JOURNAL OF BACTERIOLOGY, Sept. 1987, p. 4163–4170
0021-9193/87/094163-08$02.00/0

Vol. 169, No. 9

Cluster of Genes Controlling Synthesis and Activation of 2,3-Dihydroxybenzoic Acid in Production of Enterobactin in *Escherichia coli*

MARY SCHRODT NAHLIK, TIMOTHY P. FLEMING,† AND MARK A. McINTOSH*

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

Received 6 April 1987/Accepted 19 June 1987

The *Escherichia coli* gene cluster encoding enzymatic activities responsible for the synthesis and activation of 2,3-dihydroxybenzoic acid in the formation of the catechol siderophore enterobactin was localized to a 4.2-kilobase chromosomal DNA fragment. Analysis of various subclones and transposon insertion mutations confirmed the previously suggested gene order as *entEBG(AC)* and provided evidence to suggest that these genes are organized as three independent transcriptional units, composed of *entE*, *entBG*, and *entAC*, with the *entBG* mRNA transcribed in a clockwise direction. Plasmid-specific protein expression in *E. coli* minicells identified EntE and EntB as 58,000- and 32,500-dalton proteins, respectively, while no protein corresponding to EntG was detected. The EntA and EntC enzymatic activities could not be separated by genetic or molecular studies. A small DNA fragment encoding both activities expressed a single 26,000-dalton polypeptide, suggesting that this protein is a multifunctional enzyme catalyzing two nonsequential reactions in the biosynthetic pathway. A protein of approximately 15,000 daltons appears to be encoded by the chromosomal region adjacent to the *entAC* gene, but no known function in enterobactin biosynthesis or transport can yet be ascribed to this polypeptide.

Iron is required as an essential trace element for bacterial growth. In an external environment containing oxygen, however, iron exists in an insoluble form as a ferric hydroxide polymer, while in a mammalian host, iron is complexed with transferrin and lactoferrin in serum and body secretions, respectively. To overcome this nutritional dilemma and assimilate necessary metabolic iron stores, bacteria synthesize and secrete iron-chelating compounds called siderophores, which bind iron and are taken up into the cell by cognate membrane transport proteins. The iron is then released from the siderophore complex to be used for cellular metabolism by the bacteria.

In gram-negative enteric organisms, including *Escherichia coli* and *Salmonella typhimurium*, the native siderophore is enterobactin (enterochelin), a cyclic trimer of 2,3-dihydroxybenzoylserine (27, 32, 33). The enterobactin system in *E. coli* consists of at least seven biosynthetic genes and five transport genes, all of which map near 13 min on the *E. coli* chromosome (1). The enzymes EntC (isochorismate synthetase), EntB (2,3-dihydro-2,3-dihydroxybenzoate synthetase), and EntA (2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase) catalyze the conversion of chorismate, an intermediate of aromatic amino acid synthesis, to 2,3-dihydroxybenzoic acid through the intermediates isochorismate and 2,3-dihydro-2,3-dihydroxybenzoic acid (41). L-Serine and 2,3-dihydroxybenzoic acid are subsequently activated and utilized to produce enterobactin by an uncharacterized process that requires the polypeptides EntD, EntE, EntF, and EntG. It has been proposed that these proteins form a multienzyme complex, enterochelin synthetase (15), although the specific function of each enzyme has not been established. EntE is proposed to catalyze an ATP-PP_i exchange reaction that results in the activation of 2,3-dihydroxybenzoic acid, and EntF appears to activate serine by a similar mechanism. EntD is physically associated with EntF and EntG and may attach this enzyme complex to the cell membrane. EntG function has not been characterized.

Transport of the ferric enterobactin complex involves the outer membrane receptor protein FepA (16, 17, 39) and an energized membrane state associated with the TonB protein (14). Recent evidence implicates a periplasmic binding protein, FepB, and inner membrane transport functions, FepC (31), FepD, FepE, and perhaps FepF (28), in the uptake process. Once internalized, iron is reductively removed from the ligand complex by the *fes* gene product (17, 22).

The enterobactin genes span approximately 24 kilobase pairs (kb) of the *E. coli* chromosome, and a variety of molecular strategies were used to isolate and partially characterize their genetic and spatial organization (7, 11, 12, 20, 21, 28–31). Physical analyses of isolated *ent* gene fragments coupled with genetic complementation studies indicated that the genes are organized in the following order (clockwise on the *E. coli* chromosome): *entD*, *fepA*, *fes*, *entF*, *fepE*, *fepC*, *fepD*, *fepB*, *fepF*, *entE*, *entB*, *entG*, and *ent(AC)*. The translational products of *fepA*, *fes*, *entF*, *fepE*, *fepC*, *fepB*, *fepF*, and *entE* were identified by examining protein expression in *E. coli* minicells (7, 12, 28, 30).

A molecular analysis of the five *ent* biosynthetic activities encoded by genes which map to the right of *fepB* was initiated with the isolation of recombinant plasmids containing these genes from *E. coli* chromosomal libraries (M. S. Nahlik and M. A. McIntosh, Abstr. Annu. Meet. Am. Soc. Microbiol. 1986, K137, p. 216). Each of these gene products (except perhaps EntG) is directly involved in the biosynthesis of 2,3-dihydroxybenzoic acid or its activation in the subsequent production of enterobactin. This region has also been isolated on a hybrid λ bacteriophage, and BAL 31-generated deletions suggested that the order of transcription is counterclockwise (30). Five proteins were produced in

* Corresponding author.
† Present address: National Cancer Institute, Bethesda, MD 20205.

TABLE 1. Bacterial strains

Strain	Genotype	Reference or source
AN90	*thi leuB proC trpE lacY mtl xyl rpsL azi fhuA tsx supA entD*	9
AN93	Like AN90 but *entE*	I. G. Young
AN191	Like AN90 but *entC*	I. G. Young
AN192	Like AN90 but *entB*	I. G. Young
AN193	Like AN90 but *entA*	I. G. Young
AN93-60	Δ*recA* derivative of AN93	This study
AN191-60	Δ*recA* derivative of AN191	This study
AN192-60	Δ*recA* derivative of AN192	This study
AN193-60	Δ*recA* derivative of AN193	This study
AN462	*pabA his-4 arg-3 ilv-7 aroE purE rpsL rpsE ΔentABG415* (Mu$^+$)	38
AN489	*pabA his-4 arg-3 ilv-7 aroE purE rpsL ΔentABG415* (Mu$^+$) *pheA363 pheO352 tyrA*	38
χ984	*minA minB pdxC purE his rpsL xyl ilv met*	13
MC4100	F$^-$*araD139* Δ(*lacIOPZYA*)*U169 rpsL thiA*	5
MC4160	Δ*recA* derivative of MC4100	This study
DB89-301[a]	*purE*::Tn*5* Δ*trpE*	D. Berg
JC10284	*srlR*::Tn*10 srlC srlD metB mtl gatC gatA malA xyl rpsL sup* Δ(*srlR-recA*)	10

[a] The Tn*5* insertion into *purE* has been previously described (35).

minicells and maxicells, and one of these (63,000 daltons) was identified as EntE.

