JOURNAL OF BACTERIOLOGY, Feb. 1989, p. 784–790
0021-9193/89/020784-07$02.00/0

Vol. 171, No. 2

Nucleotide Sequence and Transcriptional Organization of the *Escherichia coli* Enterobactin Biosynthesis Cistrons *entB* and *entA*

MARY SCHRODT NAHLIK, TIMOTHY J. BRICKMAN, BRADLEY A. OZENBERGER,† AND MARK A. McINTOSH*

Department of Microbiology, School of Medicine, University of Missouri-Columbia, Columbia, Missouri 65212

Received 5 August 1988/Accepted 2 November 1988

The nucleotide sequence of a 2,137-base-pair DNA fragment expressing enterobactin biosynthesis functions defined the molecular boundaries and translational products of the *entB* and *entA* genes and identified a closely linked downstream open reading frame encoding an uncharacterized protein of approximately 15,000 daltons (P15). The sequence revealed that an independent protein-coding sequence corresponding to an EntG polypeptide was not situated in the genetic region between the *entB* and *entA* cistrons, to which the EntG$^-$ phenotype had been genetically localized. As a result, the biochemical nature of the EntG function in the biosynthetic pathway requires reevaluation. The EntA polypeptide displayed significant similarities at the amino acid level to the pyridine nucleotide-binding domains of several members of a family of alcohol-polyol-sugar dehydrogenase enzymes, consistent with its function as the enzyme catalyzing the final step of dihydroxybenzoate biosynthesis. An additional role for EntA in the isochorismate synthetase activity of EntC was strongly implicated by genetic evidence. Evidence from the nucleotide sequence of this region and newly constructed *ent-lacZ* fusion plasmids argues strongly that these genes are linked in an iron-regulated *entCEBA* (*P15*) polycistronic operon.

Genes encoding the three enzymatic activities (EntC, EntB, and EntA) catalyzing the synthesis of 2,3-dihydroxybenzoic acid from chorismic acid have been localized to the right end of the enterobactin gene cluster at min 13 on the *Escherichia coli* chromosome by molecular cloning (12, 16, 22) and transposon mutagenesis (13, 16) studies. Interspersed among these genes were two additional loci encoding activities (EntE and EntG) necessary to convert 2,3-dihydroxybenzoic acid and L-serine to the functional catechol siderophore enterobactin (also designated enterochelin), a cyclic triester of 2,3-dihydroxybenzoylserine. On the basis of genetic complementation studies using the *ent* mutant strains which originally delineated this biosynthetic pathway (14, 27, 31), the order *entEBG(AC)* was deduced for this genomic region (16). These studies identified the *entE* and *entB* gene products, indicated that synthesis of a 26,000-dalton (Da) protein was necessary to restore siderophore production to strains carrying the *entC401* or *entA403* allele, and localized the gene encoding a 15-kDa protein of unknown function to the region just downstream of that encoding the multifunctional 26-kDa protein.

However, the accumulated molecular data have not resolved several important questions concerning the functions encoded by this genetic region and its transcriptional organization. No protein product specific for EntG activity has been identified, although the function has been genetically localized (13, 16). Genetic lesions defining the EntA and EntC enzymes have been linked to the single locus encoding the 26-kDa protein (13, 16), in contrast to biochemical evidence that the enzymatic activities are separable by ion-exchange chromatography (30). Finally, conflicting evidence has been presented concerning the transcriptional linkage of these *ent* genes. It is clearly established that chromosomal *ent-lacZ* fusions are iron regulated (6) and that isolated *ent* loci on multicopy plasmids can be expressed to the extent of measuring genetic complementation in low-iron growth studies (12, 13, 22) or siderophore production assays (16). However, it is not clear whether this independent expression is iron inducible and therefore reflects the natural in vivo transcriptional configuration or simply results from detectable expression levels mediated by nonspecific promoter elements functioning in the recombinant plasmid constructs.

In this study, we determined the nucleotide sequence of a 2,137-base-pair (bp) region from within the *entE* gene to downstream of the locus encoding the uncharacterized 15-kDa protein and identified three contiguous open reading frames corresponding to the products of *entB*, *entA*, and the 15-kDa protein gene. An open reading frame that would represent *entG* does not exist in this genetic sequence. Data bank searches identified regions of similarity between EntA and the cofactor-binding domains of other NADP-dependent polyol-alcohol-sugar dehydrogenases. The accompanying report (19) describes the nucleotide sequence of a 1,400-bp region upstream of *entE* which encodes the EntC protein and displays extensive similarities to the aminobenzoate synthetases, TrpE and PabB. The prospect that EntA and EntC function together to convert chorismate to isochorismate is discussed. The DNA sequence data herein, along with *ent-lacZ* fusion experiments, provide strong evidence for an iron-inducible *entCEBA* (*P15*) polycistronic operon controlled by the genetic region upstream of *entC*.

MATERIALS AND METHODS

Bacterial strains, plasmids, and bacteriophage. *E. coli* strains used in this study were HB101 (*hsdS20 recA13 ara proA2 lacY1 galK2 rpsL20 xyl mtl supE44*) (2), MC4160 [*araD rpsL thi* Δ(*lacIPOZYA*)U169 Δ(*srlR-recA*)] (21), and

* Corresponding author.
† Present address: Department of Biology, Yale University, New Haven, CT 06511-8112.

