JOURNAL OF BACTERIOLOGY, Mar. 1987, p. 1061–1070
0021-9193/87/031061-10$02.00/0

Vol. 169, No. 3

Structural Analysis of the *Escherichia coli* K-12 *hisT* Operon by Using a Kanamycin Resistance Cassette

PEGGY J. ARPS AND MALCOLM E. WINKLER*

Department of Molecular Biology, Northwestern University Medical School, Chicago, Illinois 60611

Received 29 August 1986/Accepted 2 December 1986

We constructed a series of recombinant plasmids containing a kanamycin resistance (Km^r) cassette upstream from, within, and downstream from *hisT*, which encodes the tRNA modification enzyme pseudouridine synthase I. These Km^r insertions were then crossed directly into the bacterial chromosome. We determined growth characteristics, assayed in vivo *hisT* expression, and mapped in vivo *hisT* operon transcripts for the Km^r insertion mutants. We also analyzed polypeptides synthesized in minicells from plasmids containing Km^r cassettes. The combined results from these experiments demonstrate new features concerning the structure and expression of the complex operon that contains *hisT*. We show that the minimum size of the operon is approximately 3,500 base pairs and that it contains at least four genes, which are arranged in the order *usg-2* (*pdxB*), *usg-1*, *hisT*, and *dsg-1* and encode polypeptides with apparent molecular masses of 42,000, 45,000, 31,000, and 17,000 daltons, respectively. Of these genes, only the functions of *usg-2* (*pdxB*) and *hisT* are known, and genetic evidence suggests that these two genes do not require *usg-1* or *dsg-1* for function. *usg-2* (*pdxB*) is required for growth of bacteria on minimal medium at 37°C. In contrast, the three genes at the end of the *hisT* operon are dispensable and form a transcription unit that is expressed from a relatively strong internal promoter. The phenotypes of the Km^r insertion mutants and results from gene expression experiments further confirm the position of the internal promoter and locate additional genetic signals in the DNA sequence around *hisT*. The experiments reported here also indicate several interesting properties of the Km^r cassette as a tool for probing complex operons.

The formation of pseudouridine residues in bacterial tRNA molecules is an important model for understanding the genetic and physiological roles played by posttranscriptional modification of stable RNA molecules (see references in reference 13). To learn about the molecular control mechanisms of genes that encode tRNA modification enzymes, we are analyzing the structure and expression of the *hisT* gene of *Escherichia coli* K-12. The gene product of *hisT* is pseudouridine synthase I (PSUI), which synthesizes pseudouridine residues present in the anticodon stem and loop of over half of all bacterial tRNA species (19). Because the function of so many cellular tRNA species is influenced by PSUI-mediated pseudouridine modification, mutations in *hisT* have strong pleiotropic effects on cellular metabolism and regulation (see references in reference 13). For this reason, *hisT* mutants have been particularly valuable in studies of the effects of undermodification on translation (17), attenuation (7, 11), and cellular physiology (11, 18, 19).

Our recent results establish that *hisT* is a member of a complex operon in *E. coli* K-12 strains (1, 13). Using a combination of approaches, we showed that there are at least two genes, which we temporarily designated *usg-2* and *usg-1*, upstream from *hisT* in the operon (Fig. 1). The stop codon of *usg-1* overlaps the start codon of *hisT* by a single nucleotide. This arrangement seems to cause translational coupling between *usg-1* and *hisT*, even though these two genes appear to be structurally, functionally, and evolutionarily unrelated (1). Furthermore, in quasi in vivo systems such as minicells, molar expression of the *usg-1* gene product is at least 10-fold greater than that of PSUI. Although the basis for this differential gene expression is not yet understood, it might in part reflect the preponderance of rare codons used in *hisT* (1). Upstream of *usg-1*, there is a third gene, *usg-2*, that is also part of the *hisT* operon (Fig. 1). An insertion mutation in *usg-2* results in a set of unusual auxotrophic phenotypes, which is the subject of the accompanying paper (2). Expression in vivo of *usg-1* and *hisT* occurs from at least two promoters, P_{up} and P_{int} (Fig. 1). The internal promoter, P_{int}, was localized near the end of *usg-2*. Finally, previous results from S1 nuclease mapping experiments showed that in vivo transcripts extend at least 105 nucleotides beyond the end of the *hisT* coding region to the site labeled *Cla*I (3836) in Fig. 1 and suggested that there are additional genes downstream from *hisT* in the operon (1).

The important approach devised by Walker and his associates (20) to insert antibiotic resistance cassettes at specific positions in the bacterial chromosome potentially should be extremely useful in structural analysis of complex operons such as *hisT*. In this method, a cassette, which consists of an antibiotic resistance gene flanked by polylinker cloning sites, is inserted into a specific chromosomal restriction site contained on a recombinant plasmid. The recombinant plasmid is linearized with a restriction enzyme so that at least 300 base pairs (bp) of chromosomal DNA remain flanking the cassette. Transformation of a *recBC sbc* mutant with the linearized plasmid allows a double recombination to occur that brings the cassette into the bacterial chromosome precisely at the site where it was cloned into the recombinant plasmid. Since antibiotic resistance cassettes are not transposons, insertion mutations constructed by this method are stable and can easily be moved between genetic backgrounds by generalized transduction with P1 bacteriophage.

In previous experiments, we made limited use of this method of specific insertion mutagenesis to construct two mutants that contain kanamycin resistance (Km^r) cassettes in the bacterial chromosome at sites labeled *Hin*dIII (1518) or *Pst*I (2579) in Fig. 1 (1). These two mutants (NU400 and NU399; Table 1) were useful in our initial characterization of

* Corresponding author.

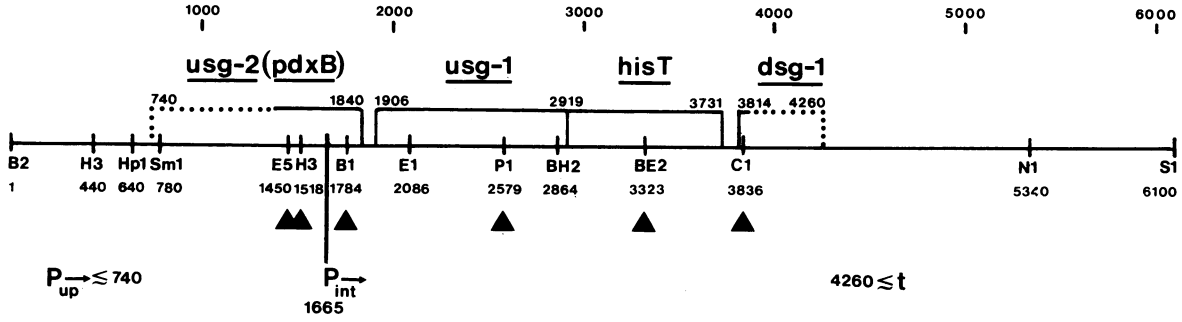

FIG. 1. Restriction map showing the structure of the *hisT* operon. Distances between *Bgl*II (1) and *Eco*RV (1450) and between *Cla*I (3836) and *Sal*I (6100) are based on mapping with restriction enzymes, while distances between *Eco*RV (1450) and *Cla*I (3836) are based on the DNA sequence (1). The boxes and numbers above the boxes show the coding regions of genes in the operon, where solid and dotted lines indicate exact and approximate boundaries, respectively. The position of the internal promoter (P_{int}) and approximate positions of at least one other promoter (P_{up}) and the terminator at the end of the operon (t) are given. The solid triangles (▲) designate the sites where Km^r cassettes were cloned in both possible orientations. Abbreviations: B1, *Bgl*I; B2, *Bgl*II; BE2, *Bst*EII; BH2, *Bss*HII; C1, *Cla*I; E1, *Eco*RI; E5, *Eco*RV; H3, *Hin*dIII; Hp1, *Hpa*I; N1, *Nru*I; P1, *Pst*I; S1, *Sal*I; Sm1, *Sma*I.

the *hisT* operon. In both constructions, transcription of the *kan* and *hisT* genes was in the same direction so that the insertions will be designated *usg-2*::Km^r(*Hin*dIII)> and *usg-1*::Km^r(*Pst*I)>, respectively. The *usg-1*::Km^r(*Pst*I)> mutation abolished *hisT* expression, which appeared to be consistent with translational coupling between expression of these two genes (1, 13). In contrast, the *usg-2*::Km^r(*Hin*dIII)> insertion reduced *hisT* expression by only about 50%. Although this result was important because it revealed the existence of *usg-2*, it was not possible to interpret the remaining *hisT* expression in the *usg-2*::Km^r(*Hin*dIII)> mutant in terms of P_{int} expression, because information was lacking about the polarity characteristics of the Km^r cassette (1).

