The *entD* Gene of the *Escherichia coli* K12 Enterobactin Gene Cluster

By PETER E. CODERRE AND CHARLES F. EARHART*

Department of Microbiology, The University of Texas, Austin, TX 78712, USA

(Received 9 May 1989; revised 25 July 1989; accepted 3 August 1989)

The *Escherichia coli entD* gene encodes a product necessary for the synthesis of the iron-chelating and transport molecule enterobactin (Ent); cells harbouring *entD* mutations fail to grow in iron-deficient environments. For unknown reasons, it has not been possible to identify the *entD* product. The nucleotide sequence of the *entD* region has now been determined. An open reading frame extending in the same direction as the adjacent *fepA* gene and capable of encoding an approximately 24 kDa polypeptide was found; it contained a high percentage of rare codons and two possible translational start sites. Complementation data suggested that EntD proteins truncated at the carboxy terminus retain some activity. Two REP sequences were present upstream of *entD* and an IS186 sequence was observed downstream. RNA dot-blot hybridizations demonstrated that *entD* is transcribed from the strand predicted by the sequencing results. An *entD-lacZ* recombinant plasmid was constructed and shown to express low amounts of a fusion protein of the anticipated size (approximately 125 kDa). The evidence suggests a number of possible explanations for difficulties in detecting the *entD* product. Sequence data indicate that if *entD* has its own promoter, it is weak; the REP sequences suggest that *entD* mRNA may be destabilized; and translation may be slow because of the frequency of rare codons and a possible unusual start codon (UUG). The data are also consistent with previous evidence that the *entD* product is unstable.

INTRODUCTION

Iron acquisition is an essential activity for the enteric bacterium *Escherichia coli* which, under low-iron conditions, accomplishes this by means of a highly specific iron-transport pathway. The system common to all *E. coli* requires the synthesis and secretion of an iron-chelating agent (siderophore) known as enterobactin (Ent) as well as the synthesis of specific envelope-associated ferri-Ent retrieval components (for a review, see Earhart, 1987). Ent is a low-molecular-mass, high-affinity Fe(III)-binding compound that is a cyclic trimer of 2,3-dihydroxybenzoylserine. The Ent system is negatively regulated by means of the Fur protein (for a review, see Bagg & Neilands, 1987).

The biosynthesis of Ent is initiated with the synthesis of 2,3-dihydroxybenzoate from chorismate by means of the *entC*, *entB* and *entA* gene products. In a poorly understood process, 2,3-dihydroxybenzoate is then combined with L-serine and cyclized in a series of reactions catalysed by the *entD*, *entE*, *entF* and *entG* gene products (Greenwood & Luke, 1976, 1980; Woodrow *et al.*, 1979). Recent nucleotide sequencing data (Liu *et al.*, 1989; Nahlik *et al.*, 1989) indicate that EntG activity may be associated with the carboxy terminus of EntB; evidence obtained in this laboratory (J. Staab, unpublished observations) supports this idea.

The transport and biosynthetic genes for the Ent system are clustered at approximately minute 13.5 on the *E. coli* chromosome. The gene order of the left end of this region is *entD*, *fepA*, *fes* and *entF* (Coderre & Earhart, 1984; Fleming *et al.*, 1985). While the *entD* and *entF* products are part of the Ent synthetase, the *fes* product is involved in iron utilization and the *fepA* product

Abbreviations: Ent, enterobactin; ORF, open reading frame.

serves as the outer-membrane receptor for ferri-Ent. Maxicells programmed with plasmid pMS101, which carries these genes, were used to demonstrate that 81, 42, and 115 kDa polypeptides correspond to the *fepA*, *fes* and *entF* gene products, respectively. No protein could be assigned to the *entD* gene. Recently, Pettis *et al.* (1988) reported the iron-controlled, divergent transcription of the *fepA* and *fes-entF* genes and suggested that a single, overlapping promoter region containing two Fur-binding sites was responsible.