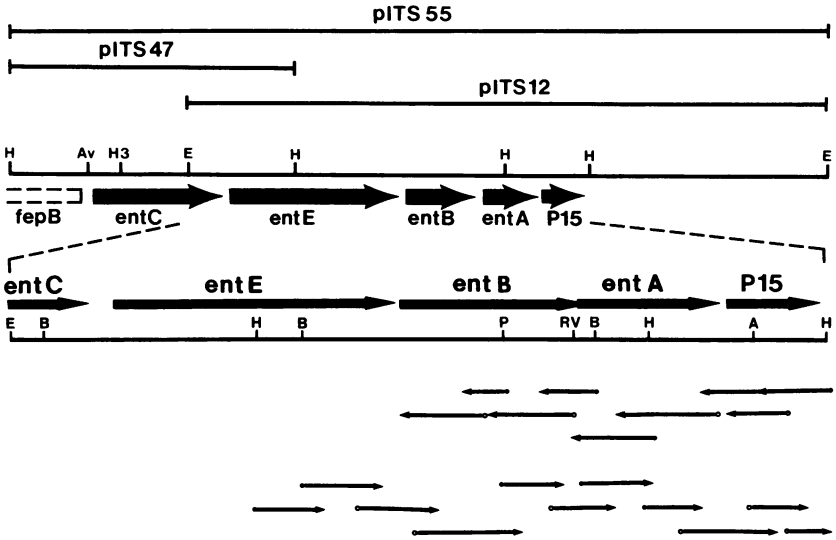

FIG. 1. DNA-sequencing strategy. pITS12 (16) and pITS47 (20) have been described elsewhere. pITS55 was constructed by ligation of the 7.0-kb *Eco*RI fragment of pITS12 to the 5.0-kb *Eco*RI fragment of pITS47. Sequence data were obtained from both DNA strands from a point 508 bp upstream of the *Pvu*II site in *entB* to the *Hpa*I site immediately downstream of the 15-kDa coding region. Additional sequence was determined on one strand for the 3′ end of *entE* from the *Hpa*I site, reading toward *entB*. Closed circles on arrows represent data obtained by using the M13 17-mer universal primer from designated restriction enzyme cleavage sites; open circles designate positions of synthetic oligonucleotide primers. The short arrows (bottom) represent the direction and approximate extent of sequence data from each primer site. Abbreviations: E, *Eco*RI; B, *Bst*EII; H, *Hpa*I; P, *Pvu*II; RV, *Eco*RV; A, *Acc*I.

JM101 [*supE thi* Δ(*lac-proAB*) (F′ *traD36 proAB lacI*^{q}*ZΔM15*)] (29). The β-galactosidase expression vector pMC1403 (4) was a generous gift of M. Casadaban. Constructions of the recombinant plasmids pITS12 (16), pITS47 (20), and pITS55 (19) were previously described. The extent of enterobactin-related DNA sequences in these plasmids is detailed in Fig. 1. Specific DNA fragments from these plasmids were isolated and cloned into the modified M13 phages M13mp18 and M13mp19 (18) for DNA sequence analysis.

Media and chemicals. Complete media were LB (Luria-Bertani) broth (15) for growth and maintenance of plasmid-containing strains and 2× TY broth (per liter: 16 g of tryptone [Difco Laboratories, Detroit, Mich.], 10 g of yeast extract, and 5 g of NaCl) for growth of JM101 used to propagate recombinant M13 phage. LB was supplemented with either 200 μM 2,2′-dipyridyl (Sigma Chemical Co., St. Louis, Mo.) or 20 μM $FeSO_4$ and 10 mM sodium citrate to produce the iron-deficient or iron-rich conditions, respectively, used for β-galactosidase assays. Ampicillin (Ap) was used at a concentration of 50 μg/ml for plasmid selection. *o*-Nitrophenyl-β-D-galactoside (ONPG) and isopropyl-β-D-galactoside (IPTG) were purchased from Sigma; 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal) was purchased from Boehringer Mannheim Biochemicals (Indianapolis, Ind.). [α-^{32}P]dATP (600 Ci/mmol) was purchased from ICN Radiochemicals (Irvine, Calif.).

β-Galactosidase assay. The method of Miller (15) was used to assay the β-galactosidase activity of chloroform-treated cells harboring various *ent-lacZ* fusions and grown in LB with low or high iron concentrations. The values presented (Table 1) represent the means of three independent β-galactosidase assays for each strain.

DNA manipulations. Restriction endonucleases and T4 DNA ligase were purchased from Promega Biotec, Inc. (Madison, Wis.) and used according to the recommendations of the manufacturer. Plasmids and the replicative form of M13 phage were isolated by the procedure of Birnboim and Doly (1). Purification was completed by precipitation of double-stranded DNA in the presence of 800 mM NaCl and 6.5% polyethylene glycol as recommended by Promega. $CaCl_2$-treated cells were transformed with plasmid DNA or transfected with phage DNA as described previously (5).