In this paper, we report the construction of additional strains containing Km^r cassettes inserted in both transcriptional orientations at several positions upstream from, within, and downstream from *hisT* on recombinant plasmids and in the bacterial chromosome. By assaying *hisT* expression in vivo for each of these new constructs and analyzing the expression of gene products in minicell experiments, we identified the reading frame and gene product of *usg-2*, quantitated the relative in vivo expression of P_{up} and P_{int}, and discovered a downstream gene, *dsg-1* (Fig. 1), in the *hisT* operon. In addition, the results reported here reveal several interesting properties of the Km^r cassette as a genetic tool to analyze the structure of complex operons. In the accompanying paper, Km^r cassette insertions are used to show that *usg-2* is actually *pdxB* (2).

MATERIALS AND METHODS

Materials. Restriction endonucleases, T4 DNA ligase used in cloning, and polynucleotide kinase used for 5′ end labeling were purchased from New England BioLabs, Inc. (Beverly, Mass.). DNA polymerase I large fragment used to fill in sites was purchased from Boehringer Mannheim Biochemicals (Indianapolis, Ind.). RNase-free RQ1 DNase and S1 nuclease used to map transcripts were purchased from Promega Biotec (Madison, Wis.) and Bethesda Research Laboratories, Inc. (Gaithersburg, Md.), respectively. Antibiotics and biochemicals were from Sigma Chemical Co. (St. Louis, Mo.) and P-L Biochemicals, Inc. (Milwaukee, Wis.). Ingredients for culture media were from Difco Laboratories (Detroit, Mich.). 3a70B scintillation cocktail and dimethyldichlorosilane-treated glass wool used in PSUI enzyme assays were purchased from Research Products International and Alltech Associates, Inc. (Applied Science Labs, State College, Pa.), respectively. [5-^{3}H]uridine (~20 Ci/mmol) used to make substrate for PSUI enzyme assays was purchased from ICN Pharmaceuticals, Inc. (Irvine, Calif.). [γ-^{32}P]ATP (>5,000 Ci/mmol) used to 5′ end label fragments and L-[^{35}S]methionine (>800 Ci/mmol) used to label polypeptides in minicells were purchased from Amersham Corp. (Arlington Heights, Ill.).

Bacterial strains, media, and growth conditions. The bacterial strains constructed for this study are listed in Table 1. Strains used in the experiments were isogenic derivatives of a single-colony isolate of a W3110 prototroph (NU426; Table 1). Markers were moved between strains by generalized transduction with P1 *kc* phage as described previously (15). Transductants were tested for release of P1 *kc* phage during growth at 42°C (8); none of the strains used in the experiments appears to be a P1 *kc* lysogen. Minicell experiments were performed with transformants of a single-colony isolate of strain P678-54 (NU282; Table 1). Plasmids were introduced into NU282 by a standard transformation protocol that included treatment of bacteria with 10 mM $MgSO_4$ and, in later steps, 50 mM $CaCl_2$ (9, 12). The structure of each recombinant plasmid was rechecked by analysis with restriction enzymes after isolation from the minicell-producing strain.

Bacteria were cultured in LB medium supplemented with 30 μg of L-cysteine per ml (LB + Cys) and in Vogel-Bonner minimal (E) medium supplemented with 0.4% glucose (9). When required, antibiotics were added to growth media at concentrations suggested in references 9 and 12. Other details about culture conditions are given in the tables and figures.

Plasmids. The recombinant plasmids constructed for this study are listed in Table 1. The kanamycin resistance (Km^r) cassette was isolated from plasmid pMB2190, which was obtained from B. Nichols. The Km^r cassette in pMB2190 is formally similar to the kanamycin resistance GenBlock marketed by Pharmacia, Inc. (Piscataway, N.J.). B. Nichols has noted that the Km^r cassette in pMB2190 has acquired a spontaneous internal deletion of about 200 bp in a region between the polylinker and the start of the *kan* gene that contained an inverted repeat (personal communication). Thus, the properties of the Km^r cassette reported in this

TABLE 1. Bacterial strains and plasmids

Strain or plasmid	Genotype[a]	Source[b] or reference
E. coli K-12		
JC7623	*arg ara his leu pro recB21 recC22 sbcB15 thr*	A. J. Clark (20)
NU282	P678-54 *ara azi gal leu lacY malA minA minB rpsL thi thr tonA tsx xyl*	P. Matsumura (3)
NU399	NU426 *usg-1*::Km^r(*Pst*I)>	Arps et al. (1)
NU400	NU426 *pdxB*::Km^r(*Hin*dIII)>	1
NU402	NU426 *pdxB*::<Km^r(*Hin*dIII)	1
NU426	W3110 prototroph	C. Yanofsky collection
NU596	JC7623 *pdxB*::<Km^r(*Bgl*I)	Transformation with linearized pNU97
NU597	JC7623 *pdxB*::Km^r(*Bgl*I)>	Transformation with linearized pNU98
NU598	JC7623 *usg-1*::<Km^r(*Pst*I)	Transformation with linearized pNU99
NU599	JC7623 *pdxB*::<Km^r(*Eco*RV)	Transformation with linearized pNU102
NU600	JC7623 *pdxB*::Km^r(*Eco*RV)>	Transformation with linearized pNU103
NU601	JC7623 *hisT*::<Km^r(*Bst*EII)	Transformation with linearized pNU105
NU602	JC7623 *hisT*::Km^r(*Bst*EII)>	Transformation with linearized pNU106
NU603	JC7623 *dsg-1*::<Km^r(*Cla*I)	Transformation with linearized pNU107
NU604	JC7623 *dsg-1*::Km^r(*Cla*I)>	Transformation with linearized pNU108
NU605	NU426 *pdxB*::<Km^r(*Bgl*I)	NU426X P1 *kc* (NU596)
NU606	NU426 *pdxB*::Km^r(*Bgl*I)>	NU426X P1 *kc* (NU597)
NU607	NU426 *usg-1*::<Km^r(*Pst*I)	NU426X P1 *kc* (NU598)
NU608	NU426 *pdxB*::<Km^r(*Eco*RV)	NU426X P1 *kc* (NU599)
NU609	NU426 *pdxB*::Km^r(*Eco*RV)>	NU426X P1 *kc* (NU600)
NU610	NU426 *hisT*::<Km^r(*Bst*EII)	NU426X P1 *kc* (NU601)
NU611	NU426 *hisT*::Km^r(*Bst*EII)>	NU426X P1 *kc* (NU602)
NU612	NU426 *dsg-1*::<Km^r(*Cla*I)	NU426X P1 *kc* (NU603)
NU613	NU426 *dsg-1*::Km^r(*Cla*I)>	NU426X P1 *kc* (NU604)
Plasmids		
pACYC184	Replicon P15A; Cm^r Tc^r	Chang and Cohen (5)
pBR322	Replicon ColE1; Ap^r Tc^r	Bolivar et al. (4)
pLC28-44	Replicon ColE1; *hisT*$^+$ *purF*$^+$	Koduri and Gots (10)
pMB2190	Km^r in pBR327 derivative; Ap^r Km^r	B. Nichols collection
Ψ210	pBR322 (*hisT*$^+$ *purF*$^+$); Ap^r	Marvel et al. (13)
Ψ300	pBR322 (*Hin*dIII (1518)-*Cla*I (3836)); Ap^r	13
pNU84	pACYC184 (*Hin*dIII (1518)-*Cla*I (3836)); Cm^r	13
pNU85	pNU84 [*usg-1*::Km^r(*Pst*I)>]; Cm^r Km^r	Km^r into *Pst*I (2579)
pNU86	pBR322 (*Hpa*I (640)-*Cla*I (3836)); Ap^r	See Results
pNU93	pBR322 (*Bgl*II (1)-*Cla*I (3836)); Ap^r	See Results
pNU94	pACYC184 (*Bgl*II (1)-*Cla*I (3836)); Cm^r	See Results
pNU97	pNU86 [*pdxB*::<Km^r(*Bgl*I)]; Ap^r Km^r	Km^r into *Bgl*I (1784)
pNU98	pNU86 [*pdxB*::Km^r(*Bgl*I)>]; Ap^r Km^r	Km^r into *Bgl*I (1784)
pNU99	pNU84 [*usg-1*::<Km^r(*Pst*I)]; Cm^r Km^r	Km^r into *Pst*I (2579)
pNU100	pNU94 [*pdxB*::<Km^r(*Eco*RV)]; Cm^r Km^r	Km^r into *Eco*RV (1450)
pNU101	pNU94 [*pdxB*::Km^r(*Eco*RV)>]; Cm^r Km^r	Km^r into *Eco*RV (1450)
pNU102	pNU86 [*pdxB*::<Km^r(*Eco*RV)]; Ap^r Km^r	Km^r into *Eco*RV (1450)
pNU103	pNU86 [*pdxB*::Km^r(*Eco*RV)>]; Ap^r Km^r	Km^r into *Eco*RV (1450)
pNU104	pBR322 (*Hin*dIII (1518)-*Nru*I (5340)); Tc^r	See Results
pNU105	pNU104 [*hisT*::<Km^r(*Bst*EII)]; Km^r Tc^r	Km^r into *Bst*EII (3323)
pNU106	pNU104 [*hisT*::Km^r(*Bst*EII)>]; Km^r Tc^r	Km^r into *Bst*EII (3323)
pNU107	pNU104 [*dsg-1*::<Km^r(*Cla*I)]; Km^r Tc^r	Km^r into *Cla*I (3836)
pNU108	pNU104 [*dsg-1*::Km^r(*Cla*I)>]; Km^r Tc^r	Km^r into *Cla*I (3836)
pNU109	pBR322 (*Hin*dIII (1518)-*Sal*I (6100)); Ap^r	See Results
pNU110	pNU109 [*hisT*::<Km^r(*Bst*EII)]; Ap^r Km^r	Km^r into *Bst*EII (3323)
pNU112	pNU109 [*dsg-1*::<Km^r(*Cla*I)]; Ap^r Km^r	Km^r into *Cla*I (3836)