Previous workers (Greenwood & Luke, 1976; Woodrow *et al.*, 1979) have suggested that the *entD* product associates with the products of the *entF* and *entG* genes to form the multienzyme complex Ent synthetase, that it has unstable activity, possibly stabilized by the *entF* product (Greenwood & Luke, 1980), and that the Ent synthetase complex may associate with the cytoplasmic membrane through the *entD* polypeptide. This study was undertaken to further characterize the *entD* gene and its elusive product.

METHODS

Bacterial strains and plasmids. The relevant characteristics and sources of *Escherichia coli* K12 strains and plasmids and bacteriophages used in this study are shown in Tables 1 and 2, respectively.

Media and reagents. Minimal medium agar plates and siderophore detection (CAS) medium were prepared according to the recipes of Howard-Flanders *et al.* (1964) and Schwyn & Neilands (1987), respectively. LB medium, which was used for routine growth of bacteria, and iron-depleted M9 medium have been described (Elish *et al.*, 1988). When appropriate, antibiotics (Sigma) were added to final concentrations of 0.2 mg carbenicillin ml⁻¹, 0.05 mg tetracycline ml⁻¹ and 0.05 mg neomycin ml⁻¹. Isopropyl β -D-thiogalactopyranoside (IPTG) was from Sigma. 5-Bromo-4-chloro-3-indoyl β -D-galactopyranoside (X-gal) was obtained from Bethesda Research Labs (BRL), and used at a final concentration of 0.04 mg ml⁻¹. L-[³⁵S]Methionine and [α -³⁵S]thio-dATP were obtained from Dupont New England Nuclear Corporation (NEN).

Labelling of plasmid-encoded proteins. The procedure of Meagher *et al.* (1977) was used to study the expression of plasmid-borne genes in minicell strain P678-54.

Electrophoresis of proteins. Protein samples were electrophoresed on an 11% (w/v) polyacrylamide-SDS gel system (Lugtenberg *et al.*, 1975). Prior to electrophoresis, samples were placed in solubilization buffer and boiled for 5 min (Laemmli, 1970). Proteins were visualized by a modification of the staining procedure of Fairbanks *et al.* (1971). [³⁵S]Methionine labelled proteins were detected by autoradiography on Kodak XRP-1 film.

Plasmid isolation and transformation. The alkaline lysis procedure of Birnboim & Doly (1979) was used to isolate plasmid DNA. Further purification was done by caesium chloride/ethidium bromide density-gradient centrifugation (Maniatis *et al.*, 1982). CaCl₂-treated cells were transformed with plasmid DNA by the method of Cohen *et al.* (1972).

DNA manipulations. Restriction enzymes were obtained from BRL or New England Biolabs (NEB) and were used according to the manufacturer's instructions. Tn5 transpositional mutagenesis was performed according to the procedure of deBruijn & Lupski (1984). Deletions in the sense strand were generated with T4 polymerase (Dale *et al.*, 1985) in M13 single-stranded DNA harbouring a 2.25 kb *Bam*HI-*Eco*RV insert from pPC12. Resulting *Eco*RI-*Hind*III fragments containing shortened regions were subcloned into pBR322. BAL31 (BRL) was used to generate deletions in the antisense strand of pMS101 DNA according to the manufacturer's instructions. Blunt-ended *Pst*I fragments were ligated into M13mp19 and subsequent *Eco*RI-*Hind*III fragments were ligated into pBR322. Plasmid DNA was analysed on horizontal agarose slab gels in TBE buffer (McDonnell *et al.*, 1977).

DNA sequencing. The chain-termination method of Sanger *et al.* (1977), using deoxyadenosine 5'- α -[³⁵S]thiotriphosphate as label (Biggen *et al.*, 1983), was used to analyse DNA sequences. Clones for DNA sequencing were generated as described above and nucleotide sequence data were analysed using the DNASTar program (DNASTar, Madison, Wisconsin, USA).