DNA sequencing. DNA fragments generated by using the strategic restriction endonuclease cleavage sites shown in Fig. 1 or random fragments generated by *Sau*3A or *Alu*I were cloned into the double-stranded replicative form DNA of M13mp18 and M13mp19 (18). Recombinants were transfected into JM101 (29) to produce the desired single-stranded templates. These were purified by the method outlined in the M13 cloning and sequencing handbook distributed by Amersham Corp. (Arlington Heights, Ill.) except that after phenol extraction, template DNA was extracted twice with phenol-chloroform (50:50) and once with chloroform before ethanol precipitation. Nucleotide sequences were determined unambiguously on both strands by the dideoxy-chain termination method of Sanger et al. (24), modified to use the Sequenase enzyme (United States Biochemicals, Cleveland, Ohio) as recommended by the supplier. Duplicate reactions, using dITP in place of dGTP in one, resolved compression areas and ensured that each base was read at least twice. All oligonucleotide primers for this analysis were synthesized by the University of Missouri DNA Core Facility (sponsored by the University of Missouri Molecular Biology Program) on a model 380A DNA synthesizer (Applied Biosystems, Inc., Foster City, Calif.). Reactions were analyzed on 0.4-mm 8% acrylamide sequencing gels, which were exposed to X-Omat RP5 film (Eastman Kodak Co., Rochester, N.Y.) at −70°C for visualization of bands. DNA sequences were examined by using the Microgenie sequence analysis program (23) distributed by Beckman Instruments, Inc. (Palo Alto, Calif.).

RESULTS

Nucleotide sequence of *entB*, *entA*, and an uncharacterized downstream gene encoding a 15-kDa protein. The previously described hybrid plasmid pITS12 (16) contains a 7.0-kilobase-pair (kb) *Eco*RI fragment which complemented the enterobactin mutant alleles *entA403*, *entB402*, *entC401*, *entE405*, and Δ*ent*(*AGB*)*415* and expressed four major polypeptides, estimated by gel electrophoresis at 58,000, 32,500, 26,000, and 15,000 Da. Analysis of the physical organization of this DNA fragment localized the coding regions for these four proteins to a span of approximately 4 kb, oriented them in the order just described, and identified them genetically as the products of *entE*, *entB*, *entAC*, and an unknown gene, respectively. We have designated this latter gene *P15* (and its product P15) for simplicity. Those studies failed to identify the product of *entG*, although they indicated that EntG activity is either related to the EntB protein or encoded by a small cistron localized to a 250-bp region between *entB* and *entA*. In addition, it was suggested on the basis of complementation experiments with strains harboring the *entC401* and *entA403* alleles that both EntC (isochorismate synthetase) and EntA (2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase) activities were related to the single 26,000-Da protein.

To address these specific uncertainties, we determined the nucleotide sequence of a 2.2-kb region from within *entE* to a site downstream of *P15*. Discrete DNA fragments specific for both strands of pITS12 (Fig. 1) were generated by restriction endonucleases *Acc*I, *Bst*EII, *Eco*RV, *Hpa*I, and *Pvu*II and cloned into bacteriophages M13mp18 and -mp19. Nucleotide sequences obtained from these fragments were overlapped with those from random fragments generated by *Alu*I and *Sau*3A. Sequencing gaps on either strand were completed by using synthetic oligonucleotide primers complementary to the sequences flanking the gap regions (Fig. 1). The DNA sequence from a point 510 bp upstream of the *Pvu*II site in *entB* to the *Hpa*I site immediately downstream of *P15* (a span of 2,137 bp) was determined unambiguously on both strands (Fig. 2).

Three contiguous open reading frames were identified from the primary nucleotide data. The most upstream coding region extended from base positions 76 to 930 and specified the 285-amino-acid EntB protein, with a calculated molecular weight of 32,558. Seven bases upstream of the predicted translation initiation codon was a sequence (5′-GGAGA) resembling the consensus ribosome-binding site (5′-GGAGG) (25). Immediately downstream, and overlapping the EntB termination codon at base position 933, was an open reading frame encoding a 248-residue polypeptide with a deduced molecular weight of 26,253. This protein (designated EntA in Fig. 2 for reasons discussed below) corresponded to the previously identified EntAC protein (16) and was also preceded by a potential ribosome-binding sequence (5′-GAGG). The downstream open reading frame spanned from bases 1682 to 2092 and encoded a 137-amino-acid protein with a calculated size of 14,972 Da. Seven bases upstream of its translation initiation codon was a purine-rich sequence that may bind ribosomes (11). This protein (P15) has no known function in enterobactin biosynthesis or transport. An insertion mutation, constructed in vitro by ligating a kanamycin resistance gene into the unique *Cla*I site in *P15* followed by homologous recombination into the chromosome, displayed no obvious phenotypes with respect to the enterobactin system (T. J. Brickman, unpublished data). The deduced molecular weights for these three protein sequences closely resembled those estimated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (16, 22). Examination of this nucleotide sequence in all three translational reading frames did not identify a separate coding sequence for an independently translated EntG protein. Since the EntB- and EntA-coding regions were directly linked, there was no intervening region to encode EntG, and no open reading frame that overlapped the EntB-EntA junction sequences was obvious. This observation will be discussed below in reference to previous reports of a separate *entG* cistron.