[a] < or > indicates that the direction of transcription of *kan* in the Km^r cassette is opposite to or the same as *hisT*, respectively. The restriction sites and the numbers used to designate them refer to the map in Fig. 1. For simplicity, *usg-2* (*pdxB*) is referred to as *pdxB* as shown in the accompanying paper (2). Km^r (restriction site) signifies the presence of a kanamycin resistance cassette cloned into that site, whereas Km^r used as a phenotype indicates kanamycin resistance. Cm^r, Chloramphenicol resistant; Tc^r, tetracycline resistant; Ap^r, ampicillin resistant.

[b] In plasmid constructions, two fragments with noncompatible ends were joined by converting sites to blunt ends before ligation (see Materials and Methods).

paper apply to the one isolated from pMB2190 and might not apply to other Km^r cassettes. However, there is a strong possibility that analogous deletions have occurred in other Km^r cassette constructions propagated on multicopy plasmids.

DNA manipulations, restriction fragment isolation, and cloning were done by well-established methods as described previously (1, 13). Ends of restriction fragments containing dissimilar ends were filled in (5′ overhang) or trimmed back (3′ overhang) by reactions catalyzed by the large fragment of DNA polymerase I; the resulting blunt ends could then be ligated together (see references 12 and 13). The structure of each plasmid was verified by extensive analysis with restriction enzymes. When the Km^r cassette was inserted into a site, the integrity of the regions flanking the cassette was checked by determining that the closest possible restriction sites which could be unambiguously analyzed were intact. The direction of *kan* gene transcription, which is known

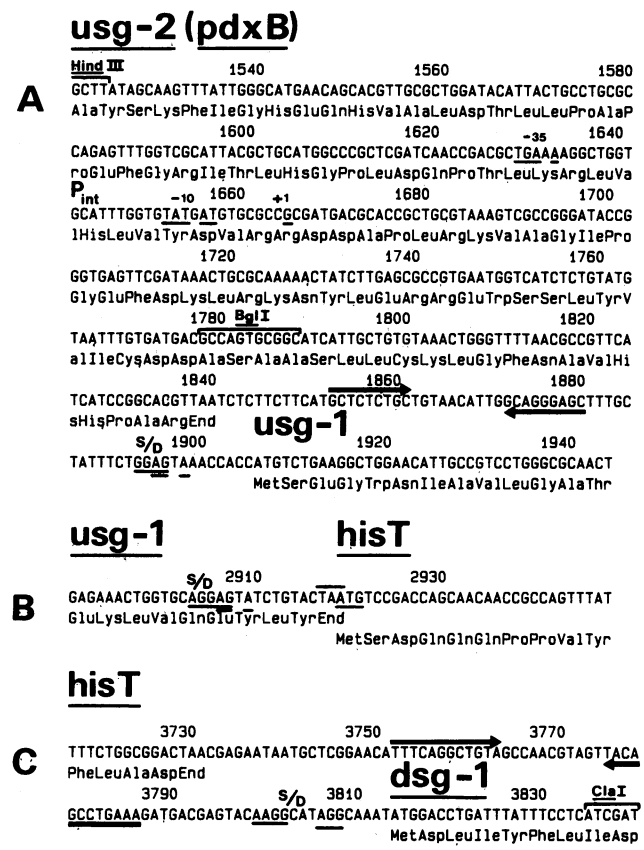

FIG. 2. Nucleotide sequence and genetic features of intercistronic regions in the *hisT* operon. The positions of restriction sites used in constructions, the internal promoter (P_{int}), and Shine-Dalgarno sequences (S/D) that precede coding regions are indicated. Only the start of *usg-1* is putative (see reference 1); all other translation starts and stops are based on experiments described in the text and on an amino acid sequence (PSUI, reference 1). The inverted arrows mark strong dyad symmetries found between coding regions. (A) The end of the *usg-2* (*pdxB*) coding region and the intercistronic region between *usg-2* (*pdxB*) and the most likely start of *usg-1*. (B) The overlap of the stop codon of *usg-1* with the start codon of *hisT*. (C) The intercistronic region between *hisT* and *dsg-1*.

from the DNA sequence (16), was determined by digesting with enzymes that cut both inside the Km^r cassette (usually *Cla*I, *Hin*dIII, or *Xho*I) and in the plasmid.

S1 nuclease mapping of transcripts. Total bacterial RNA that was free of DNA was isolated from bacteria growing exponentially in LB + Cys medium at 37°C by a hybrid of several methods (see reference 1). Briefly, a 40-ml culture was grown to about 5×10^8 cells per ml and then chilled. Bacteria were collected by low-speed centrifugation for 5 min at 4°C, and the pellet was rapidly suspended with vortexing in 10 ml of 30 mM Tris hydrochloride–0.1 M NaCl–1.0 mM EDTA–1.0% (wt/vol) sodium dodecyl sulfate–7.0 M urea (pH 7.3) that was heated to 96°C. After the pellet dissolved, the extract was heated for 90 s longer at 96°C and then poured directly into a flask containing a mixture of 7 ml of $CHCl_3$, 7 ml of H_2O-saturated, redistilled phenol, and 3.0 ml of 1.0 M sodium acetate (pH 5.2) that was preheated to 65°C. The flask was rapidly swirled at 65°C for 4 min. The contents were poured into a 50-ml tube and centrifuged at top speed ($1,520 \times g$) in a swinging bucket rotor of a tabletop centrifuge for 10 min at room temperature. The upper phase was added to a second flask containing 7 ml of $CHCl_3$ and 7 ml of H_2O-saturated phenol preheated to 65°C, and the extraction and centrifugation were repeated. The upper phase was added to a third flask containing 14 ml of $CHCl_3$ preheated to 65°C, and the extraction and centrifugation were completed again. The upper phase was added to a 30-ml Corex centrifuge tube containing 1.3 ml of 3.0 M sodium acetate–1.0 mM EDTA (pH 7.9). After mixing, 7 ml of the solution was distributed to another 30-ml Corex tube, and 21 ml of 100% ethanol was added to each tube. Each RNA pellet was collected by centrifugation, washed with 70% ethanol, and suspended in 0.4 ml of 0.3 M sodium acetate. Each RNA solution was extracted four times with anhydrous ethyl ether, and the RNA was reprecipitated by adding 1.0 ml of 100% ethanol. Each pellet was washed four times with 70% ethanol, dried under vacuum at room temperature, and suspended in 0.4 ml of H_2O. The two RNA solutions were combined, distributed into 100-μl aliquots, and stored at −70°C. The approximate concentration of RNA was determined spectrophotometrically from the A_{260} value.