Construction of *entD-lacZ* gene fusions. A 330 bp *Hae*III fragment (bases 96-397, Fig. 2, bases 3701-3730 from pGEM) from pPC01G, containing the SP6 promoter, the presumed *entD* ribosome-binding site and part of the structural gene for *entD*, was purified by agarose gel electrophoresis and then cloned into the *Sma*I site of pMC1403 (Casadaban *et al.*, 1980). This vector allows the insertion of a DNA fragment in-frame and prior to the 9th codon of the *lacZ* gene. The *lac* strain TB1, an MC4100 derivative, was transformed with the ligation mixture and ampicillin-resistant transformants were picked. Plasmid DNA was isolated and restriction mapping was used to determine the presence and orientation of the insert in the plasmid vector.

Assays. Clones were tested for β -galactosidase production by streaking onto minimal medium containing ampicillin and X-gal; succinate and dipyrindyl replaced glucose to test for iron regulation. Enzyme assays (Miller, 1972) were done on cells grown in minimal medium supplemented with glycerol and either 0.02 mM-FeCl₃ or no iron.

Table 1. *E. coli* strains

Strain	Description	Source/reference
AB1515	F ⁻ <i>fhuA purE42 proC14 leu-6 trpE38 thi-1</i>	Coli Genetic Stock Center
AN90	<i>entD purE</i> ⁺ derivative of AB1515	J. B. Neilands
AN90 <i>recA</i>	<i>recA</i> derivative of AN90	J. R. Pierce
UT5600	<i>entD</i> and <i>fepA</i> deletion derivative of RW193 (RW193 is an <i>entA403 purE</i> ⁺ <i>fhuA</i> ⁺ derivative of AB1515)	Elish <i>et al.</i> (1988)
P678-54	<i>thr leu thi supE lacY fhuA gal mal xyl ara mtl minA minB</i>	J. R. Walker/Adler <i>et al.</i> (1967)
JM103	<i>rpsL endA1 thi-1 hsdR17 supE44 traD36 sbcB</i> $\Delta(lac\ proA,B)/F'$ <i>proA,B lacI^s lacZ</i> Δ M15	J. R. Walker/Yanisch-Perron <i>et al.</i> (1985)
TB1	<i>ara</i> $\Delta(lac\ proA,B)$ <i>rpsL hsdR</i> ϕ 80 ^R <i>lacZ</i> Δ M15	I. J. Molineux

Table 2. *Plasmids and bacteriophages*

Strain	Description	Source/reference
pBR322	Ap ^R Tc ^R	
pMS101	pBR322 derivative, Ap ^R , <i>entD</i> ⁺ <i>fes</i> ⁺ <i>fepA</i> ⁺ <i>entF</i> ⁺	I. G. Young
pPC5-4	pMS101 derivative, <i>entD</i> :: Tn5	This study
pPC12	pBR322 derivative, <i>entD</i> ⁺	Coderre & Earhart (1984)
M13mp8, M13mp18, M13mp9, M13mp19		J. R. Walker/Yanisch-Perron <i>et al.</i> (1985)
λ 467	<i>b221 rex</i> :: Tn5 <i>cI</i> 857 <i>Oam29 Pam80</i>	Berg <i>et al.</i> (1979)
M13mp1813a18	mp18 derivative containing 1.1 kb of <i>dnaZX</i>	J. R. Walker
pGEM-blue	pBR322 derivative containing multiple cloning site from pUC18	Promega Biotec
pT7/T3-18, pT7/T3-19	pBR322 derivatives containing multiple cloning site from pUC18/pUC19; contain T3 and T7 promoters	I. J. Molineux
pMC1403	pBR322 derivative containing the <i>lacZ</i> structural gene but missing the first 8 codons	I. J. Molineux
pPC17LZ	pMC1403 derivative containing a 330 bp <i>Hae</i> III fragment from pPC01G in-frame with <i>lacZ</i>	This study
pPC18LZ	pMC1403 derivative containing a 330 bp <i>Hae</i> III fragment from pPC01G out-of-frame with <i>lacZ</i>	This study

Ent was estimated using the phenolate assay of Arnow (1937).