Protein similarities between EntA and a family of dehydrogenases. The 26-kDa protein sequence was used for homology comparisons with sequences obtained from the National Biomedical Research Foundation protein data bank. This search revealed a region of significant similarity at the amino terminus of EntA (Fig. 3) to members of an extended family of short alcohol-polyol-sugar dehydrogenases, including *Drosophila* alcohol dehydrogenase (26), *Bacillus megaterium* glucose dehydrogenase (9), and *Klebsiella aerogenes* ribitol dehydrogenase (10). Critical to these similarities is the strategic conservation and spacing of glycine residues predicted to function in coenzyme-binding domains (10) of the initial two β-strands (designated βA and βB) at the amino ends of such pyridine nucleotide-linked enzymes. The similarities detected would predict that the 26,000-Da protein functions as the NADP-dependent dehydrogenase responsible for the conversion of 2,3-dihydro-2,3-dihydroxybenzoate to 2,3-dihydroxybenzoate (i.e., EntA). This prediction is corroborated by the previously described genetic complementation of the *entA403* lesion (16) and by the in vitro biochemical conversion of the substrate to the expected product in the presence of purified 26-kDa protein (C. T. Walsh, personal communication).

Whereas these data implicate the 26-kDa protein in EntA enzymatic activity, genetic evidence for additional involvement of this protein in EntC function has been established by the isolation of two independent Tn*5* insertion mutations, pMS142 (13) and T6 (16), and mini-*kan* insertion mutation M3 (16), all of which mapped in the 26-kDa protein gene and yet failed to complement both the *entA403* and *entC401* lesions. Data bank analysis, however, did not reveal extensive homologies at the amino acid level between the 26-kDa protein and the other chorismate-utilizing enzymes, TrpE and PabB, predicted from earlier reports (7, 17) if this enzyme were EntC. These similarities were apparent in the 44-kDa protein encoded by the gene located upstream of *entE*, as detailed in the accompanying report (19). That study identified the 44-kDa protein as EntC, and as a result, the 26-kDa protein is redefined as EntA. However, it should be mentioned that a small region of amino acid homology of unknown significance was identified between EntA, TrpE, and PabB in the data bank search (data not shown). Whether this region of the EntA protein is involved, along with the 44-kDa protein, in the complete EntC enzymatic activity awaits molecular delineation of *entA* lesions (e.g., *entC401*) affecting EntC activity.

Iron-mediated transcriptional linkage between *ent* biosynthetic genes. The lack of intercistronic nucleotide regions containing canonical promoter elements separating the three open reading frames which define *entB*, *entA*, and *P15* (Fig. 2) indicates that these genes are controlled in vivo by a regulatory region located upstream of *entE*. The fact that chromosomal *lacZ* fusions to *entC*, *entE*, and *entA* are all iron regulated to similar extents (6) further suggests that the iron-responsive control elements for these genes are located

