottom strand). We initially identified this repeat because the stem-loop structure generated in single-stranded DNA made it difficult to sequence this region. If this stem-loop structure is stable enough to cause severe compressions in our gels containing 7 M urea, we assume it might be significant in RNA conformation. If the weaker of the two P_C promoters (P_{C2}) was used, this sequence would lie at the 5′ end of the RNA transcript, with possible consequences for stability or regulation or both. However, until we know which of the possible P_C promoters is (are) used in vivo, the significance of this interesting structure remains speculative.

ACKNOWLEDGMENTS

We thank L. I. Grossman, in whose laboratory some of these experiments were performed, for his continuing interest in the project; several of our colleagues in the *nif* field for sharing their Ntr consensus sequences with us; and Robert Osuna and Anthony Schwacha for their invaluable help in assuring the accuracy of the sequence and for determining the presence and location of the *Aha*II and *Hga*I sites.

This work was supported by Public Health Service grant AI 15822 from the National Institutes of Health to R.A.B.

LITERATURE CITED

1. **Adams, J., T. Kinney, S. Thompson, L. Rubin, and R. B. Helling.** 1979. Frequency-dependent selection for plasmid-containing cells of *Escherichia coli*. Genetics **91**:627–637.
2. **Ausubel, F. M.** 1984. Regulation of nitrogen fixation genes. Cell **37**:5–6.
3. **Bender, R. A., K. A. Janssen, A. D. Resnick, M. Blumenberg, F. Foor, and B. Magasanik.** 1977. Biochemical parameters of glutamine synthetase from *Klebsiella aerogenes*. J. Bacteriol. **129**:1001–1009.
4. **Bender, R. A., P. M. Snyder, R. Bueno, M. Quinto, and B. Magasanik.** 1983. Nitrogen regulation system of *Klebsiella aerogenes*: the *nac* gene. J. Bacteriol. **156**:444–446.
5. **Boylan, S. A., and R. A. Bender.** 1984. Genetic and physical maps of *Klebsiella aerogenes* for histidine utilization (*hut*). Mol. Gen. Genet. **193**:99–103.
6. **Boylan, S. A., L. J. Eades, K. A. Janssen, M. I. Lomax, and R. A. Bender.** 1984. A restriction enzyme cleavage map of the histidine-utilization (*hut*) genes of *Klebsiella aerogenes* and deletions lacking regions of *hut* DNA. Mol. Gen. Genet. **193**:92–98.
7. **Buck, M.** 1986. Deletion analysis of the *Klebsiella pneumoniae* nitrogenase promoter: importance of spacing between conserved sequences around −12 and −24 for activation by the *nifA* and *ntrC* (*glnG*) products. J. Bacteriol. **166**:545–551.
8. **deCrombrugghe, B., S. Busby, and H. Buc.** 1984. Cyclic AMP receptor protein: role in transcription activation. Science **224**: 831–838.
9. **Goldberg, R. B., R. A. Bender, and S. L. Streicher.** 1974. Direct selection for P1-sensitive mutants of enteric bacteria. J. Bacteriol. **118**:810–814.
10. **Goldberg, R. B., and R. Hanau.** 1980. Regulation of *Klebsiella pneumoniae hut* operons by oxygen. J. Bacteriol. **141**:745–750.
11. **Goldberg, R. B., and B. Magasanik.** 1975. Gene order of the histidine utilization (*hut*) operons in *Klebsiella aerogenes*. J. Bacteriol. **122**:1025–1031.
12. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11**:2237–2255.
13. **Hirschman, J., P.-K. Wong, K. Sei, J. Keener, and S. Kustu.** 1985. Products of nitrogen regulatory genes *ntrA* and *ntrC* of enteric bacteria activate *glnA* transcription *in vitro*: evidence that the *ntrA* product is a sigma factor. Proc. Natl. Acad. Sci. USA **82**:7525–7529.
14. **Holmes, D. S., and M. Quigley.** 1984. A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. **114**:193–197.
15. **Hunt, T. P., and B. Magasanik.** 1985. Transcription of *glnA* by purified *Escherichia coli* components: core RNA polymerase and the products of *glnF*, *glnG*, and *glnL*. Proc. Natl. Acad. Sci. USA **82**:8453–8457.
16. **Lederberg, E. M., and S. N. Cohen.** 1974. Transformation of *Salmonella typhimurium* by plasmid deoxyribonucleic acid. J. Bacteriol. **119**:1072–1074.
17. **Magasanik, B.** 1978. Regulation in the *hut* system, p. 373–387. *In* J. H. Miller and W. S. Reznikoff (ed.), The operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
18. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
19. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
20. **Morrison, D. A.** 1977. Transformation in *Escherichia coli*: cryogenic preservation of competent cells. J. Bacteriol. **132**: 349–351.
21. **Mulligan, M. E., D. K. Hawley, R. Entriken, and W. R. McClure.** 1984. *Escherichia coli* promoter sequences predict *in vitro* RNA polymerase selectivity. Nucleic Acids Res. **12**:789–800.
22. **Nieuwkoop, A. J., S. A. Boylan, and R. A. Bender.** 1984. Regulation of *hutUH* operon expression by the catabolite gene activator protein-cyclic AMP complex in *Klebsiella aerogenes*. J. Bacteriol. **159**:934–939.
23. **Prival, M. J., and B. Magasanik.** 1971. Resistance to catabolite repression of histidase and proline oxidase during nitrogen-limited growth of *Klebsiella aerogenes*. J. Biol. Chem. **246**: 6288–6296.
24. **Salser, W., R. F. Gesteland, and A. Bolle.** 1967. *In vitro* synthesis of bacteriophage lysozyme. Nature (London) **215**: 588–591.
25. **Shalloway, D., T. Kleinberger, and D. M. Livingston.** 1980. Mapping of SV40 DNA replication origin region binding sites for the SV40 T antigen by protection against exonuclease III digestion. Cell **20**:411–422.
26. **Shine, J., and L. Dalgarno.** 1974. The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Sci. USA **71**:1342–1346.
27. **Siebenlist, U., R. B. Simpson, and W. Gilbert.** 1980. *E. coli* RNA polymerase interacts homologously with two different promoters. Cell **20**:269–281.
28. **Tyler, B. M., and R. B. Goldberg.** 1976. Transduction of chromosomal genes between enteric bacteria by bacteriophage P1. J. Bacteriol. **125**:1105–1111.