Transcriptional studies. Total cellular RNA was isolated by the method of von Gabain *et al.* (1983) and blotted onto GeneScreen membrane filters (White & Bancroft, 1982). Each top row dot contained 25 μ g RNA, with dots in each subsequent row containing half the amount of RNA present in the previous row. Filters were hybridized according to the manufacturer's instructions with [³⁵S]UTP-labelled RNA probes (10⁶ c.p.m.) synthesized from appropriate plasmid DNA (Melton *et al.*, 1984). These plasmids contain a multiple cloning site bracketed by bacteriophage promoters; the pGEM-blue plasmid contains promoters for the SP6 and T7 RNA polymerases while the pT7/T3 plasmids contain promoters for the T3 and T7 RNA polymerases. SP6 and T7 RNA polymerases were obtained from Promega Biotec and NEB, respectively.

RESULTS

Construction of mutant plasmids

Previous studies of the 10.6 kb *Hind*III fragment within pMS101 determined the location of the *entD*, *fepA*, *fes* and *entF* genes (Coderre & Earhart, 1984). Also, a 4.0 kb *Eco*RI fragment from pMS101 cloned into pBR322 to yield pPC12 was shown to genetically complement AN90 (*entD*).

Fig. 1. Restriction map, cloning strategy for the *entD* sequence, and complementing activities of *entD*-containing plasmids. A 4.1 kb region extending from the start of *fepA* to a region downstream of *entD* is indicated at the top of the figure. Pertinent restriction sites (B, *Bam*HI; H2, *Hinc*II; P, *Pst*I, RI, *Eco*RI; RV, *Eco*RV; S, *Stu*I; U, *Pvu*I) and Tn5 insertion sites are indicated. Tn5 insertions that do or do not inactivate a gene are indicated by filled and open circles, respectively. The locations of *fepA* and *entD* are designated by heavy arrows. Below the *entD* region, each line represents the region of an individual clone sequenced. Arrows indicate the terminal nucleotide in each determination; arrows pointed to the right indicate that the clone was of the anti-sense DNA strand, while arrows pointed to the left indicate that the clone was of the sense DNA strand. The AN90 complementing activity of the subclone plasmids is indicated by a plus or minus after the plasmid name.

Transposon mutagenesis of pMS101 with Tn5 was performed and two pertinent mutations in the *entD* region were obtained (Fig. 1), one of which inactivated *entD*. Three different Tn5 insertions were found that inactivated *fepA* and each remained *entD*⁺ in complementation studies.

A number of deletion mutants in the *entD* region were generated in order to further characterize the extent of the *entD*-complementing region (Fig. 1) as well as to generate clones for DNA sequencing.

Genetic complementation by mutant plasmids

Complementation patterns of the *entD* deletion and Tn5 insertion mutation plasmids in AN90 and AN90*recA* strains were determined on CAS and low-iron plates. Growth on succinate/dipyridyl plates and the production of orange haloes on CAS blue agar plates were considered positive tests for genetic complementation.

AN90 cells were complemented by deletion plasmids pPC372, pPC1002 and pPC5172 (Fig. 1). AN90 cells harbouring pPC5172 produced smaller orange haloes than those produced by AN90 cells harbouring pPC372 or pPC1002. Deletion plasmids pPC832, pPC442, and the pMS101*entD*::Tn5 plasmid (pPC5-4) did not complement AN90 by either assay; AN90 cells harbouring pPC5122, pPC5152 and pPCA192 exhibited weak growth on succinate/dipyridyl plates but did not produce orange haloes on CAS blue plates.

Nucleotide sequence of *entD*

The strategy shown in Fig. 1 was used to determine the nucleotide sequence of both strands of a 902 bp region encompassing *entD* (Fig. 2); the sequence of the last 175 bases (nos 903–1077) was determined from one strand only.