This report corroborates and extends the previous studies. We isolated an 11.6-kb *Hin*dIII fragment and a 7-kb *Eco*RI fragment which contain the *entACGBE* genes. Complementation studies using subclones and insertion mutations constructed from the original fragments delineated the location of each of these genes on the 7-kb *Eco*RI fragment, and protein products were identified for *entE*, *entB*, and *entAC* through the use of these subclones in minicell analysis.

MATERIALS AND METHODS

Organisms, media, and enzymes. Table 1 lists all the *E. coli* strains used in this study. The plasmids pBR328 (36), encoding resistance to ampicillin, tetracycline, and chloramphenicol, pACYC184 (6), encoding resistance to tetracycline and chloramphenicol, and pUC18 (26), encoding ampicillin resistance and β-galactosidase, were used as cloning vectors. The recombinant bacteriophage λNK1105 (37) was used to deliver the kanamycin resistance transposable element (p*tac* + mini-*kan*) as described below. P1 transduction experiments to transfer relevant genetic markers between strains were performed by the method of Miller (25). *recA* deletion derivatives of each strain were constructed by isolating P1 transductants which carried the *srl*::Tn*10* and Δ*recA* markers from JC10284. Spontaneous loss of Tn*10* was selected (4), and retention of the *recA* deletion was assayed by sensitivity to UV irradiation. Luria broth, M9 minimal medium (25), and Tris-succinate iron-limiting medium (24) have been previously described. Supplements were added to minimal media at the following concentrations: amino acids and adenine, 50 μg/ml; thiamine, 10 μg/ml; glucose, 4 g/liter; and succinate, 30 mM. Ampicillin was used at 25 μg/ml, chloramphenicol was at 30 μg/ml, tetracycline was at 10 μg/ml, and kanamycin was at 50 μg/ml. Restriction nuclease enzymes and T4 DNA ligase were purchased from Promega Biotech, Madison, Wis., Bethesda Research Laboratories, Gaithersburg, Md., International Biotechnologies, Inc., New Haven, Conn., and New England BioLabs, Inc., Beverly, Mass.

Construction of the *E. coli* chromosomal libraries. *E. coli* MC4100 was grown in Luria broth, and the chromosomal DNA was isolated by the procedure of Schleif and Wensink (34). Chromosomal DNA was completely digested with either *Eco*RI or *Hin*dIII. The resulting DNA fragments were electrophoresed on a 0.7% low-melting-point agarose gel, and since it had been previously shown (20) that the 2,3-dihydroxybenzoic acid-related genes were located on a 6.6-kb *Eco*RI fragment and an 11.6-kb *Hin*dIII fragment, the region from 6 to 12 kb was size selected and inserted into *Eco*RI- or *Hin*dIII-cleaved, alkaline phosphatase-treated pBR328. Colonies were screened for insertional inactivation of either the Cm^r gene (*Eco*RI) or the Tc^r gene (*Hin*dIII). Colonies containing recombinant plasmids were pooled, and the hybrid plasmid DNA was isolated by the procedure of Birnboim and Doly (3). This pooled plasmid DNA formed the *E. coli* partial genomic banks.

Complementations. Recombinant plasmid libraries were used to transform bacterial strains containing chromosomal mutations in *entA*, *entC*, *entB*, *entE*, or Δ*entABG* by using $CaCl_2$-treated bacterial cells and a brief (2-min) heat pulse at 42°C (8). Complementation was monitored in bioassays (11) designed to detect the production of 2,3-dihydroxybenzoic acid or enterobactin with the indicator strains AN193 or AN90, respectively.

Restriction mapping and ligation. Restriction endonucleases and bacteriophage T4 DNA ligase were used according to the recommendations of the suppliers. Endonuclease reactions were terminated by heating at 68°C for 5 min. Restriction nuclease fragments were analyzed on horizontal 0.7 or 1.0% agarose gels in 0.04 M Tris-acetate–2 mM EDTA buffer (pH 8.2) (23).

Minicell isolation and protein electrophoresis. Minicells were isolated from strain χ984 and labeled with [^{35}S]methionine as previously described (12). The labeled minicell preparations were boiled for 5 min in sample buffer (2% sodium dodecyl sulfate, 5% 2-mercaptoethanol, 10% [vol/vol] glycerol, 62.5 mM Tris, pH 6.8) before separation on a 12% or a 10 to 18% acrylamide gradient with 3% urea gels according to the method of Laemmli (19). Proteins were stained with Coomassie blue, and labeled peptides were identified by autoradiography on Kodak XAR-2 X-ray film.

Mini-*kan* insertions. pITS23 (Fig. 1) was transformed into MC4160, and the resulting strain was infected with λNK1105 at a multiplicity of infection of less than 1. λNK1105 contains the Tn*903* Km^r gene flanked by the first 70 base pairs (bp) of IS*10*R and IS*10*L from Tn*10* (37). The Tn*10* transposase gene is present in λNK1105 to provide the necessary *trans*-active transposition signal. Movement of this mini-*kan* element into the pITS23 5.9-kb *Eco*RI-*Pvu*II DNA insert was selected by plating the cells on Luria broth plates containing ampicillin, tetracycline, and kanamycin. Plasmid DNA was isolated from these cells and used to transform MC4160 to eliminate any contaminating λ DNA which may have been carried during the first plasmid isolation. Plasmid DNA from these transformants was mapped to localize the site of the mini-*kan* element insertion mutation, and relevant plasmids containing the mini-*kan* element in the insert fragment were then transformed into *entA*, *entB*, *entC*, *entE*, and Δ*entABG* mutants for genetic complementation analysis.