Before use in S1 nuclease mapping experiments, the total bacterial RNA was treated with RNase-free RQ1 DNase to remove traces of DNA. Total RNA (250 μg) was digested with about 14 U of RQ1 DNase in 40 mM Tris-hydrochloride–10 mM NaCl–6 mM $MgCl_2$ (pH 7.9) for 15 min at 37°C. A 100-μl portion of phenol saturated with 10 mM Tris-hydrochloride–1 mM EDTA (pH 7.9) (TE) was added to kill the reaction. After centrifugation, the upper phase was

TABLE 2. *hisT* expression in bacteria containing Km^r cassettes inserted in or near *hisT*

Strain	Position of Km^r cassette in chromosome[a]	Doubling time (min) in LB + Cys at 37°C[b]	Relative PSUI sp act[c]
1. W3110 prototroph (NU426)	None	27	≡1.0
2. NU608	<*Eco*RV (1450)	27	0.4
NU609	*Eco*RV (1450)>	27	0.4
3. NU402	<*Hin*dIII (1518)	28	0.5
NU400	*Hin*dIII (1518)>	29	0.5
4. NU605	<*Bgl*I (1784)	31	0
NU606	*Bgl*I (1784)>	31	<0.06
5. NU607	<*Pst*I (2579)	35	0
NU399	*Pst*I (2579)>	32	0
6. NU610	<*BstE*II (3323)	33	0
NU611	*BstE*II (3323)>	32	0
7. NU612	<*Cla*I (3836)	27	0.6
NU613	*Cla*I (3836)>	27	0.6

[a] Numbers refer to the restriction map in Fig. 1. < or > indicates that the direction of transcription of *kan* in the Km^r cassette is opposite to or the same as *hisT*, respectively.

[b] Overnight starter cultures of all strains except NU426 contained 50 μg of kanamycin per ml in addition to LB + Cys. Final cultures, which lacked antibiotic, were inoculated with a small volume of starter culture (<2%) and were incubated with vigorous shaking. Values are averages of three independent determinations.

[c] PSUI specific activity was determined for exponentially growing bacteria as described previously (13). PSUI specific activity in the parent strain NU426 was 560 cpm of ^3H released per min per mg of protein of cellular extract prepared by sonic disruption. Values are averages of at least two independent determinations.

removed to an empty tube, the phenol was washed with 100 μl of fresh TE, the wash was combined with the original upper phase, and the RNA solution was extracted with ether. Finally, the RNA was precipitated by adding ethanol, washed with 70% ethanol, dried, and suspended in H_2O. Except for the RNA isolation, the S1 nuclease mapping experiments were performed exactly as described before (1).

Other methods. The Km^r cassette was crossed into the bacterial chromosome from linearized plasmids as described previously (1, 20). The correct location of the Km^r cassette in the chromosome of strains NU399 and NU400 was verified by Southern hybridization analysis (2), which confirmed that the method of specific insertion mutagenesis was working properly in our strains. PSUI enzyme assays and minicell experiments were performed as detailed earlier (13). Antibiotic selection was maintained in all cultures of minicell-producing strains before the isolation of minicells. Molecular weights of the *usg-2*, *dsg-1*, and *dsg-2* polypeptides were determined from sodium dodecyl sulfate-polyacrylamide gels relative to the following standard polypeptides: *cat* (22,500), *trpA* (28,700), *bla* (30,000), *hisT* (30,400), *kan* (30,700), *usg-1* (45,000), *trpB* (45,700), *trpC* (55,700), and *purF* (56,400) (13).

RESULTS

Cloning the *usg-2* and *dsg-1* regions. In earlier experiments, we characterized and determined the nucleotide sequence of the *Hin*dIII (1518)-*Cla*I (3836) restriction fragment that contains *hisT* (1, 13). Throughout this paper, the numbers used to designate restriction sites refer to the map in Fig. 1. In our initial constructions, this 2,318-bp fragment was ligated to a *Hin*dIII-*Cla*I fragment of pBR322 to give plasmid Ψ300 and to a *Hin*dIII-*Bam*HI fragment of pACYC184 to give plasmid pNU84 (Table 1) (13). The *Hin*dIII (1518)-*Cla*I (3836) fragment was originally isolated from a 22-kilobase Clarke-Carbon plasmid, pLC28-44. Therefore, to further characterize the region upstream from the *hisT* operon, we completed a limited restriction map of pLC28-44 upstream of the *Hin*dIII (1518) site (Fig. 1). We then cloned this region by ligating together the following restriction fragments: (i) *Bgl*II (1)-*Cla*I (3836) of pLC28-44 to *Bam*HI-*Cla*I of pBR322 to give pNU93; (ii) *Bgl*II (1)-*Cla*I (3836) of pLC28-44 to *Bam*HI-*Hin*dIII of pACYC184 to give pNU94 (in which the *Cla*I (3836) and *Hin*dIII sites were filled in before ligation); and (iii) *Hpa*I (640)-*Cla*I (3836) of pLC28-44 to *Bal*I-*Cla*I of pBR322 to give pNU86 (Table 1). The minicell experiments described later in this paper and complementation experiments detailed in the accompanying paper (2) confirm that plasmids pNU86, pNU93, and pNU94 each contain an intact, functional copy of *usg-2*.

We isolated the region downstream from *hisT* by starting with the 14-kilobase plasmid Ψ210, which is a subclone of pLC28-44 that was described previously (13). After completing a limited restriction map of Ψ210, we ligated the *Hin*dIII (1518)-*Nru*I (5340) fragment of Ψ210 to the *Hin*dIII-*Sca*I fragment of pBR322 to give pNU104 and the *Hin*dIII (1518)-*Sal*I (6100) fragment of Ψ210 to the *Hin*dIII-*Sal*I fragment of pBR322 to give pNU109 (Fig. 1; Table 1). Results from minicell experiments presented later in this paper suggest that the operon may extend only about 600 bp downstream from *hisT* (Fig. 1).

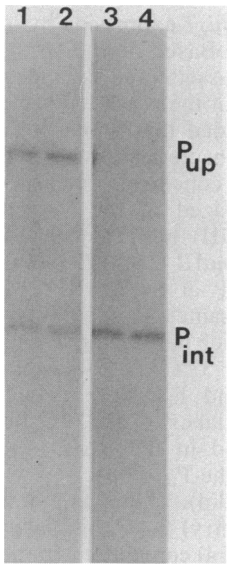

FIG. 3. S1 nuclease mapping of in vivo transcripts synthesized in the W3110 parent (NU426) and in a *usg-2* (*pdxB*)::<Km^r(*Hin*dIII) insertion mutant (NU402). Approximately 100 μg of total bacterial RNA, which was isolated from the two strains during exponential growth in LB + Cys medium at 37°C, was hybridized to 5′-end-labeled template strand of the *Hin*dIII (1518)-*Eco*RI (2086) fragment as described in Materials and Methods. After treatment with S1 nuclease for 15 min (lanes 1 and 3) or 30 min (lanes 2 and 4), samples were analyzed by gel electrophoresis (see Materials and Methods). The positions of protected DNA segments corresponding to transcription initiations at P_{up} and P_{int} are indicated. Lanes 1 and 2, NU426 RNA; lanes 3 and 4, NU402 RNA. Laser densitometer tracing indicated that the ratio of P_{up} to P_{int} transcripts was 1.1 or <0.1 for lanes 1 and 2 or 3 and 4, respectively.

Construction and growth characteristics of Km^r insertion mutants. We cloned single Km^r cassettes into the plasmids described in the preceding section to give insertions at each of the six restriction sites marked by triangles in Fig. 1 (Table 1). For each construction, the integrity of the DNA flanking the Km^r cassette and the orientation of the *kan* gene were checked by analysis with restriction enzymes (see Materials and Methods). Plasmids were obtained with the *kan* gene inserted at each restriction site in the same or in the opposite transcriptional direction as *hisT* (Table 1). These plasmids were then used to cross the Km^r cassettes into the bacterial chromosome by the method outlined in the Introduction. As the final step in these constructions, all Km^r insertions were transduced into an isogenic W3110 prototrophic background.