Computer-assisted analysis of potential coding sequences demonstrated the existence of one large open reading frame (ORF). Depending upon the translational initiation codon (AUG,

Fig. 2. Nucleotide sequence of the *entD* gene. A 1077 bp region was sequenced. Amino acid assignments for each codon in the proposed reading frame are indicated in one-letter code below the first base of each codon. The termination codon is designated with an asterisk and the proposed -35 and -10 promoter sites are underlined. The two REP sequences upstream of *entD* are indicated with arrows. The boxed nucleotides represent a consensus sequence of unknown function. The proposed Shine-Dalgarno sequence is both under- and overlined. Potential initiation codons are indicated by asterisks above the bases. The palindromic sequences of the proposed termination structure are indicated by arrows downstream of the ORF. The DNA sequence homologous to IS186 is indicated with line and arrow beneath the nucleotides. Relevant restriction sites are indicated above the nucleotide sequence.

position 149, or UUG, position 130) used, this ORF could encode a 23 606 Da (209 amino acids) or a 24 194 Da (215 amino acids) polypeptide, respectively. [Two considerations led to the inclusion of UUG as a possible *entD* initiation codon: (i) its placement with respect to a potential ribosome-binding site (see below); and (ii) the facts that the *entD* product has not been detected and that, other factors being the same, gene expression is reduced when the initiation codon is UUG rather than AUG or GUG (Reddy *et al.*, 1985).] Either of these start sites is consistent with the complementation data. Table 3 shows that an unusually high fraction – 16/209 (7.66%) and 17/215 (7.91%) – of the codons are classified as 'rare' (Sharp & Li, 1986).

Table 3. *Codon usage in EntD*

Rare codons are indicated by asterisks. Numbers in parentheses indicate additional codons used for the larger EntD protein.

GCA Ala	9	CAG Gln	5	UUG Leu	2 (3)	UAA Ter	0
GCC Ala	5		12		18 (19)	UAG Ter	0
GCG Ala	4	GAA Glu	6	AAA Lys	7	UGA Ter	0
GCU Ala	5	GAG Glu	10	AAG Lys	2		0
	23		16		9	ACA Thr	4
*AGA Arg	1	*GGA Gly	3	AUG Met	3	ACC Thr	2
*AGG Arg	0	GGC Gly	4		3	ACG Thr	4
*CGA Arg	1	*GGG Gly	1	UUC Phe	1	ACU Thr	3
CGC Arg	1	GGU Gly	2	UUU Phe	9		13
*CGG Arg	2		10		10	UGG Trp	4
CGU Arg	4	CAC His	6	CCA Pro	1		4
	9	CAU His	5	*CCC Pro	2	UAC Tyr	2
AAC Asn	1		11	CCG Pro	3	UAU Tyr	4
AAU Asn	3 (4)	*AUA Ile	5	CCU Pro	2		6
	4 (5)	AUC Ile	4		8	GUA Val	3
GAC Asp	3	AUU Ile	8	AGC Ser	3	GUC Val	4 (5)
GAU Asp	6		17	AGU Ser	2	GUG Val	3
	9	*CUA Leu	1	UCA Ser	0 (2)	GUU Val	2
UGC Cys	2	CUC Leu	3	UCC Ser	2		12 (13)
UGU Cys	3	CUG Leu	8	*UCG Ser	0 (1)		
	5	CUU Leu	0	UCU Ser	3		
CAA Gln	7	UUA Leu	4		10 (13)	Total	209 (215)

Also, the putative EntD protein would contain four Cys residues, which is consistent with previous enzymic data (Greenwood & Luke, 1976) indicating that a thiol group is essential for EntD activity. No genes with significant sequence homology to *entD* were found in the DNASTar DNA sequence database.

Two REP sequences (highly conserved inverted repeats) located upstream of the *entD* gene and downstream of the *fepA* gene (Fig. 2) were identified, in agreement with the work of others (Lundrigan & Kadner, 1986; Pettis *et al.*, 1988). These sequences may form stem-loop structures which stabilize upstream messages, presumably by inhibiting 3'-5' exonucleases, and are frequently located between genes that are expressed at different levels (Newbury *et al.*, 1987*a, b*).