Tn*5* insertions. pITS12 (Fig. 1) was used to transform strain DB89-301, which contains a Tn*5* in the chromosomal

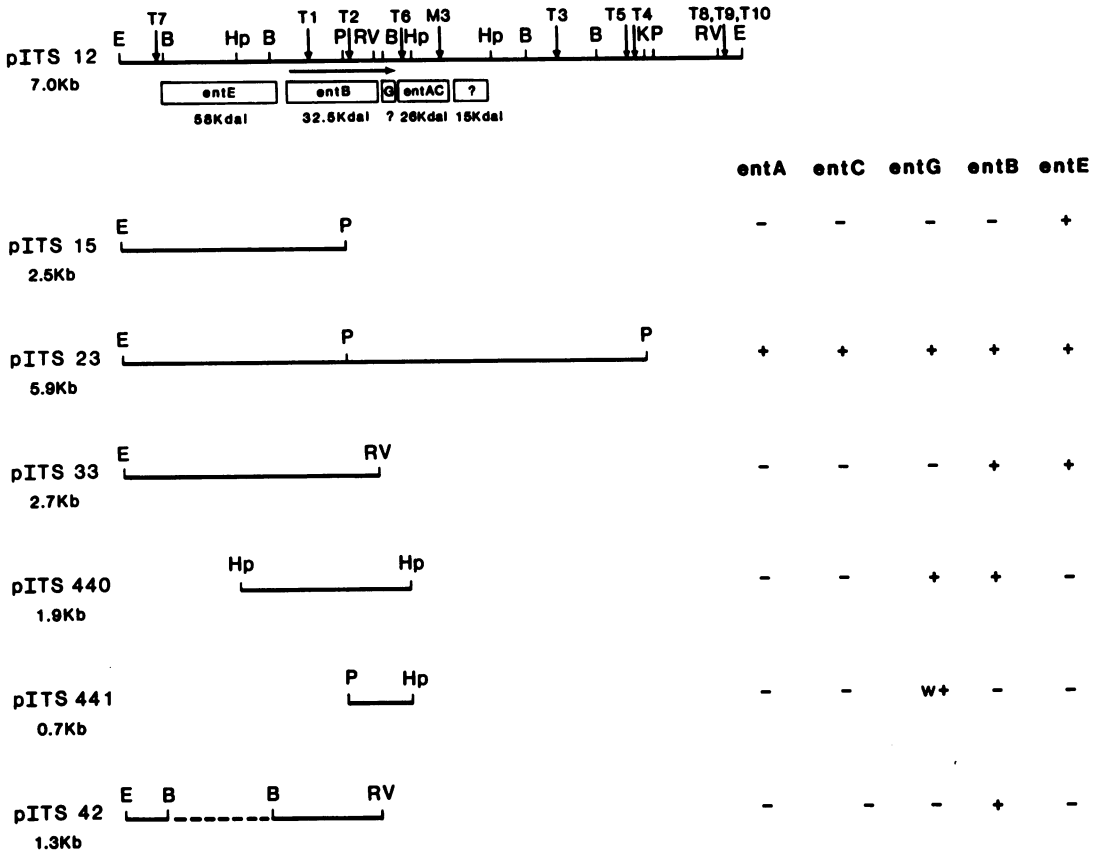

FIG. 1. Subclones defining *entB* and *entE*. pITS12, containing a 7-kb *Eco*RI fragment, is depicted at the top along with key restriction sites. There are no *Bam*HI, *Bal*I, *Sma*I, *Xho*I, or *Xba*I sites in pITS12. The location of Tn*5* (T) and mini-*kan* (M) insertion elements in pITS12 and pITS23 are indicated as arrows above pITS12. The *entE*, *entB*, *entG*, and *entAC* genes and their protein products are schematically illustrated below pITS12. pITS23, containing a 5.9-kb *Eco*RI-*Pvu*II fragment from pITS12, was used to generate a family of subclones to localize the gene boundaries for *entE* and *entB*. pITS15 contained the left 2.5-kb *Eco*RI-*Pvu*II fragment, and pITS33 contained the 2.75-kb *Eco*RI-*Eco*RV fragment, both cloned into the corresponding sites on pBR328. pITS440 contained a 1.4-kb *Hpa*I fragment, and pITS441 contained a 700-bp *Pvu*II-*Hpa*I fragment ligated into the *Sma*I site of pUC18. pITS42 is a derivative of pITS33, created by deletion of the internal 1.2-kb *Bst*EII fragment. The sizes (in kilobase pairs) under each recombinant plasmid refer to the length of the cloned DNA fragment and do not include the vector length. The depicted subclones were transformed into the appropriate enterobactin mutants, and complementation of the genes was determined by bioassay. B, *Bst*EII; E, *Eco*RI; RV, *Eco*RV; Hp, *Hpa*I; K, *Kpn*I; P, *Pvu*II; +, positive complementation; w+, weak complementation; −, lack of complementation; - - -, location of a deleted *Bst*EII fragment.

purE gene (35). Transformation mixtures were plated on Luria broth plates containing ampicillin and kanamycin. Isolated colonies were picked, grown to mid-log phase, and subsequently plated on Luria broth containing ampicillin and 250 μg of kanamycin per ml to select for multiple copies of Tn*5* (2). Plasmid DNA was isolated from the pooled high Km^r colonies and used to transform enterobactin mutants. The map locations of these insertions were determined by restriction enzyme analysis, and the insertional inactivation of the genes was determined by bioassay.

RESULTS

Isolation of the 2,3-dihydroxybenzoic acid gene cluster. AN193-60, containing the *entA403* mutation, was transformed with *Eco*RI and *Hin*dIII partial genomic libraries, and $EntA^+$ transformants were selected by growth on iron-deficient plates. The presence of *entA* on the plasmid was confirmed in a positive bioassay that results from the production of 2,3-dihydroxybenzoic acid by AN193-60 colonies containing plasmids with the complementing *entA* gene.

One recombinant plasmid was isolated from each library. pITS10 contained an 11.6-kb *Hin*dIII fragment, and pITS12 contained a 7.0-kb *Eco*RI fragment, which is a part of the 11.6-kb *Hin*dIII fragment. The DNA fragments incorporated into pITS10 and pITS12 are located at the right end of the 24 kb of DNA containing all characterized enterobactin genes. pITS1, which has been previously described (12), is at the left end of this 24 kb of DNA. pITS10 and pITS12 were transformed into *recA* derivatives of the *entA*, *entC*, *entB*, and *entE* mutant strains as well as the two *entG* deletion strains. Positive bioassays with these transformants revealed that both plasmids carried all five genes. The two cloned fragments containing the *entACGBE* gene cluster correspond to those contained in pMS112 and pMS111 (20) and pCP410 (30).