Bacteria containing a Km^r cassette at *Eco*RV (1450) or *Bgl*I (1784) (Fig. 1) had the same set of unusual auxotrophic phenotypes as those with an insertion at *Hin*dIII (1518). This observation demonstrates that insertion into any of these three closely spaced chromosomal sites most likely inactivates the same gene, which we called *usg-2*. The distinctive physiology of *usg-2*::Km^r mutants is described in the accompanying paper (2). Inspection of the DNA sequence from this region indicates that only one possible open reading frame is consistent with gene inactivation at *Hin*dIII (1518) and *Bgl*I (1784); therefore, this open reading frame must represent the carboxyl terminus of the *usg-2* gene product (Fig. 2; see Discussion).

Mutants with a Km^r insertion at *Pst*I (2579), *Bst*EII (3323), or *Cla*I (3836) (Fig. 1) remained prototrophic and grew on LB + Cys and minimal (E) plus glucose media at 25, 30, 37, and 42°C (data not shown). However, insertion at *Pst*I (2579) or *Bst*EII (3323), both of which are in *usg-1* and *hisT*, respectively (Fig. 1), did cause the bacteria to grow about 15% slower than the W3110 parent in LB + Cys medium at

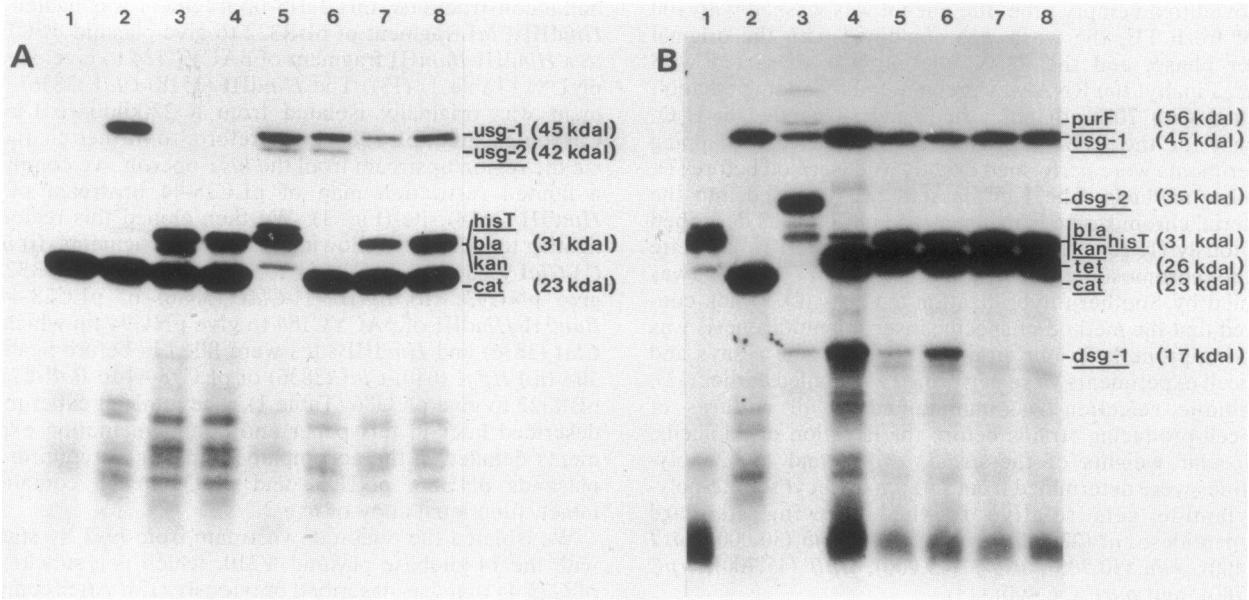

FIG. 4. Polypeptides synthesized in minicells from recombinant plasmids containing Km^r cassettes upstream, within, and downstream from *hisT* in the operon. The plasmids, which are listed in Table 1 and described in the text, were transformed into strain P678-54 (NU282), and minicells were isolated and labeled with [^{35}S]methionine as described in the Materials and Methods. An approximately equal number of trichloroacetic acid-precipitable counts were added to the lanes in each panel. The locations and molecular masses of gene products encoded by chromosomal inserts (*usg-1*, *usg-2*, *hisT*, *dsg-2*, and *dsg-1*), by cloning vectors (*bla*, *cat*, *tet*), and by the Km^r cassette (*kan*) are indicated. (A) Analysis of expression upstream from *hisT* in the operon. Lanes: 1, pACYC84; 2, pNU84; 3, pNU85; 4, pNU99; 5, pNU93; 6, pNU94; 7, pNU100; 8, pNU101. (B) Analysis of expression downstream from *hisT* in the operon. Lanes: 1, pBR322; 2, pNU84; 3, Ψ210; 4, pNU104; 5, pNU105; 6, pNU106; 7, pNU107; 8, pNU108. kdal, Kilodaltons.

37°C (Table 2, entries 1, 5, and 6). This reproducible reduction in growth rate is characteristic of bacteria that lack PSUI activity (11). The *usg-2*::Km^r(*BglI*) mutants also grew about 15% slower than the parent and the other two *usg-2*::Km^r mutants (Table 2, entries 1, 2, 3, and 4). Finally, bacteria containing a Km^r insertion at *Cla*I (3836) had essentially the same growth characteristics as the W3110 parent on all media and at all temperatures tested (Table 2, entries 1 and 7; data not shown). Elsewhere in this paper, we show that *Cla*I (3836) is in a downstream gene of the *hisT* operon, which we temporarily designated *dsg-1*. Clearly, *dsg-1*, like *usg-1* and *hisT*, is dispensable for growth of bacteria on defined media.

Expression of *hisT* in Km^r insertion mutants. Results from previous S1 nuclease mapping experiments suggest that an internal promoter, P_{int}, exists and is located at position 1665 between *Hin*dIII (1518) and *Bgl*I (1784) (Fig. 1) (1). If the interpretation of these experiments was correct, then a transcriptional block placed at *Bgl*I (1784) should eliminate *hisT* expression from P_{up} and P_{int}, whereas a transcriptional block placed at *Eco*RV (1450) or *Hin*dIII (1518) should allow *hisT* to be expressed only from P_{int}. The results presented in Table 2 confirm these predictions. The *usg-2*::<Km^r(*BglI*) insertion eliminated *hisT* expression, while the *usg-2*::Km^r(*BglI*)> insertion reduced *hisT* expression by at least 17-fold (Table 2, entry 4). This result shows that the Km^r cassette is extremely polar in the bacterial chromosome and that there are no additional promoters between P_{int} and *hisT*. Furthermore, elimination of *hisT* expression accounts for the reduced growth rate of *usg-2*::Km^r(*BglI*) mutants (Table 2) (see above). Finally, the lack of *hisT* expression in *usg-2*::Km^r(*BglI*) mutants shows that P_{int} occurs in the coding region of *usg-2* as depicted in Fig. 1 and 2.

In contrast, *usg-2*::Km^r(*Eco*RV) and *usg-2*::Km^r(*Hin*dIII) insertions reduced *hisT* expression by only 50 to 60% (Table 2, entries 2 and 3). Based on the strong polarity characteristics of the Km^r cassette and the independence of the data on the orientation of the cassette (Table 2, entries 2 and 3), it might be concluded that about 50% of *hisT* expression originates at P_{int} in bacteria growing in LB + Cys medium at 37°C. To test this conclusion further, we mapped in vivo transcripts synthesized in the W3110 parent strain and *usg-2*::<Km^r(*Hin*dIII) mutant with S1 nuclease (Fig. 3). Single-stranded *Hin*dIII (1518)-*Eco*RI (2086) DNA that was 5′ labeled with ^{32}P at its *Eco*RI end was hybridized to approximately the same amount of total RNA prepared from the two strains, which had been growing exponentially in LB + Cys medium at 37°C. In the parent strain, transcripts initiated at P_{up} and P_{int} were detected in nearly equal amounts (Fig. 3, lanes 1 and 2). In contrast, the only transcripts detected in the *usg-2*::<Km^r(*Hin*dIII) mutant originated at P_{int}; the P_{up} transcripts were completely gone (Fig. 3, lanes 3 and 4). The apparent slight increase (~1.3-fold by densitometry) in P_{int}-initiated transcripts in the mutant (lanes 3 and 4) compared with the parent (lanes 1 and 2) might simply reflect small differences in the amount of total RNA used in each hybridization reaction. Thus, these data support the conclusion that about 50% of *hisT* transcription initiates at P_{int} in bacteria growing in rich medium. In addition, the data directly confirm that the Km^r cassette inserted into the bacterial chromosome can cause strong transcriptional polarity.