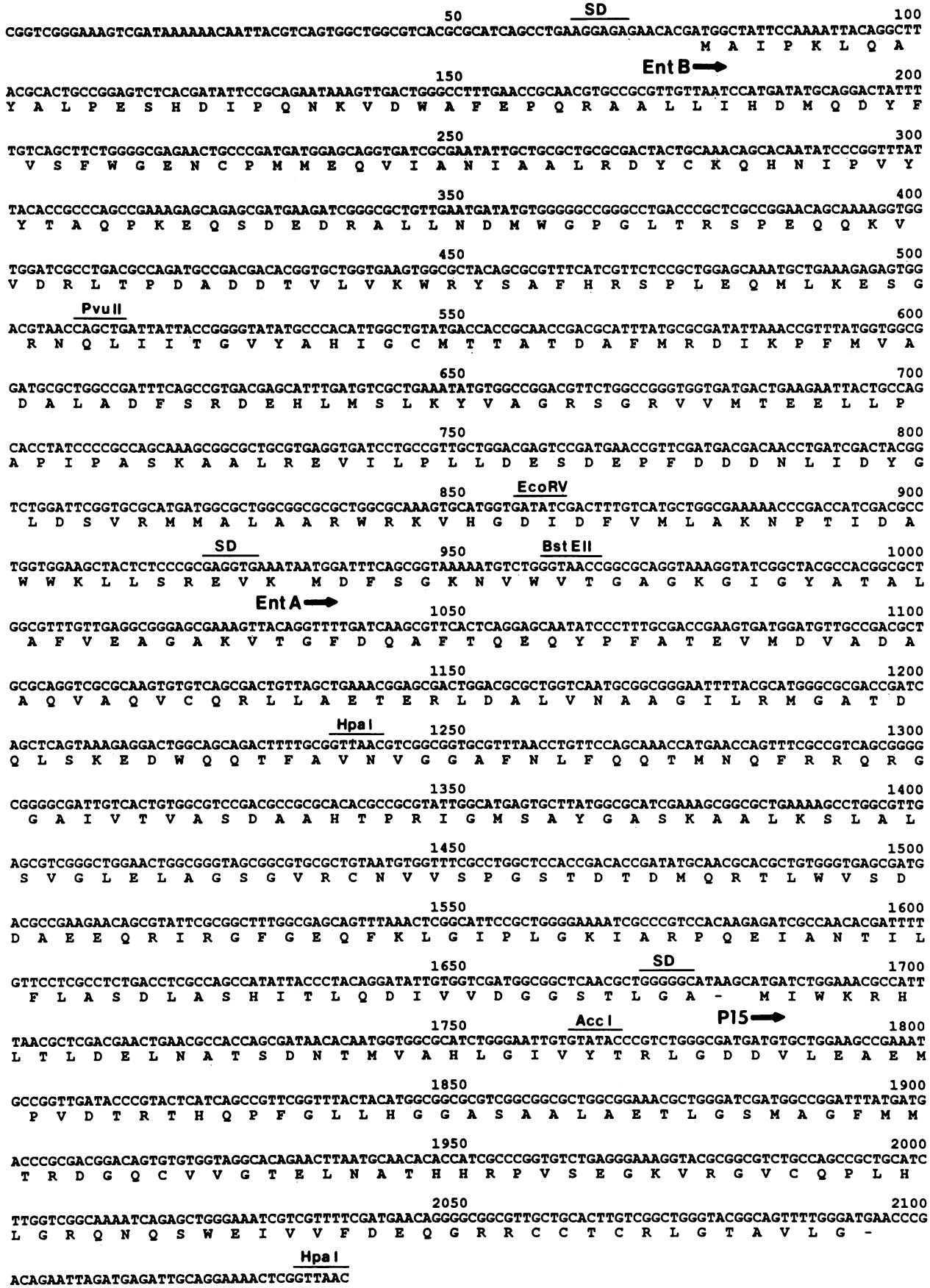

FIG. 2. Nucleotide sequence of *entB*, *entA*, and a tandem gene (*P15*) encoding the 15-kDa protein. Relevant restriction sites are designated. Open reading frames are identified by one-letter amino acid codes. Potential ribosome-binding sites (SD) upstream of three open reading frames are indicated.

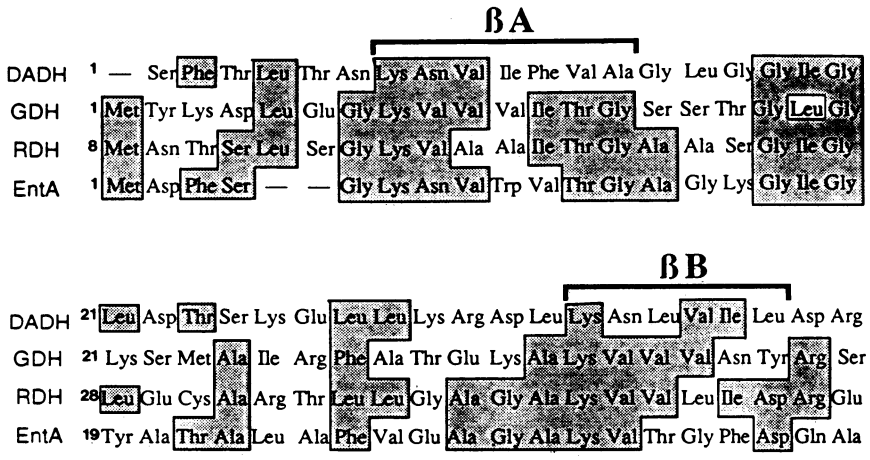

FIG. 3. Amino acid alignments between EntA and pyridine nucleotide-dependent dehydrogenases. Homologous segments of EntA, *Drosophila* alcohol dehydrogenase (DADH), the *B. megaterium* glucose dehydrogenase (GDH), and ribitol dehydrogenase (RDH) of *K. aerogenes* are aligned. Numbers after enzyme designations indicate amino acid positions in the respective proteins. Regions of similarity are enclosed in shaded boxes. The initial two β strands (βA and βB) characteristic of cofactor-binding domains (10) are identified.

between the divergently transcribed *fepB* and *entC* loci (Fig. 1). To test this hypothesis, a series of Ent-LacZ protein fusions was constructed so that expression of each hybrid protein was controlled only by its immediate upstream region. Restriction sites in each *ent* coding sequence were chosen so that insertion of that DNA fragment into the expression vector pMC1403 (4) created in-frame translational fusions with *lacZ* (Fig. 4). These plasmid constructs, as well as a previously described iron-regulated *entF-lacZ* fusion in pMC1403 (20), were transformed into MC4160 (Δ*lac*), and β-galactosidase activity was monitored relative to iron concentrations in the growth medium (Table 1). Only the *entC-lacZ* fusion gene in pITS305 was iron regulated similarly to the *entF-lacZ* fusion control; the remaining plasmids expressed low-level or no enzyme activity regardless of the availability of iron. These data, in the context of