Several subclones and plasmids containing deletions of various sizes were constructed. Since pITS12 contained all of the 2,3-dihydroxybenzoic acid gene activities, most of the subcloning and genetic analysis was accomplished with this plasmid (Fig. 1). It is not known at this time if there are genes related to the ferric enterobactin system contained within the 4.1 kb of DNA of pITS10 located to the right of the 7-kb *Eco*RI fragment. pITS12 was digested to completion with *Eco*RI and then partially digested with *Pvu*II. A 5.9-kb DNA

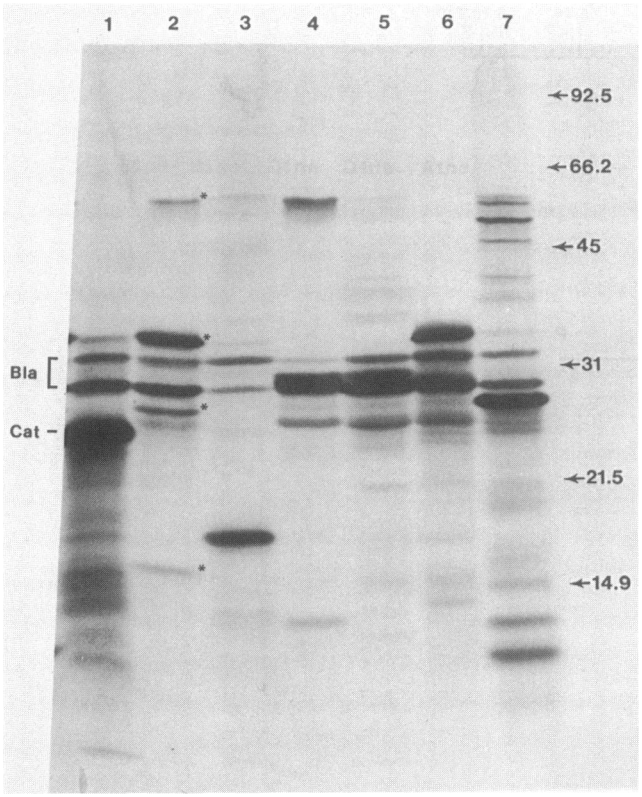

FIG. 2. Translational products of *entE*, *entB*, and *entAC*. Plasmid-specified translational products were analyzed from minicells containing the indicated plasmids. Minicell isolation and labeling were described in Materials and Methods, and proteins were separated on a 10 to 18% acrylamide gradient gel in the presence of 3% urea. Minicell products of the following plasmids are depicted: 1, pBR328; 2, pITS23; 3, pITS15; 4, pITS33; 5, pITS42; 6, pITS440; and 7, T2 insertion into pITS12. The β-lactamase (Bla) and chloramphenicol acetyltransferase (Cat) proteins are indicated on the left. Four easily detected, plasmid-specified proteins are indicated for pITS23 (lane 2) by asterisks (*), beginning with the 58,000-dalton protein and decreasing in size as follows: 32,500, 26,000, and 15,000 daltons. The molecular weight standards, indicated on the right, were lysozyme (14,900), soybean trypsin inhibitor (21,500), carbonic anhydrase (31,000), ovalbumin (45,000), bovine serum albumin (66,200), and phosphorylase *b* (92,500).

fragment was isolated and ligated into *Pvu*II-*Eco*RI-cut pBR328 to create pITS23 (Fig. 1). Complementation studies provided evidence that it expressed all five biosynthetic enzyme activities.

In addition to subcloning, Tn*5* and p*tac* + mini-*kan* insertional mutagenesis was used to define the locations of the enterobactin cistrons. Ten independent Tn*5* insertions and one mini-*kan* insertion were isolated. These insertions were mapped (Fig. 1), and the complementation groups were determined by transformation and bioassay. Insertion T7 is located on the extreme left end of pITS12, and its presence did not affect any of the five enzymatic activities. Insertions T3, T4, T5, T8, T9, and T10, located on the right, also did not inactivate any of the genes. Four other insertions, T1, T2, T6, and M3, inactivated various enterobactin cistrons and are discussed below.

Characterization of *entE* and *entB*. Subclones were constructed from pITS23 in order to isolate the individual gene loci (Fig. 1). pITS15 encoded only EntE activity and corresponds to pCP1492 of Pickett et al. (30). The subclone pITS33 contained an additional 250-bp *Pvu*II-*Eco*RV fragment extending to the right of the *Pvu*II site in pITS15 and resulted in the additional expression of EntB activity. An internal 1.2-kb *Bst*EII fragment was deleted from pITS33 to create the recombinant pITS42, which continued to express EntB activity but no longer synthesized EntE.

To define the protein products for these genes, the various subclones were transformed into the minicell-producing strain χ984 (Fig. 2). The 5.9-kb *Pvu*II-*Eco*RI fragment in pITS23 expressed three easily detectable protein products with molecular weights of approximately 58,000, 32,500, and 26,000 and a fourth, less obvious polypeptide near 15,000. A previous report (30) identified five proteins from the protein profile of the 7.0-kb *Eco*RI fragment inserted into the vector pACYC184, corresponding to the four peptides above and a 27,500-dalton protein, which may have been obscured in the pITS clones by the β-lactamase proteins. However, when the 7.0-kb *Eco*RI insert of pITS12 was cloned into pACYC184 in both orientations to form pITS212A and pITS212B (see below), only the four polypeptides identified in pITS23 and a weakly detectable protein that migrated slightly faster than the 26,000-dalton protein were common between these two clones. This additional protein was detected in all subsequent subclones regardless of the nature of their DNA inserts or locations relative to vector sequences, and its intensity varied from clone to clone. pITS15 complemented the *entE* mutation and expressed the 58,000-dalton polypeptide, which represents the previously described EntE (30), and an additional strong 20,000-dalton polypeptide. pITS33, which has both EntE and EntB activities, expressed the 58,000-dalton protein as well as a 29,000-dalton polypeptide, whereas its *Bst*EII deletion derivative pITS42, which still produced the 29,000-dalton protein and complemented *entB*, did not produce the 58,000-dalton protein and did not complement the *entE* mutation. These data suggest that the 20,000- and 29,000-M_r polypeptides are truncated forms of the mature 32,500-M_r EntB protein and that the 29,000-M_r polypeptide retained EntB activity even though it was truncated at the carboxy terminus. The complete 32,500-dalton EntB protein was expressed and active in pITS23 and pITS440, and its amino terminus must be located in the 0.5 kb between the *Bst*EII and *Pvu*II sites of pITS42. The cloning vector pBR328 and all recombinant derivatives which expressed Tc^r produced a weakly detectable protein product that migrated coincidentally with the 32,500-dalton EntB protein in this gel system. That this protein is the product of the *tet* gene was confirmed by pITS33, which was deleted for the *tet* promoter and amino terminus and did not express the comigrating polypeptide (Fig. 2, lane 4). The Tn*5* insertion mutations T1 and T2 confirmed that the strongly expressed 32,500-dalton protein represents the EntB product. T1 and T2 are located, respectively, 0.4 kb to the left and slightly to the right of the *Pvu*II site. These two insertions did not complement *entB* mutations, and correspondingly, the protein profiles of these two plasmids did not show the 32,500-M_r protein. T2 did, however, appear to encode a truncated derivative of EntB which did not have activity (Fig. 2). The presence of the truncated proteins in pITS15, pITS33, pITS42, and T2 suggests that *entB* is transcribed from left to right.