Several potential promoter sequences were found upstream of the ORF. A potential -35 RNA polymerase binding site (TTGCCA) which matched with five out of six of the consensus sequence bases (TTGACA) was present 62 bp upstream of the UUG codon. Further downstream, a potential -10 RNA polymerase binding site (TATATT) matched with the consensus bacterial Pribnow box (TATAAT) in five out of six positions. The putative promoter sequences were unusual in that the distance between the last base of the -35 sequence and the first base of the -10 sequence was six bases rather than the 14-17 bases which is normally observed. The DNA sequence TGATAAGCGT was observed between the possible -10 site and a potential Shine-Dalgarno ribosome-binding site (Fig. 2). Sequences similar to this are also found upstream of the vitamin B₁₂ receptor gene, *btuB*, and the iron-regulated genes *fhuA*, *iutA* and *fepA* (Lundrigan & Kadner, 1986).

A potential ribosome-binding site (Shine-Dalgarno sequence) was centred at position 117. This sequence, AGGCAGT, matched at five out of seven positions when compared with consensus bacterial sequences (Shine & Dalgarno, 1974). UUG and AUG initiation codons were positioned 7 and 27 bases, respectively, downstream of the potential Shine-Dalgarno sequence. A potential stem and loop structure [$\Delta G(25^\circ\text{C}) = -18.8 \text{ kcal mol}^{-1}$ ($-78.7 \text{ kJ mol}^{-1}$) (Tinoco *et al.*, 1973)] was located downstream of *entD* between bases 798 and 829 (indicated with arrows in Fig. 2) and may represent a rho-dependent transcriptional termination site.

Further downstream of the proposed *entD* structural gene and its corresponding stem and loop structure, a base sequence that was homologous to a portion of the IS186 sequence was present.

Fig. 3. Construction of plasmids for RNA probe synthesis. (a) The top line is a restriction map of the 4.0 kb *EcoRI* insert from pPC12 which includes *entD* and a portion of the *fepA* region. The *entD* and *fepA* genes are indicated by heavy arrows. pPC01G contains a 1.1 kb *SmaI*-*BamHI* fragment from pPC12 cloned into the multiple cloning site (MCS) of pGEM. pPC818G contains a 1.1 kb fragment from the *fepA* region cloned into the MCS of pGEM. The direction of the SP6 and the T7 promoters in pPC01G and pPC818G, respectively, are indicated by short arrows above the insert. (b) The top line is a restriction map of the *dnaZX* region (Yin *et al.*, 1986). pCXZG contains a 1.1 kb insert from M13mp812a18 cloned into the MCS of pGEM. The direction of the T7 promoter in pCXZG is indicated by a short arrow above the insert. A, *AvaI*; B, *BamHI*; Bs, *BstI*; RV, *EcoRV*; H2, *HincII*; K, *KpnI*; N, *NruI*; P, *PstI*; Pv, *PvuI*; RI, *EcoRI*; Sm, *SmaI*.

From positions 885 to 1077 (Fig. 2), there was a 97% homology to the terminal 193 nucleotides of the *IS186* sequence (Kothary *et al.*, 1985). This region downstream of *entD* also shared similar *HaeIII*, *BamHI* and *PstI* restriction sites with *IS186* (Chong *et al.*, 1985).

RNA dot-blot hybridizations

Transcriptional studies utilized radioactively labelled RNA probes derived from DNA sequences for the *fepA*, *dnaZX* and *entD* regions for RNA dot-blot hybridizations (Fig. 3); the *dnaZX* probe served as a control as this region is believed to be expressed constitutively. The probes were hybridized to total cellular RNA from strains AB1515, AN90(pMS101), AN90(pPC12) and UT5600. Only the results obtained with RNA from AN90(pMS101) are shown (Fig. 4).