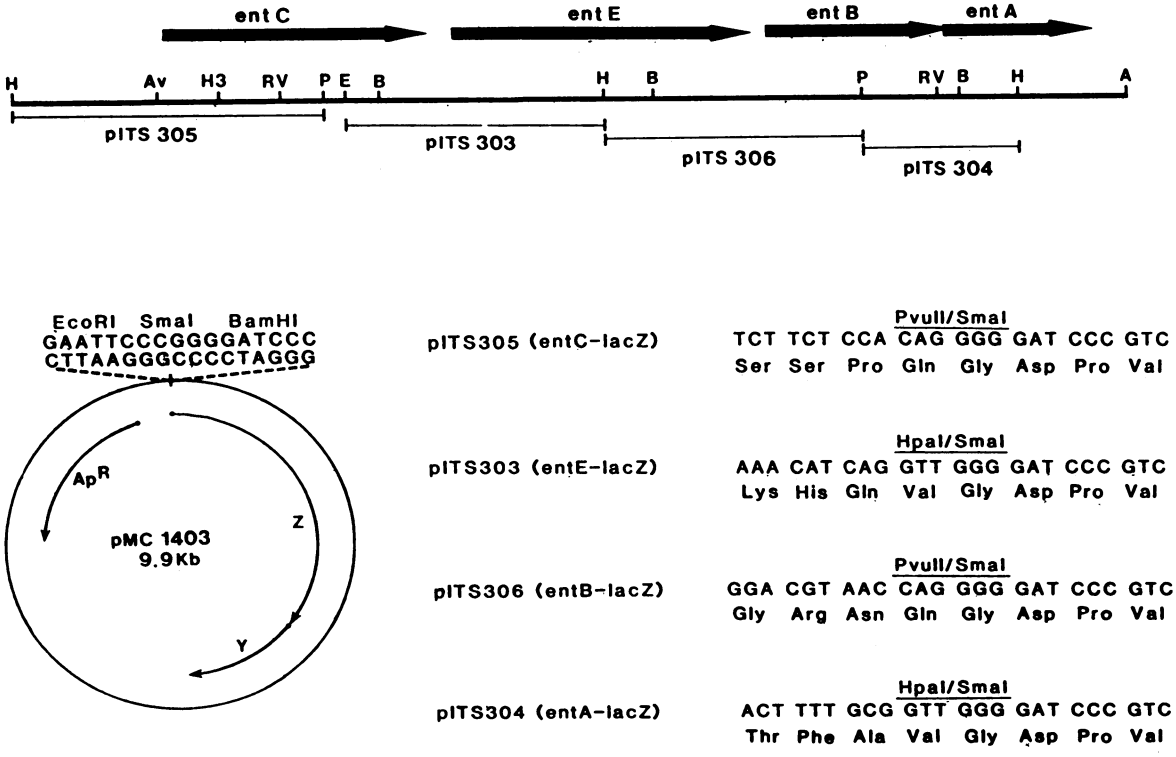

FIG. 4. Construction of plasmids containing *ent-lacZ* fusions. Designated restriction fragments were ligated into pMC1403 digested with *Eco*RI-*Sma*I (pITS303) or *Sma*I (pITS304, -305, and -306) to produce in-phase protein fusions between Ent enzymes and β-galactosidase. Cloned fragment orientation was verified by restriction enzyme mapping, and fusion junctions were determined by DNA sequence analysis. Abbreviations: A, *Acc*I; Av, *Ava*I; B, *Bam*HI; E, *Eco*RI; H, *Hpa*I; H3, *Hin*dIII; P, *Pvu*II; RV, *Eco*RV; Y, *lacY*; Z, *lacZ*.

TABLE 1. β-Galactosidase activity of *lacZ* fusion plasmids in response to iron concentration

Fusion plasmid	β-Galactosidase activity (U)	
	High Fe	Low Fe
pITS305 (*entC-lacZ*)	4	1,000
pITS303 (*entE-lacZ*)	43	51
pITS306 (*entB-lacZ*)	53	51
pITS304 (*entA-lacZ*)	0	0
pITS312 (*entF-lacZ*)	227	1,579

the previous iron-controlled chromosomal fusions and the nucleotide sequence analysis, provide strong evidence that *entC*, *entE*, *entB*, *entA*, and *P15* form an operon controlled in vivo by an iron-responsive operator-promoter region located upstream of the *entC* gene.

DISCUSSION

Attempts to delineate the structural organization of the group of *ent* genes clustered within a 4-kb region on the clockwise end of the large complex of enterobactin biosynthesis and transport loci have relied on genetic and biochemical complementation studies using a limited collection of *ent* mutant strains. These investigations have produced confusing and inconsistent reports concerning the gene order and transcriptional linkage of the *ent* genes. In an effort to determine the precise genetic configuration of these genes, we established the complete nucleotide sequence of a 2,137-bp DNA fragment extending from an undefined region in *entE* (75 bp upstream of *entB*) to a position 45 bases downstream of *P15* and containing the complete *entB*, *entA*, and *P15* genes. The deduced amino acid sequences of these three genes represent protein products whose calculated sizes agree well with previous estimates (16, 22). The identified cistrons are tightly spaced, with no extensive intergenic regions containing consensus transcriptional initiation signals, which suggests cotranscriptional expression of these three genes from a region upstream of the defined sequence.

The most obvious contradiction established by the nucleotide sequence of this region is the lack of an independent protein-coding sequence corresponding to EntG. Although previous studies had failed to detect a polypeptide associated with this function (16), the deletion which defines *entG* was clearly localized to a genetic region (13, 16) now known to encode the carboxyl end of EntB and the amino end of EntA. It was observed in one of those studies (16) that full expression of $EntG^+$ activity in the *ent(ABG)415* deletion strain required that the entire EntB protein be encoded by the complementing plasmid. Weak expression was detected when the carboxyl end of EntB was produced as a hybrid protein, fused in frame to the amino termini of various proteins during the plasmid constructions, with the exception of one hybrid plasmid from which no obvious protein fusion was produced. These observations led to the alternative considerations that the EntB protein also provided EntG function or that *entG* was transcriptionally linked to *entB* and encoded a small, and perhaps unstable, protein product with an undefined enzymatic function (16). The latter premise is clearly eliminated by the current study, which indicates that *entG* does not exist as a separate genetic unit.