Characterization of *entG*. *entG* was originally identified as part of a Mu-induced deletion that also included *entA* and *entB* (38). The location of *entG* was defined between *ent*(*AC*) and *entB entE* by Laird and Young (21), who cloned DNA fragments surrounding Tn*5* insertion mutations. No characterized *entG* point mutation exists. Therefore, the deletion

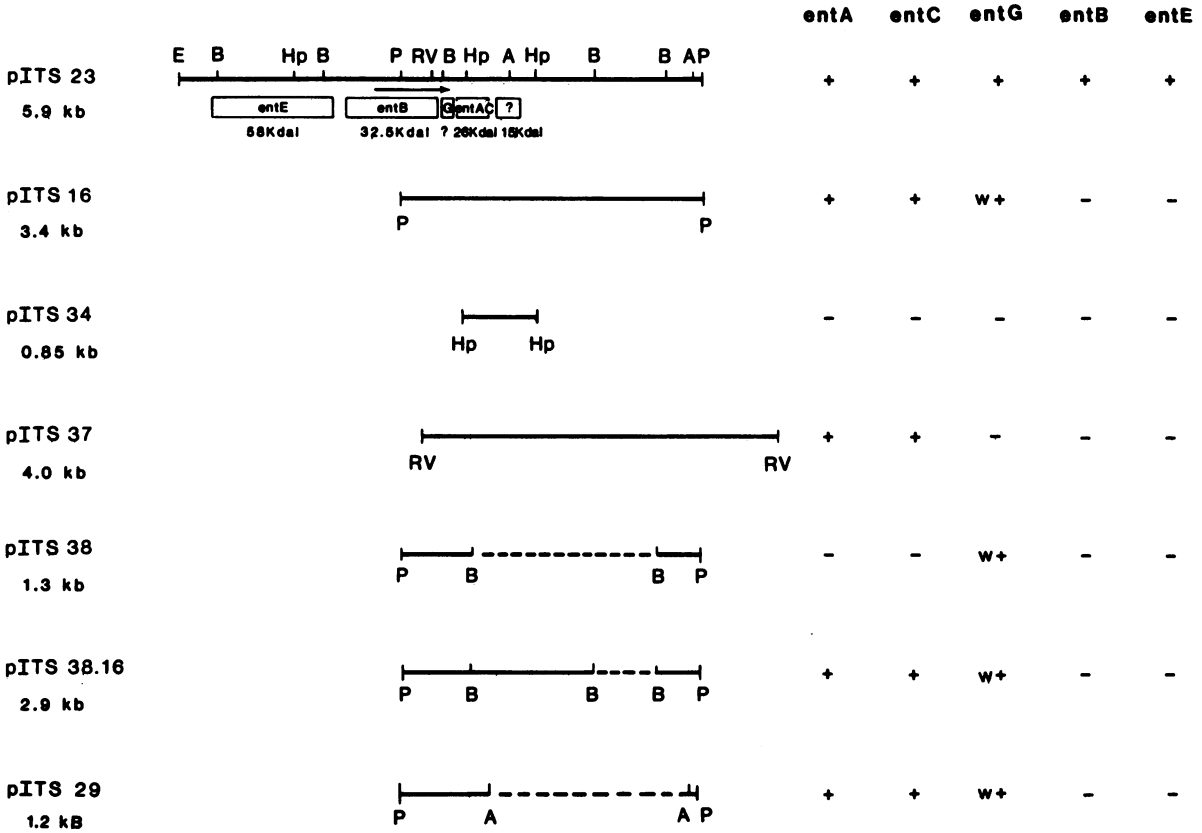

FIG. 3. Subclones defining *entAC*. The *entE*, *entB*, *entG*, and *entAC* genes and their protein products are schematically depicted below pITS23. The genes are drawn to scale, and the boundaries of each gene are represented relative to restriction sites as in Fig. 1. Individual restriction fragments from pITS23 were isolated and cloned into the appropriate restriction sites of pBR328 to produce the depicted subclones. The insert fragment sizes are given with each plasmid designation. The 0.85-kb *Hpa*I fragment was ligated into the *Eco*RV site of pBR328 to produce pITS34. The deletion plasmids pITS38 and pITS29 were constructed by digestion of pITS16 with *Bst*EII and *Acc*I, respectively, followed by ligation of the linear fragments. pITS38.16 was formed by ligation of the 1.6-kb *Bst*EII fragment into the single *Bst*EII site of pITS38. The dashed lines represent the extent of deleted DNA in these subclones. The depicted subclones were transformed into the appropriate enterobactin mutants, and complementation of the genes was determined by bioassay of the transformants. Restriction sites are indicated only if relevant to the indicated subclones. A, *Acc*I; B, *Bst*EII; E, *Eco*RI; RV, *Eco*RV; Hp, *Hpa*I; P, *Pvu*II; +, positive complementation; w+, weak complementation; −, lack of complementation; - - -, location of a deletion.

strains AN462 and AN489 (35) were used to assay EntG activity in the presence of 50 μM 2,3-dihydroxybenzoic acid to bypass the *entA* and *entB* lesions.

These strains were transformed with the previously described subclones and bioassayed on low-iron–2,3-dihydroxybenzoic acid medium by using AN90 as the indicator strain. The haloes of growth on the bioassay plate varied with each plasmid used in the transformation. Plasmids which complemented *entB* and contained a full-length EntB protein also complemented *entG* (Fig. 1 and 3). The subclones with fragments containing the left *Pvu*II site on one end, including pITS16, pITS29, pITS38.16, and pITS441, complemented *entG*, but the growth haloes were smaller and weaker than those produced by pITS23 and pITS440. This suggests that *entG* is located between the *Pvu*II site and the *Hpa*I site and that *entG* is not expressed fully in the absence of *entB*. This was supported by the T2 insertion located in *entB* which also inactivated *entG*. Poor but detectable expression of *ent* genes downstream on a polycistronic transcript has also been observed, in that low levels of *entF* expression can be measured in this bioassay in the absence of the strong promoter that controls the *fes-entF* transcript (29).

Protein profiles of the subclones shown in Fig. 1 were examined to define the *entG* protein product. pITS440, which complemented *entB* as well as *entG*, expressed only the EntB protein (Fig. 2). The 700-bp *Pvu*II-*Hpa*I fragment contains approximately 420 bp which code for the carboxy terminus of EntB. The coding region for the EntAC protein (see below) also extends into this fragment, leaving less than 230 bp to code for an EntG protein, which suggests that if *entG* encodes a protein, it would have a molecular weight of less than 9,000 and may not be resolved in our gel system.