The plasmid pPC818G was used to study *fepA* transcription (Fig. 3). The RNA probe generated by the T7 RNA polymerase was transcribed from the antisense (non-coding) strand and this RNA probe, but not that generated by the SP6 RNA polymerase, did hybridize to RNA

Fig. 4. Autoradiograms of RNA dot-blot hybridizations of RNA probes to RNA from AN90(pMS101). Odd-numbered lanes contain RNA from AN90(pMS101) grown in deferrated minimal medium supplemented with 0.02 mM-FeCl₃; even-numbered lanes contain RNA prepared from AN90(pMS101) grown in deferrated minimal medium without additional iron. Lanes 1 and 2, pPC01G SP6 probe; lanes 3 and 4, pPC01G T7 probe; lanes 5 and 6, pPC818G T7 probe; lanes 7 and 8, pPC818G SP6 probe; lanes 9 and 10, pPCXZG SP6 probe; lanes 11 and 12, pPCXZG T7 probe.

from AB1515, AN90(pMS101) and AN90(pPC12). That is, the coding strand was in agreement with sequence data of others (Lundrigan & Kadner, 1986; Pettis *et al.*, 1988). This same probe did not hybridize to RNA from UT5600, a strain deleted for *fepA* and *entD*. Transcription of *fepA* was iron-regulated; hybridization to RNA from cells grown in deferrated M9 minimal medium was enhanced compared to hybridization to RNA from cells grown in the same medium supplemented with 0.02 mM-FeCl₃, particularly when plasmid pMS101, which contains *fepA* as well as *entD*, was present (Fig. 4). Iron starvation and iron-replete conditions were verified by assaying culture supernatant fluids for phenolates.

The plasmid pPCXZG was constructed from M13mp812a18. The T7 but not the SP6 RNA polymerase generated RNA probe hybridized to all total cellular RNA preparations as anticipated (Yin *et al.*, 1986) and transcription of *dnaZX* was not iron-regulated (Fig. 4).

Strand-specific RNA probes containing the *entD* gene were constructed. SP6 RNA polymerase generated probes from pPC01G hybridized to total cellular RNA from AB1515, AN90(pMS101) and AN90(pPC12) but not UT5600, indicating that *entD* transcription was from the same strand as *fepA*. In AB1515 cells, the amount of *entD* transcript was low. As with *fepA*, hybridization was enhanced when RNA was isolated from plasmid-containing strains. Iron regulation of *entD* transcription was most obvious when the plasmid pMS101, which contains a Fur box upstream of *fepA*, was present (Fig. 4).

EntD-LacZ protein fusions

An *entD-lacZ* gene fusion was constructed as described in Methods; plasmids containing the *Hae*III fragment in both the correct and incorrect orientation were obtained. TB1 cells harbouring the recombinant plasmids were tested for β -galactosidase production by two different methods. Cells were streaked both on L plates containing ampicillin and X-gal and on minimal medium containing succinate, dipyrindyl, ampicillin and X-gal. Cells harbouring plasmids containing the insert in the incorrect orientation (pPC18LZ) produced white colonies on both media while AB1515 cells and cells harbouring the recombinant plasmid with the insert in the proper orientation (pPC17LZ) produced blue colonies on both media.

To determine if β -galactosidase production was iron-regulated, enzyme assays were performed on cells grown in minimal medium supplemented with glycerol and either 0.02 mM-FeCl₃ or no iron (Table 4). Cells harbouring pPC17LZ produced approximately the same low

Fig. 5. Autoradiogram of proteins synthesized in minicells programmed with pMC1403 and pMC1403 recombinant plasmids. [³⁵S]Methionine-labelled proteins from minicells harbouring recombinant plasmids were electrophoresed on an 11% polyacrylamide gel. The arrow indicates the EntD-LacZ fusion protein. Lane 1, no plasmid; lane 2, pMC1403; lane 3, pPC17LZ; lane 4, pPC18LZ. The positions of the molecular mass markers are shown on the right: myosin (heavy chain), phosphorylase *b*, bovine serum albumin, ovalbumin, α -chymotrypsinogen, β -lactoglobulin and lysozyme.