On the basis of the ability of the Km^r cassette to cause polarity and to disrupt genes, we expected insertion into the coding regions of *usg-1* or *hisT* to eliminate *hisT* expression (Table 2, entries 5 and 6). However, it should be noted that because the polarity caused by the Km^r cassette appears to be so strong, the lack of *hisT* expression in *usg-1*::Km^r(*Pst*I)

mutants can no longer be interpreted as indirect evidence for translational coupling between expression of *usg-1* and *hisT* (see Introduction and reference 1). Surprisingly, insertion of the Km^r cassette into *Cla*I (3836), which is 105 bp downstream from *hisT*, also reproducibly reduced *hisT* expression in vivo (Table 2, entry 7). In the next section, we show that *Cla*I (3836) is contained in the coding region for a 17,000-dalton *dsg-1* polypeptide. Therefore, it is possible that the *dsg-1* gene product is necessary for full *hisT* expression. However, this explanation is unlikely for several reasons (see Discussion); instead, we think it more likely that *hisT* and *dsg-1* are encoded by a polycistronic mRNA molecule whose stability is decreased by insertion of the Km^r cassette. Thus, based on our results it appears that insertion of the Km^r cassette into a multigene operon can reduce in vivo expression of both upstream and downstream genes.

Expression of *hisT* operon genes in minicells. We reasoned that the knockout mutations and polarity caused by the Km^r cassettes in vivo could also be used to study the structure of the *hisT* operon in recombinant plasmids. In addition, we wanted to identify the *usg-2* and putative *dsg-1* gene products and to determine their expression relative to the *usg-1* and *hisT* gene products. Consequently, we examined the polypeptides synthesized in minicells from the recombinant plasmids and their Km^r derivatives described above. To our surprise, the strong polarity caused by the Km^r cassette in the bacterial chromosome was not apparent for expression from plasmids in minicells. Further analysis demonstrated that reduced polarity is a property of the minicell system, since the Km^r cassette caused strong polarity on expression from the same recombinant plasmids contained in exponentially growing bacteria rather than in minicells.

Plasmid pNU84 contains the *Hin*dIII (1518)-*Cla*I (3836) fragment in vector pACYC184 and expressed the *usg-1* and *hisT* (PSUI) gene products as reported before (Fig. 4A, lane 2) (13). Insertion of the Km^r cassette into pNU84 at *Pst*I (2579) eliminated the *usg-1* gene product (Fig. 4A, lanes 3 and 4) and confirmed that this region encodes *usg-1* (Fig. 1). Concomitant changes in the amount of *hisT* gene product (PSUI) could not be measured in these experiments because the *kan* and *hisT* polypeptides comigrated in our gel system. Figure 4A also shows the polypeptides synthesized from plasmids pNU93 (lane 5) and pNU94 (lane 6), which contain the *Bgl*II (1)-*Cla*I (3836) fragment cloned into pBR322 and pACYC184, respectively. Comparison of polypeptides synthesized from pACYC184 (lane 1), pNU84 (lane 2), and pNU94 (lane 6) revealed that a 42,000-dalton polypeptide was synthesized from the region upstream of *usg-1* and *hisT*. Furthermore, insertion of the Km^r cassette into the *Eco*RV (1450) site of pNU94 caused the 42,000-dalton polypeptide to disappear (Fig. 4A, lanes 7 and 8). Since we established above that *Eco*RV (1450) is in *usg-2*, we conclude that this 42,000-dalton polypeptide is the *usg-2* gene product. Because the stop codon of *usg-2* occurs at position 1,840 (Fig. 2 and Discussion), the approximate molecular mass of 42,000 daltons for the *usg-2* gene product localizes P_{up} near or before position 740 (Fig. 1).

In analogous minicell experiments, we found that plasmid pNU86, which contains the *Hpa*I (640)-*Cla*I (3836) fragment in vector pBR322, also expressed the 42,000-dalton polypeptide (data not shown). This result is consistent with the observation that pNU86 and pNU94 complement *usg-2* mutants (2). However, the ratio of *usg-2* polypeptide to *usg-1* polypeptide was noticeably less from pNU86 than from pNU94 (Fig. 4A, lane 6; data not shown). One explanation for this lower relative expression is that the 640-bp region upstream of *usg-2*, which is present in pNU94 but not in pNU86, contains a promoter or other genetic element that enhances *usg-2* expression. This conjecture remains to be tested. Insertion of the Km^r cassette at *Eco*RV (1450), *Hin*dIII (1518), or *Bgl*I (1784) in pNU86 eliminated the 42,000-dalton polypeptide and again supports the conclusion that this polypeptide is the *usg-2* gene product. Unexpectedly, the *usg-2*::<Km^r(*Bgl*I) insertion, which eliminated chromosomal *hisT* expression (Table 2), was only moderately polar on *usg-1* expression from plasmids in minicells (data not shown). This discrepancy in polarity levels caused by insertion of the Km^r cassette into the chromosome versus into plasmids in minicells is considered below. Nevertheless, because the pNU86 [*usg-2*::<Km^r(*Bgl*I)] plasmid (pNU97) was used to construct the *usg-2*::<Km^r(*Bgl*I) mutant (NU605), the expression of intact *usg-1* gene product from the plasmid verifies that the lack of *hisT* expression in the mutant (NU605; Table 2) was not caused by an undetected disruption of *usg-1* that occurred during the constructions.

Figure 4B presents an analysis of expression of genes located downstream from *hisT*. Plasmid pNU104, which contains the *Hin*dIII (1518)-*Nru*I (5340) fragment in vector pBR322, expressed the *usg-1* and *hisT* gene products and a 17,000-dalton polypeptide (Fig. 4B, lane 4). Insertion of the Km^r cassette into pNU104 at *Cla*I (3836) specifically eliminated the 17,000-dalton polypeptide (Fig. 4B, lanes 7 and 8). Other minicell experiments showed that a 2-bp insertion mutation constructed by filling in the *Cla*I (3836) site of pNU104 also specifically eliminated the 17,000-dalton polypeptide (data not shown). Because in vivo transcripts extend from *hisT* to at least *Cla*I (3836) (see Introduction), we conclude that the *dsg-1* gene that encodes the 17,000-dalton polypeptide is part of the *hisT* operon. Furthermore, inspection of the DNA sequence between the end of *hisT* and *Cla*I (3836) indicates that only one translational start codon can account for knockout and frameshift mutations at *Cla*I (3836). Therefore, this start codon must correspond to the amino terminus of the *dsg-1* gene product (Fig. 2; see Discussion), and *dsg-1* must extend to about position 4,260 to accommodate the coding region for a 17,000-dalton polypeptide (Fig. 1).

Insertion of the Km^r cassette into pNU104 at *BstE*II (3323) substantiated that *hisT* and *dsg-1* are members of the same operon. The *hisT*::<Km^r(*BstE*II) insertion in pNU104 was moderately polar on *dsg-1* expression and reduced the amount of *dsg-1* polypeptide by about three- to fourfold (Fig. 4B, lane 5). In contrast, the *hisT*::Km^r(*BstE*II)> insertion in pNU104 did not appreciably affect *dsg-1* expression (Fig. 3, lane 6). Thus, unlike chromosomal insertions (Table 2), the Km^r cassette is only moderately polar in one orientation on expression from plasmids in minicells.