Characterization of *entA* and *entC*. To define the *entA* and *entC* cistrons and characterize their protein products, subclones were constructed based on the restriction maps of pITS12 and pITS23. pITS16, which contained a 3.4-kb *Pvu*II fragment, and pITS37, which contained a 4.0-kb *Eco*RV fragment, both from pITS12, complemented *entA* and *entC* mutants (Fig. 3). The 0.85-kb *Hpa*I fragment isolated in pITS34, however, expressed neither activity, and pITS38, which was created as a result of the deletion of two internal *Bst*EII fragments from pITS16, gave similar results. When the 1.6-kb *Bst*EII fragment was ligated in the proper orientation into pITS38 at the single *Bst*EII site to create pITS38.16, both *entA* and *entC* activities were restored. These observations localized *entA* and *entC* to the 1.7-kb region between the leftmost *Eco*RV site in pITS37 and the second *Bst*EII site in pITS38.16. A smaller subclone

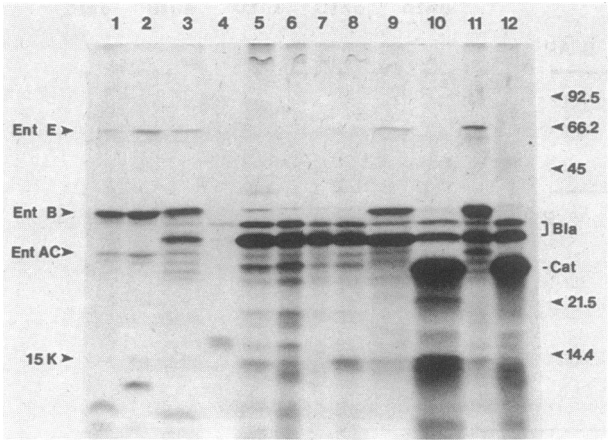

FIG. 4. Translational product of *entAC*. Minicells were transformed with the depicted subclones. Plasmid-specific proteins were synthesized in the presence of [^{35}S]methionine, separated on a 12% polyacrylamide gel, and visualized by autoradiography. Protein profiles of the following plasmids are depicted: 1, pITS212A; 2, pITS212B; 3, pITS12; 4, pITS15; 5, pITS16; 6, pITS29; 7, pITS38; 8, pITS38.16; 9, pITS23-M3; 10, pITS34; 11, pITS23; and 12, pBR328. pITS212A and pITS212B are derivatives of pACYC184 which contained the 7-kb *Eco*RI fragment from pITS12, cloned in opposite orientations. The four proteins of interest encoded by pITS212A, pITS212B, pITS12, and pITS23 are marked with arrows, and the protein names are indicated. Their apparent molecular weights are 58,000, 32,500, 26,000, and 15,000 from top to bottom. Migration of the molecular weight standards is indicated on the right (in kilodaltons), as is the location of the β-lactamase (Bla) and chloramphenicol acetyltransferase (Cat) proteins.

(pITS29) complementing both mutations was created by deleting the internal 2.3-kb *Acc*I fragment from pITS16, suggesting that both *entA* and *entC* are located within the 0.85-kb *Eco*RV-*Acc*I fragment. *entB* was shown to extend approximately 50 bp beyond the *Eco*RV site, and *entG* is also located in this fragment, leaving less than 800 bp to encode both EntA and EntC. This amount of DNA does not have the capacity to encode both the 26,000- and the 15,000-dalton proteins unless large overlapping coding regions exist. Tn*5* insertion mutations mapped to this region suggest that the EntA and EntC activities are encoded by the same genetic sequence. Insertion T6 was located 0.5 kb to the right of the T2 insertion into *entB* and inactivated both *entA* and *entC*. The M3 insertion was mapped another 0.5 kb to the right of the T6 insertion. This insertion also inactivated both *entA* and *entC*.

Although *entA* and *entC* could not be separated genetically, minicell expression analysis of these subclones localized the 26,000- and the 15,000-M_r proteins (Fig. 4). pITS16 expressed both proteins. The 0.85-kb *Hpa*I insert in pITS34 encoded the 15,000-M_r protein, as well as a 19,000-M_r polypeptide, which may represent a fusion protein or a truncated derivative of the 26,000-M_r protein; this subclone did not complement either *entA* or *entC*. The pITS38 subclone encoded neither protein, but when the 1.6-kb *Bst*EII fragment was inserted into pITS38, it not only restored *entA* and *entC* activities, but it also encoded both the 26,000- and 15,000-M_r proteins. The discriminating evidence was provided by pITS29, which contained both *entA* and *entC* activities but encoded only the 26,000-M_r protein. On the basis of these data, the EntA and EntC activities appear to reside in the same protein and the 15,000-M_r protein does not appear to be correlated with a known enterobactin gene. Unfortunately, protein profiles of two insertion mutation plasmids (T6 and M3) could not establish whether the 26,000-M_r protein defined by the subcloning data as *entAC* was present, since one of the Km^r proteins also migrated with that apparent molecular weight.

DISCUSSION

The molecular genetic data presented in this report have defined the spatial organization of four structural genes (*entE*, *entB*, *entG*, and *entAC*), which encode the five enzymatic activities in *E. coli* responsible for the production of 2,3-dihydroxybenzoic acid and its subsequent utilization for the synthesis of the catechol siderophore enterobactin. The gene boundaries for these four genes were localized, and their protein products were identified (Figs. 1 and 3).

The *entE* gene is located within the 2.5-kb *Eco*RI-*Pvu*II fragment of pITS15 and was not affected by the T7 insertion mutation which is located 400 bp to the right of the left-end *Eco*RI site. The EntE protein migrated in sodium dodecyl sulfate-polyacrylamide–urea gels, with an apparent molecular weight of 58,000, in agreement with the size reported elsewhere (30). The detection of EntE and its activity from various subclones suggests that *entE* can be expressed in the absence of other cloned *ent* genes.

The *entB* gene is found adjacent to *entE* and extends through the *Eco*RV site that defines the right end of the pITS33 DNA insert (Fig. 1). By use of various subclones and insertion mutations, the EntB enzyme, 2,3-dihydro-2,3-dihydroxybenzoate synthetase, was characterized by gel electrophoresis as a 32,500-dalton protein. The observations that pITS15 and pITS33 produced truncated EntB polypeptides of 20,000 and 29,000 daltons, respectively, provide strong evidence that *entB* is transcribed from left to right. In addition, the latter polypeptide was functional in genetic complementation tests with an *entB* mutant strain, suggesting that the carboxy terminus of EntB is not essential for enzymatic activity.