The former concept suggests the exciting possibility that the $EntG^+$ function (likely to reside on the EntB protein) represents a level of communication between the enzymatic machinery which synthesizes the catechol moiety and the multienzyme enterobactin synthetase complex which produces the siderophore end product. For example, the absence of the protein carrying this function may prevent the synthetase complex from forming efficiently, consistent with the observation that the $EntG^-$ strain AN462 also produces very low levels of EntD, EntE, and EntF activities (27). Direct physical interaction between the protein expressing EntG activity and the synthetase complex has been firmly established by gel filtration and ion-exchange chromatography (3, 8, 28). Overproduction and purification of the now molecularly defined EntB protein can be pursued to evaluate the ability of this protein to directly associate with components of enterobactin synthetase.

Although data bank searches did not identify any proteins sharing structural homologies with EntB or P15, significant similarities were established between the 26-kDa EntA protein and an extended family of short-chain alcohol-polyol-sugar dehydrogenases, primarily in regions established as cofactor-binding domains (10). This observation is consistent with previous biochemical evidence (30) that EntA contains NADP-dependent dehydrogenase activity. In a recent genetic study (16), the *entA403* and *entC401* mutations were clearly mapped to a single genetic locus encoding this 26-kDa protein, and independent insertion mutations from that and a previous study (13) failed to complement both lesions. These observations suggested (16) that this polypeptide was required for both isochorismate synthetase and 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase activities. Moreover, in light of the previous report that these two enzymatic activities could be biochemically separated by chromatography on DEAE-cellulose (30), it was proposed that both the genetic and biochemical data are consistent with the concept, among others, that either activity is the result of association with another cellular factor or protein. The accompanying study (19) firmly establishes by mutational analysis and amino acid matrix comparisons that the 44-kDa protein encoded by the genetic region upstream of *entE* is EntC, but it does not rule out the possibility, indicated by the genetic evidence, that an additional component residing on the EntA molecule is required for enzymatic activity. Confirmation would require establishment of a requirement for both proteins during the in vitro conversion of chorismate to isochorismate along with corroborating molecular characterization of *entA* alleles defective in EntC activity.

The transcriptional organization of this *ent* gene cluster has not been elucidated by the previous investigations. Data from Mu d(Ap^r *lac*) operon fusions (6) suggested that these genes form an iron-inducible operon transcribed in a clockwise direction on the *E. coli* chromosome, although the gene order derived from that investigation is not consistent with current molecular analyses. In contrast, the expression of individual *ent* loci from this region on multicopy plasmids (12, 13, 16, 22) and the apparent lack of polarity effects as a result of insertional mutagenesis in these loci (13, 16) led to the conclusions that the *ent* genes from this region can be independently transcribed. These conclusions are valid when it is considered that expression of these plasmid-borne genes in mutant strains allowed growth on iron-deficient media or production of detectable enterobactin levels. However, the iron inducibility of this expression or the possibility that expression from nonspecific vector promoters could produce levels of enterobactin enzymes sufficient to account for the apparent complementation was not thoroughly examined in those studies. The nucleotide sequence data derived herein argue strongly that independent promoter elements

are not present upstream of the *entB*, *entA*, and *P15* cistrons. This hypothesis was confirmed by constructing *lacZ* fusions in these cistrons, as well as *entE* and *entC*, but under the control of the immediate upstream region only. These plasmids established that only the *entC-lacZ* hybrid is iron regulated. In light of previous evidence that chromosomal *entA-lacZ* and *entE-lacZ* fusions are controlled by iron availability (6), these data indicate that the *entCEBA* (*P15*) cluster forms a single, iron-inducible operon, controlled from the genetic region between *fepB* and *entC*. The structure of this region is described in the accompanying report (19), and details of the molecular control elements will be published elsewhere (T. J. Brickman, B. A. Ozenberger, and M. A. McIntosh, manuscript in preparation).

ACKNOWLEDGMENTS

We thank G. S. Pettis and S. K. Armstrong for helpful discussions during the course of this study and for critical evaluation of the manuscript. We are grateful to C. T. Walsh for communication of unpublished results. We thank J. Neuner for preparation of the manuscript.

This work was supported by grant DMB8416017 (to M.A.M.) from the National Science Foundation and by Public Health Service predoctoral traineeship 1 T32 AI07276 (to T.J.B.) from the National Institute of Allergy and Infectious Diseases.