We wanted to learn whether this reduced polarity is a property of the minicell system or a general property of expression from our plasmid constructs. To distinguish between these possibilities, we took advantage of the absence of chromosomal *hisT* gene expression in strains NU605 and NU606, which contain the *usg-2*::<Km^r(*Bgl*I) and *usg-2*::Km^r(*Bgl*I)> insertions, respectively (Table 1). If the Km^r cassette is strongly polar on downstream gene expression in plasmids pNU97 (pNU86 [*usg-2*::<Km^r(*Bgl*I)]) and pNU98 (pNU86 [*usg-2*::Km^r(*Bgl*I)>]), then strains NU605(pNU97) and NU606(pNU98) should contain much lower PSUI activity than strains containing the multicopy pNU86 parent plasmid (13). In contrast, if the Km^r cassette is only weakly polar on downstream gene expression from plasmids as we observed in minicells, then

PSUI activity in strains NU605(pNU97) and NU606(pNU98) should be high compared with the single-copy, chromosomal level. The results from this type of experiment showed that the Km^r cassette reduced plasmid-encoded *hisT* expression by 38-fold in strain NU605(pNU97) (<Km^r) and by 16-fold in strain NU606(pNU98) (Km^r>) (data not shown). Analogous experiments with strains NU607(pNU99) (*usg-1*::<Km^r(*Pst*I)) and NU399(pNU85) (*usg-1*::Km^r(*Pst*I)>) also indicated extremely strong polarity (>76-fold), although this effect might in part reflect loss of translational coupling between *usg-1* and *hisT*. Thus, we conclude that the relative lack of polarity on downstream gene expression caused by insertion of the Km^r cassette into plasmids is a property of the minicell system rather than the plasmids themselves.

Returning to Fig. 4B, comparison of the polypeptides synthesized from pNU104 (lane 4) with those from plasmid pNU109, which contains the *Hin*dIII (1518)-*Sal*I (6100) fragment in vector pBR322, indicates that a 35,000-dalton polypeptide is encoded by the region that spans *Nru*I (5340) (data not shown). The band that corresponds to this second downstream (*dsg-2*) gene product is particularly prominent and, as expected, was also produced by parent plasmid Ψ210, which extends far beyond *Sal*I (6100) to at least *purF* (Fig. 4B, lane 3) (13). A *hisT*::<Km^r(*Bst*EII) insertion in pNU109 reduced the amount of *dsg-1* polypeptide by three- to fourfold but did not affect the amount of *dsg-2* gene product (data not shown). Similarly, a *dsg-1*::<Km^r(*Cla*I) insertion in pNU109 did not affect the amount of *dsg-2* polypeptide. Therefore, *dsg-2* does not seem to be part of the *hisT* operon, and *dsg-1* is possibly the last gene in the *hisT* operon. This issue should be definitively resolved by DNA sequence analysis and S1 nuclease mapping experiments which are currently in progress (C. Marvel, personal communication).

Finally, as noted before, there appears to be strong differential expression between *usg-1* and *hisT* in minicells, in maxicells, and in a coupled in vitro transcription-translation system (13). The minicell experiments presented in this paper extend this observation by showing that the *hisT* gene product appears to be synthesized at a much lower level than the other three polypeptides encoded by the operon (Fig. 4A, lanes 2 and 6, and B, lanes 2 and 4). Densitometer tracing of gels containing polypeptides labeled with [^{35}S]methionine in minicells indicates that the *usg-2* (*pdxB*), *usg-1*, *hisT*, and *dsg-1* polypeptides were synthesized in a relative ratio of 0.3:1.0:0.07:0.6. Thus, even without correcting for the unknown methionine contents of *usg-2* (*pdxB*) and *dsg-1*, it is clear that *hisT* seems to be expressed at a much lower level than adjoining genes in the operon. Experiments currently in progress should determine whether differential expression of the *hisT* operon actually occurs in exponentially growing bacteria.

DISCUSSION

The results reported in this paper reveal new aspects about the structure and expression of the complex operon that encodes the tRNA modification enzyme PSUI. Besides *hisT*, the operon contains at least three additional genes, which we have temporarily called *usg-2*, *usg-1*, and *dsg-1* (Fig. 1). Of these four genes, only *usg-2* is absolutely required for growth of bacteria on minimal (E) plus glucose medium at 37°C. The other three genes, *usg-1*, *hisT*, and *dsg-1*, are dispensable for growth on defined media at temperatures ranging from 25 to 42°C (see Results). However, *usg-2*, *usg-1*, and *hisT* mutants that lack PSUI enzyme activity did grow slower than their W3110 parent (Table 2). Presently, we do not know the functions of *usg-1* or *dsg-1*, although the mutants constructed for this study should help identify these genes. In the accompanying paper, we describe the unusual phenotypes of *usg-2* mutants, demonstrate that *usg-2* is *pdxB*, which encodes an enzyme involved in the biosynthesis of the essential coenzyme pyridoxal phosphate, and consider possible reasons for grouping *pdxB* and *hisT* together in the same operon (2).

The results presented here also show that *usg-2* (*pdxB*) starts about 2,200 bp upstream from the start of *hisT* (Fig. 1) and encodes a 42,000-dalton polypeptide (Fig. 4). At least one promoter (P_{up}) is located upstream of *usg-2* (*pdxB*) (Fig. 1 and 3). We are presently determining whether *usg-2* (*pdxB*) is the first gene in the *hisT* operon or whether additional genes precede it. The positions of insertion mutations that inactivate *usg-2* (*pdxB*) expression in vivo are consistent with only one open reading frame in the DNA sequence from the region containing the end of *usg-2* (*pdxB*). The 106 amino acids that form the carboxyl terminus of the *usg-2* (*pdxB*) polypeptide are shown in Fig. 2A. Unlike the *hisT* coding region, this segment, which makes up about 10% of *usg-2* (*pdxB*), is not particularly rich in rare codons. Expression of the *usg-2* (*pdxB*) polypeptide in minicells seems to be comparable to expression of the *usg-1* and *dsg-1* polypeptides, whereas expression of PSUI occurs at a much lower level (Fig. 4). The rare codon usage found in *hisT* has been suggested as a factor that contributes to this pattern of differential expression (1), although this notion needs to be rigorously tested.

Transcription of *usg-1*, *hisT*, and *dsg*-1 is initiated in vivo from P_{up} and from an internal promoter (P_{int}) located in the *usg-2* (*pdxB*) coding region 176 bp upstream from the end of the gene (Fig. 2A). Our data indicate that P_{int} is relatively strong in vivo; fully half of the *hisT* transcripts initiate at P_{int} in bacteria growing exponentially in rich medium at 37°C (Table 2; Fig. 3). We noted before that P_{int} contains a sequence similar to the discriminator region of stringently controlled promoters (1). The mutants constructed in this study will be useful in experiments to analyze P_{int} function independent of P_{up} (Table 2).

Inspection of the DNA sequence indicates that the end of *usg-2* (*pdxB*) and the putative start of *usg-1* are separated by a 65-bp intercistronic region (Fig. 2A). For comparison, Fig. 2B also shows the overlapping stop and start codons of *usg-1* and *hisT*; an arrangement that appears to cause translational coupling between expression of these two genes (1, 13). Results from minicell experiments confirm that *usg-1* encodes a polypeptide which migrates on sodium dodecyl sulfate-polyacrylamide gels with an apparent molecular mass of 45,000 daltons (Fig. 4) (13). Furthermore, the combined results from previous S1 nuclease mapping experiments (1) and the current gene disruption (Table 2 and text) and minicell experiments (Fig. 4) show that there is another gene in the operon (*dsg-1*) which must start 82 bp downstream from the end of *hisT* (Fig. 2C). Because *dsg-1* encodes a 17,000-dalton polypeptide (Fig. 4), the operon extends at least 500 bp beyond the end of *hisT* (Fig. 1). In addition, the lack of polarity caused by insertions in *hisT* or *dsg-1* on expression of another gene downstream from *dsg-1* suggests that *dsg-1* may be the last gene in the operon (see Results); thus, the minimum size of the *hisT* operon is approximately 3,500 bp (Fig. 1). Finally, the intercistronic region between *hisT* and *dsg-1*, like the one between *usg-2* (*pdxB*) and *usg-1*, contains a segment of especially strong dyad symmetry (Fig.

2A and C); however, we presently do not know whether these structures play functional roles.