The weak expression of *entG* in the absence of a full-length *entB* transcript and polarity effects from transposon insertion mutations suggest that *entG* and *entB* form a polycistronic transcript under the control of a promoter upstream of *entB*. Although the EntG protein was reported to be a component in the multienzyme enterochelin synthetase complex (15), its enzymatic activity has not been characterized. On the basis of the genetic complementation of various subclones, *entG* was localized to the 700-bp *Pvu*II-*Hpa*I fragment between *entB* and *entAC*, in agreement with the location suggested previously (21). The *entB* gene extends approximately 400 bp into this fragment from the *Pvu*II end, and the *entAC* gene overlaps the *Hpa*I site at the other end of this fragment, leaving very little nucleotide information to encode EntG. Preliminary sequence analysis of this region has identified a small open reading frame of 25 codons which may represent EntG. Alternatively, EntG activity could reside in the carboxy end of the EntB polypeptide. Each of the subclones which weakly expressed EntG activity (e.g., pITS16, pITS29, pITS38, and pITS38.16) resulted in an in-phase fusion between the Cat enzyme and EntB at the *Pvu*II site. These fusion proteins, which contain the 38 amino-terminal residues of the Cat protein fused to the carboxy-terminal 140 amino acids of EntB, could account for the low level of *entG* complementation. However, pITS441 contained the 700-bp *Pvu*II-*Hpa*I fragment cloned in the opposite orientation to the *lacZ* coding sequence in pUC18 and did not produce a fusion

protein with EntB. This plasmid also weakly complemented the *entG* mutation, suggesting that poor transcription levels, and not the production of fusion proteins, account for this weak expression. No assignable EntG polypeptide was identified in minicell expression experiments. It is possible that the EntG protein is unstable or not resolved in our gel system. The EntD component of enterochelin synthetase, which also has no defined enzymatic activity, was also not detected when plasmid subclones carrying the *entD* structural gene were expressed in *E. coli* minicells (7, 12). The apparent small size of EntG and EntD, as well as the lack of defined enzymatic activity, suggest that these polypeptides may represent structural elements around which the multienzyme complex is constructed.

Biochemical and complementation analyses defined three enzymatic activities (EntC, EntB, and EntA) required to convert the aromatic intermediate chorismate into 2,3-dihydroxybenzoic acid. The enzyme activities designated isochorismate synthetase and 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase were assigned to the genetic complementation groups *entC* and *entA*, respectively (41). However, subsequent genetic studies have not resolved these enzymatic activities into separate genetic loci (20, 21, 30), except in two instances in which the mutagenic agent was the bacteriophage Mu (11, 38). Strong evidence is provided by this study that both activities reside within the same 26,000-dalton protein, which we have designated EntAC. The *entAC* gene is located between the *Eco*RV site and the *Acc*I site, 339 and 1,229 bp, respectively, to the right of the *Pvu*II site marking the left end of the pITS16 DNA insert fragment. These boundaries are defined by the subclones pITS37 and pITS29, respectively (Fig. 3).

Multifunctional enzymes are common in the aromatic amino acid pathways which branch from chorismate (18). In the phenylalanine, tyrosine, and tryptophan biosynthetic pathways, chorismate is the substrate for the first activity of individual multifunctional enzymes which have a sequential second activity. In *Bacillus subtilis*, a multifunctional enzyme is active in nonsequential steps in the synthesis of chorismate and its subsequent conversion to prephenate. The EntAC enzyme also appears to catalyze nonsequential steps in the conversion of chorismate to 2,3-dihydroxybenzoic acid. The characterized *entA* and *entC* point mutations used in this study would then represent isolated mutations in two active sites of a single polypeptide which may function as a single protein or as a multisubunit complex in either enzymatic reaction. We have not ruled out the possibility that the *entA* and *entC* genes overlap and encode proteins of almost equal size or that one of the proteins was not detectable under the conditions used in these experiments. It is important to note that the EntA and EntC activities were separated on the basis of charge by chromatography on DEAE-cellulose (40), although neither activity fraction was analyzed by gel filtration or gel electrophoresis to identify the protein components. Both activities could reside in the same polypeptide chain and still be separable on DEAE-cellulose if (i) either activity is the result of association with another cellular factor or protein, (ii) either activity requires a posttranslational modification of the polypeptide, or (iii) the polypeptide functions alone in one enzymatic reaction but as an oligomer in the other activity.

The preceding evidence supports the original report that this gene cluster is organized into three transcriptional units (21). Tn*5* insertion mutations isolated in the previous study, as well as the insertions and subclones isolated in this study, establish that *entE*, *entBG*, and *entAC* can be expressed independently of each other, although there is the possibility that a major in vivo, iron-regulated transcript contains several of these cistrons. In this regard, it is apparent from the minicell expression data (Fig. 2 and 4) that although EntB was produced in large quantities in the absence of the *entE* gene, the expression of EntAC was reduced when *entB* was deleted.

Two reports suggest that these genes form a single operon. One report was based on large BAL 31 deletions of the 7.0-kb *Eco*RI fragment (30). Deletions which were reported to have removed the operon promoter probably also removed most of the *entAC*, *entB*, *entG*, and *entE* structural genes, resulting in a loss of complementing activity. The second report (11), from our laboratory, based the operon structure on Mu d(Ap^r *lac*) operon fusions. It is not clear why a Mu d insertion into *entA* inactivates *entC*, *entB*, and *entE* even though *entB* and *entE* can be independently transcribed. Mu phage has a propensity for causing deletions and rearrangements, although this was not likely in the previous study, since it was shown that an insertion into *entA* was precisely excised, restoring all *ent* activities. The most plausible explanation is that the complementation assays used in the previous experiments were based on an interpretation of growth on low-iron plates, which is not as sensitive an assay as the bioassays used in this study.

An additional protein of 15,000 M_r has been identified, although there is no known enterobactin gene to code for this protein. The structural gene for the 15,000-M_r protein is located to the right of the *entAC* gene between the *Acc*I site and the *Hpa*I site to its right. Site-specific mutagenesis and marker transfer of the mutation to the chromosome will be required to define what role, if any, this 15,000-M_r protein plays in the production of enterobactin.

ACKNOWLEDGMENTS

We are grateful to J. Wall and C. F. Earhart for helpful discussions during these studies and to Betty Glinsmann for construction of the Tn*5* insertion mutations. We especially thank I. G. Young, M. Casadaban, D. Berg, and B. Bachmann for furnishing bacterial strains and G. Smith and N. Kleckner for providing λ1105. Our thanks to S. Saums and K. Ehlert for preparation of the manuscript.