LITERATURE CITED

1. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
2. **Boyer, H. W., and D. Roulland-Dussoix.** 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. **41**:459–474.
3. **Bryce, G. F., and N. Brot.** 1972. Studies on the enzymatic synthesis of the cyclic trimer of 2,3-dihydroxy-*N*-benzoyl-L-serine in *Escherichia coli*. Biochemistry. **11**:1708–1715.
4. **Casadaban, M. J., J. Chou, and S. N. Cohen.** 1980. In vitro gene fusions that join an enzymatically active β-galactosidase segment to amino-terminal fragments of exogenous proteins: *Escherichia coli* plasmid vectors for the detection and cloning of translational initiation signals. J. Bacteriol. **143**:971–980.
5. **Cohen, S. N., A. C. Y. Chang, and L. Hsu.** 1972. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of *Escherichia coli* R-factor DNA. Proc. Natl. Acad. Sci. USA **69**:2110–2114.
6. **Fleming, T. P., M. S. Nahlik, and M. A. McIntosh.** 1983. Regulation of enterobactin iron transport in *Escherichia coli*: characterization of *ent*:Mu d(Ap^r *lac*) operon fusions. J. Bacteriol. **156**:1171–1177.
7. **Goncharoff, P., and B. P. Nichols.** 1984. Nucleotide sequence of *Escherichia coli pabB* indicates a common evolutionary origin of *p*-aminobenzoate synthetase and anthranilate synthetase. J. Bacteriol. **159**:57–62.
8. **Greenwood, K. T., and R. K. J. Luke.** 1976. Studies on the enzymatic synthesis of enterochelin in *Escherichia coli* K-12: four polypeptides involved in the conversion of 2,3-dihydroxybenzoate to enterochelin. Biochim. Biophys. Acta **454**:285–297.
9. **Jany, K.-D., W. Ulmer, M. Froschle, and G. Pfleiderer.** 1984. Complete amino acid sequence of glucose dehydrogenases from *Bacillus megaterium*. FEBS Lett. **165**:6–10.
10. **Jornvall, H., H. von Bahr-Lindstrom, K.-D. Jany, W. Ulmer, and M. Froschle.** 1984. Extended superfamily of short alcohol-polyol-sugar dehydrogenases: structural similarities between glucose and ribitol dehydrogenases. FEBS Lett. **165**:190–196.
11. **Kozak, M.** 1983. Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. Microbiol. Rev. **47**: 1–45.
12. **Laird, A. J., D. W. Ribbons, G. C. Woodrow, and I. G. Young.** 1980. Bacteriophage Mu-mediated gene transposition and in vitro cloning of the enterochelin gene cluster of *Escherichia coli*. Gene **11**:347–357.
13. **Laird, A. J., and I. G. Young.** 1980. Tn*5* mutagenesis of the enterochelin gene cluster of *Escherichia coli*. Gene **11**:359–366.
14. **Luke, R. K. J., and F. Gibson.** 1971. Location of three genes concerned with the conversion of 2,3-dihydroxybenzoate into enterochelin in *Escherichia coli* K-12. J. Bacteriol. **107**:557–562.
15. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. **Nahlik, M. S., T. P. Fleming, and M. A. McIntosh.** 1987. Cluster of genes controlling synthesis and activation of 2,3-dihydroxybenzoic acid in production of enterobactin in *Escherichia coli*. J. Bacteriol. **169**:4163–4170.
17. **Nichols, B. P., M. van Cleemput, and C. Yanofsky.** 1981. Nucleotide sequence of *Escherichia coli trpE*. J. Mol. Biol. **146**:45–54.
18. **Norrander, J., T. Kempe, and J. Messing.** 1983. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene **26**:101–106.
19. **Ozenberger, B. A., T. J. Brickman, and M. A. McIntosh.** 1989. Nucleotide sequence of *Escherichia coli* isochorismate synthetase gene *entC* and evolutionary relationship of isochorismate synthetase and other chorismate-utilizing enzymes. J. Bacteriol. **171**:775–783.
20. **Ozenberger, B. A., M. S. Nahlik, and M. A. McIntosh.** 1987. Genetic organization of multiple *fep* genes encoding ferric enterobactin transport functions in *Escherichia coli*. J. Bacteriol. **169**:3638–3646.
21. **Pettis, G. S., and M. A. McIntosh.** 1987. Molecular characterization of the *Escherichia coli* enterobactin cistron *entF* and coupled expression of *entF* and the *fes* gene. J. Bacteriol. **169**:4154–4162.
22. **Pickett, C. L., L. Hayes, and C. F. Earhart.** 1984. Molecular cloning of the *Escherichia coli* K-12 *entACGBE* genes. FEMS Microbiol. Lett. **24**:77–80.
23. **Queen, C., and L. J. Korn.** 1984. A comprehensive sequence analysis program for the IBM personal computer. Nucleic Acids Res. **12**:581–599.
24. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
25. **Shine, J., and L. Dalgarno.** 1974. The 3′-terminal sequence of *Escherichia coli* 16S RNA: complementarity to nonsense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA **71**:1342–1346.
26. **Thatcher, D. R., and L. Sawyer.** 1980. Secondary-structure prediction from the sequence of *Drosophila melanogaster* (fruitfly) alcohol dehydrogenase. Biochem. J. **187**:884–886.
27. **Woodrow, G. C., I. G. Young, and F. Gibson.** 1975. Mu-induced polarity in the *Escherichia coli* K-12 *ent* gene cluster: evidence for a gene (*entG*) involved in the biosynthesis of enterochelin. J. Bacteriol. **124**:1–6.
28. **Woodrow, G. C., I. G. Young, and F. Gibson.** 1979. Biosynthesis of enterochelin in *Escherichia coli* K-12: separation of the polypeptides coded for by the *entD*, *E*, *F*, and *G* genes. Biochim. Biophys. Acta **582**:145–153.
29. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.
30. **Young, I. G., and F. Gibson.** 1969. Regulation of the enzymes involved in the biosynthesis of 2,3-dihydroxybenzoic acid in *Aerobacter aerogenes* and *Escherichia coli*. Biochim. Biophys. Acta **177**:401–411.
31. **Young, I. G., L. Langman, R. K. J. Luke, and F. Gibson.** 1971. Biosynthesis of the iron transport compound enterochelin: mutants of *Escherichia coli* unable to synthesize 2,3-dihydroxybenzoate. J. Bacteriol. **106**:51–57.