As noted above, the last three genes of the *hisT* operon are transcribed from both P_{up} and P_{int}. In addition, these same three genes are dispensable for growth of the bacteria under standard laboratory conditions. Earlier results showed that *usg-1* and *hisT* appear to be functionally unrelated, even though they are closely associated in the operon (1). Our current results suggest that *dsg-1* and *hisT* are also functionally unrelated. Insertion mutations right at the start of the *dsg-1* coding region completely eliminate the *dsg-1* gene product (Fig. 4B). Surprisingly, these insertions reduce in vivo expression of the upstream *hisT* gene (Table 2); however, the effect is less than twofold. If intact *dsg-1* gene product was required for *hisT* expression, we would expect a much greater reduction in PSUI activity in *dsg-1* knockout mutants. In addition, a multicopy plasmid that contains *hisT* but lacks *dsg-1* (Ψ300) overexpresses PSUI to the same extent as a related plasmid that contains an intact copy of both genes (Ψ210) (13); again, it appears that the *dsg-1* gene product is not required for PSUI function. We will begin to understand why *usg-1*, *hisT*, and *dsg-1* are organized into a transcription unit under the control of P_{int} when we identify *usg-1* and *dsg-1*.

The results discussed in this paper are based on the use of a kanamycin resistance (Km^r) cassette to disrupt genes in the bacterial chromosome or on recombinant plasmids. Besides providing new information about the structure of the *hisT* operon, these experiments indicate several interesting properties of the Km^r cassette as a genetic tool to probe complex operons. In either orientation, the Km^r cassette seems to be strongly polar in exponentially growing cells on chromosome (Table 2)- and plasmid-encoded (see text) gene expression downstream from the gene containing the insertion. In addition, expression of genes upstream of the Km^r cassette can be reduced (Table 2), perhaps by destabilizing the 3′ end of mRNA molecules. Thus, in the bacterial chromosome and in recombinant plasmids contained in exponentially growing cells, the Km^r cassette can be a powerful tool for showing the relationship between genes in complex operons.

By contrast, the Km^r cassette is moderately polar on downstream gene expression from plasmids in minicells only when *kan* gene and operon transcription are in opposite directions (Fig. 4). When *kan* gene and operon transcription are in the same direction, expression of downstream genes from plasmids often increases (Fig. 4). Experiments described in the Results indicate that reduced polarity seems to be a property of the minicell system. This notion is supported by our observation that the classical polarity expected from certain small internal deletions in the *hisT* operon is also absent from plasmids in minicells. We suspect that the Km^r cassette supplies transcription termination sites that are highly efficient in exponentially growing cells, but not in minicells. One likely possibility is that termination factors such as rho become depleted during isolation of the minicells; this conjecture remains to be tested. We also have not investigated whether apparent lack of transcription termination on plasmids is specific to minicells prepared from strain P678-54.

Another property of the Km^r cassette used in this study is the lack of significant levels of transcription from its ends. Roth and his associates (6) observed that transcription emerges from the ends of several commonly used transposons. Consequently, these transposons are polar on downstream gene expression only when they insert upstream of strong rho-dependent termination sites. For the Km^r cassette, insertion into Rosenberg promoter analysis vectors (14) shows that transcription emerges from the Km^r cassette when *kan* and *galK* are in the same orientation; however, when *kan* and *galK* are in the opposite orientation, there is no apparent *galK* transcription initiated from within the Km^r cassette (J. Malakooti and P. Matsumura, personal communication). The data presented here are consistent with the latter result. When *kan* gene and *hisT* operon transcription are in opposite directions, there is complete polarity on downstream gene expression in exponentially growing cells (NU605; Table 2 and text). When *kan* gene and *hisT* operon transcription are in the same direction, there is still extremely strong polarity; however, now there is very slight downstream gene expression (NU606; Table 2 and text). This knowledge of the polarity characteristics of the Km^r cassette in exponentially growing bacteria and in minicells should be of practical importance when applying the new methods of specific insertion mutagenesis.

ACKNOWLEDGMENTS

We thank B. Bachmann, B. Nichols, and C. Yanofsky for bacterial strains and plasmids, M. Connolly and B. Roa for help with genetic constructions, R. Rownd and D. Womble for access to equipment, and M. Johnson and L. Taylor for help with preparation of the manuscript. We also thank C. Marvel, P. Matsumura, and B. Nichols for stimulating discussions and communicating unpublished information.

This work was supported by grant DMB-8417005 from the National Science Foundation and grant 82-41 from the Illinois Division of the American Cancer Society.

LITERATURE CITED

1. **Arps, P. J., C. C. Marvel, B. C. Rubin, D. R. Tolan, E. E. Penhoet, and M. E. Winkler.** 1985. Structural features of the *hisT* operon of *Escherichia coli* K-12. Nucleic Acids Res. **13**:5297–5315.
2. **Arps, P. J., and M. E. Winkler.** 1986. An unusual genetic link between vitamin B_6 biosynthesis and tRNA pseudouridine modification in *Escherichia coli* K-12. J. Bacteriol. **169**:1071–1079.
3. **Bartlett, D. H., and P. Matsumura.** 1984. Identification of *Escherichia coli* region III flagellar gene products and description of two new flagellar genes. J. Bacteriol. **160**:577–585.
4. **Bolivar, F., R. L. Rodriquez, P. J. Greene, M. C. Betlach, H. L. Heynecker, H. W. Boyer, J. H. Crosa, and S. Falkow.** 1977. Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene **2**:95–113.
5. **Chang, A. C. Y., and S. N. Cohen.** 1978. Construction and characterization of the amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. **134**:1141–1156.
6. **Ciampi, M. S., M. B. Schmid, and J. R. Roth.** 1982. Transposon Tn10 provides a promoter for transcription of adjacent sequences. Proc. Natl. Acad. Sci. USA **79**:5016–5020.
7. **Cortese, R., R. Landsberg, R. A. VonderHaar, H. E. Umbarger, and B. N. Ames.** 1974. Pleiotropy of *hisT* mutants blocked in pseudouridine synthesis in tRNA: leucine and isoleucine-valine operons. Proc. Natl. Acad. Sci. USA **71**:1857–1861.
8. **Curtiss, R.** 1981. Gene transfer, p. 243–265. *In* P. Gerhardt, R. G. E. Murray, R. N. Costilow, E. W. Nester, W. A. Wood, N. R. Krieg, and G. B. Phillips (ed.), Manual of methods for general bacteriology. American Society for Microbiology, Washington, D.C.
9. **Davis, R. W., D. Botstein, and J. R. Roth.** 1980. Advanced bacterial genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
10. **Koduri, R. K., and J. S. Gots.** 1980. A DNA-binding protein with specificity for *pur* genes in *Escherichia coli*. J. Biol. Chem. **255**:9594–9598.
11. **Lewis, J. A., and B. N. Ames.** 1971. Histidine regulation in *Salmonella typhimurium*. XI. The percentage of transfer

RNA^{His} charged *in vivo* and its relation to repression of the histidine operon. J. Mol. Biol. **66**:131–142.
12. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
13. **Marvel, C. C., P. J. Arps, B. C. Rubin, H. O. Kamen, E. E. Penhoet, and M. E. Winkler.** 1985. *hisT* is part of a multigene operon in *Escherichia coli* K-12. J. Bacteriol. **161**:60–71.
14. **McKenney, K., H. Shimatake, D. Court, U. Schmeissner, C. Brady, and M. Rosenberg.** 1981. A system to study promoter and terminator signals recognized by *Escherichia coli* RNA polymerase, p. 383–415. *In* J. C. Chirikjan and T. S. Papas (ed.), Gene amplification and analysis, vol. 2. Elsevier/North-Holland Publishing Co., New York.
15. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. **Oka, A., H. Sugisaki, and M. Takanami.** 1981. Nucleotide sequence of the kanamycin resistance transposon Tn903. J. Mol. Biol. **147**:217–226.
17. **Palmer, D. T., P. H. Blum, and S. W. Artz.** 1983. Effects of the *hisT* mutation of *Salmonella typhimurium* on translation elongation rate. J. Bacteriol. **153**:357–363.
18. **Spadaro, A., A. Spena, A. Santonastaso, and P. Donini.** 1981. Stringency without ppGpp accumulation. Nature (London) **291**:256–258.
19. **Turnbough, C. L., R. J. Neill, R. Landsberg, and B. N. Ames.** 1979. Pseudouridylation of tRNAs and its role in regulation in *Salmonella typhimurium*. J. Biol. Chem. **254**:5111–5119.
20. **Winans, S. C., S. J. Elledge, J. H. Krueger, and G. C. Walker.** 1985. Site-directed insertion and deletion mutagenesis with cloned fragments in *Escherichia coli*. J. Bacteriol. **161**:1219–1221.