This work was supported by grants PCM 8210415 and DMB 8416017 from the National Science Foundation.

LITERATURE CITED

1. **Bachmann, B. J.** 1983. Linkage map of *Escherichia coli* K-12, edition 7. Microbiol. Rev. **47**:180–230.
2. **Berg, D. E., M. A. Schmandt, and J. B. Lowe.** 1983. Specificity of transposon Tn*5* insertion. Genetics **105**:813–828.
3. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
4. **Bochner, B. R., H. Huang, G. L. Schieven, and B. N. Ames.** 1980. Positive selection for loss of tetracycline resistance. J. Bacteriol. **143**:926–933.
5. **Casadaban, M. J., and S. N. Cohen.** 1979. Lactose genes fused to exogenous promoters in one step using a Mu-*lac* bacteriophage: *in vivo* probe for transcriptional control sequences. Proc. Natl. Acad. Sci. USA **76**:4530–4533.
6. **Chang, A. C. Y., and S. N. Cohen.** 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. **134**: 1141–1156.
7. **Coderre, P. E., and C. F. Earhart.** 1984. Characterization of a plasmid carrying the *Escherichia coli* K-12 *entD*, *fepA*, *fes* and *entF* genes. FEMS Microbiol. Lett. **25**:111–116.
8. **Cohen, S. N., A. C. Y. Chang, and L. Hsu.** 1972. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of

Escherichia coli by R-factor DNA. Proc. Natl. Acad. Sci. USA **69**:2110–2114.
9. **Cox, G. B., F. Gibson, R. K. J. Luke, N. A. Newton, I. G. O'Brien, and H. Rosenberg.** 1970. Mutations affecting iron transport in *Escherichia coli*. J. Bacteriol. **104**:219–226.
10. **Csonka, L. N., and A. J. Clark.** 1979. Deletions generated by the transposon Tn*10* in the *srl recA* region of the *Escherichia coli* K-12 chromosome. Genetics **93**:321–343.
11. **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.
12. **Fleming, T. P., M. S. Nahlik, J. B. Neilands, and M. A. McIntosh.** 1985. Physical and genetic characterization of cloned enterobactin genomic sequences from *Escherichia coli* K-12. Gene **34**:47–54.
13. **Frazer, A. C., and R. Curtiss III.** 1975. Production, properties and utility of bacterial minicells. Curr. Top. Microbiol. Immunol. **69**:1–84.
14. **Frost, G. E., and H. Rosenberg.** 1975. Relationship between the *tonB* locus and iron transport in *Escherichia coli*. J. Bacteriol. **124**:704–712.
15. **Greenwood, K. T., and R. K. J. Luke.** 1976. Studies on the enzymatic synthesis of enterochelin in *Escherichia coli* K12: four polypeptides involved in the conversion of 2,3-dihydroxybenzoate to enterochelin. Biochim. Biophys. Acta **454**:285–297.
16. **Hancock, R. E. W., K. Hantke, and V. Braun.** 1976. Iron transport in *Escherichia coli* K-12: involvement of the colicin B receptor and of a citrate-inducible protein. J. Bacteriol. **127**: 1370–1375.
17. **Hollifield, W. C., and J. B. Neilands.** 1978. Ferric enterobactin transport system in *Escherichia coli* K-12. Extraction, assay, and specificity of the outer membrane receptor. Biochemistry **17**:1922–1928.
18. **Kirschner, K., and H. Bisswanger.** 1976. Multifunctional proteins. Annu. Rev. Biochem. **45**:143–166.
19. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
20. **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.
21. **Laird, A. J., and I. G. Young.** 1980. Tn*5* mutagenesis of the enterochelin gene cluster of *Escherichia coli*. Gene **11**:359–366.
22. **Langman, L., I. G. Young, G. E. Frost, H. Rosenberg, and F. Gibson.** 1972. Enterochelin system of iron transport in *Escherichia coli*: mutations affecting ferric-enterochelin esterase. J. Bacteriol. **112**:1142–1149.
23. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. **McIntosh, M. A., and C. F. Earhart.** 1976. Effect of iron on the relative abundance of two large polypeptides of the *Escherichia coli* outer membrane. Biochem. Biophys. Res. Commun. **70**:315–322.
25. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
26. **Norrander, J., T. Kempe, and J. Messing.** 1983. Construction of improved M13 vectors using oligodeoxynucleotide-directed mutagenesis. Gene **26**:101–106.
27. **O'Brien, I. G., and F. Gibson.** 1970. The structure of enterochelin and related 2,3-dihydroxy-*N*-benzoylserine conjugates from *Escherichia coli*. Biochim. Biophys. Acta **215**:393–402.
28. **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**:3647–3653.
29. **Pettis, G. S., and M. A. McIntosh.** 1987. Molecular characterization of the *Escherichia coli* enterobactin cistron *entF* and its coupled expression to the *fes* gene. J. Bacteriol. **169**:4154–4162.
30. **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.
31. **Pierce, J. R., and C. F. Earhart.** 1986. *Escherichia coli* K-12 envelope proteins specifically required for ferrienterobactin uptake. J. Bacteriol. **166**:930–936.
32. **Pollack, J. R., B. N. Ames, and J. B. Neilands.** 1970. Iron transport in *Salmonella typhimurium*: mutants blocked in the biosynthesis of enterobactin. J. Bacteriol. **104**:635–639.
33. **Pollack, J. R., and J. B. Neilands.** 1970. Enterobactin, an iron transport compound from *Salmonella typhimurium*. Biochem. Biophys. Res. Commun. **38**:989–992.
34. **Schleif, R. F., and P. C. Wensink.** 1981. Practical methods in molecular biology, p. 98. Springer-Verlag, New York.
35. **Shaw, K. J., and C. M. Berg.** 1979. *Escherichia coli* K-12 auxotrophs induced by insertion of the transposable element Tn*5*. Genetics **92**:741–747.
36. **Soberon, X., L. Covarrubias, and F. Bolivar.** 1980. Construction and characterization of new cloning vehicles. IV. Deletion derivatives of pBR322 and pBR325. Gene **9**:287–305.
37. **Way, J. C., M. A. Davis, D. Morisato, D. E. Roberts, and N. Kleckner.** 1984. New Tn*10* derivatives for transposon mutagenesis and for construction of *lacZ* operon fusions by transposition. Gene **32**:369–379.
38. **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.
39. **Wookey, P., and H. Rosenberg.** 1978. Involvement of inner and outer membrane components in the transport of iron and in colicin B action in *Escherichia coli*. J. Bacteriol. **133**:661–666.
40. **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.
41. **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.