Table 4. *β -Galactosidase production in TB1 and AB1515 cells*

Strain/plasmid	Presence/absence of iron	Units of enzyme
TB1/pMC1403	+	2
	—	2
TB1/pPC17LZ	+	22
	—	20
TB1/pPC18LZ	+	3
	—	7
AB1515 (+IPTG)	+	1334
AB1515 (—IPTG)	+	23

level of the enzyme under iron-replete and iron-poor conditions; the activity detected was that observed in uninduced AB1515 cells, which have a wild-type *lac* operon.

Autoradiograms of the proteins directed by pPC17LZ, pPC18LZ and pMC1403 showed strong expression of β -lactamase and, in minicells harbouring pPC17LZ, weak expression of an approximately 125000 Da protein (Fig. 5). This latter band is presumed to be a hybrid protein, a result of the fusion of approximately 10000 Da of EntD and 115772 Da of LacZ and three linker-encoded amino acids.

DISCUSSION

The following results indicate that the *entD* product is a protein: (i) an ORF that was transcribed from the predicted strand and whose size corresponded to complementation data was detected; (ii) construction details for plasmid pPC17LZ, which encodes an EntD-LacZ fusion protein, showed that the ORF was as anticipated and therefore functional for the synthesis of the hybrid protein; (iii) the size of the hybrid protein was consistent with the sequencing data. The results regarding the nature of the *entD* product are therefore in agreement with earlier studies (Greenwood & Luke, 1976; Woodrow *et al.*, 1979) on Ent synthetase components.

Nonetheless, we have been unable to detect the EntD protein. In addition to procedures previously described (Coderre & Earhart, 1984), the following procedures and variations have also been unsuccessful: urea denaturing gels, two-dimensional polyacrylamide gels, high pH and low pH non-denaturing discontinuous polyacrylamide gels, use of high and low copy number plasmids, including placing *entD* under the control of a bacteriophage T7 promoter and infecting cells with T7 (Dunn & Studier, 1983) or inducing a chromosomally-encoded T7 RNA polymerase (Studier & Moffat, 1986), use of a variety of radioactive amino acids for labelling, addition of protease inhibitors or iron chelating agents to *in vitro* DNA-directed protein synthesizing systems, examination of subcellular fractions, and use of subclones containing *entD* and *fepA*, as well as *entD* alone.

The current results offer several possible explanations, which are not mutually exclusive, for our inability to observe the EntD protein. The most obvious of these is the presence of the REP sequences, which, if *fepA* and *entD* are transcribed together on a polycistronic message (see below), may be stabilizing the upstream *fepA* mRNA and destabilizing the *entD* part of the transcript. The dot-blot results, in which *entD* mRNA was readily detected, could be misleading as they do not distinguish between functional and non-functional mRNA. A second explanation is that inefficient translation may cause the *entD* gene to be expressed at low levels. Poor translation is indicated by the relatively high percentage of rare codons (Table 3); in *E. coli*, rare codons occur at a frequency of $3.5\% \pm 2.6\%$ and $8.8\% \pm 1.2\%$ in highly expressed proteins and regulatory proteins, respectively (Konigsberg & Godson, 1983). Less than optimal translational efficiency is also suggested by the initiation codon, regardless of which (AUG or UUG) is used. For AUG, the ribosome-binding site which best fits the consensus sequence is 27 bases upstream, a distance far greater than the 5–13 base spacing that is considered optimal (Gold, 1988). An alternative Shine-Dalgarno sequence (ATCAGTC) is centred six bases upstream of the AUG codon, but, as it matches the consensus sequence in only three of seven positions, it too would be expected to have poor translational yield. UUG, which is the initiation codon for several *E. coli* proteins including NADH dehydrogenase (Poulis *et al.*, 1981), ribosomal protein S20 (Mackie, 1981), and adenylate cyclase (Roy *et al.*, 1983), is properly positioned with respect to the putative ribosome-binding site but, for adenylate cyclase at least, UUG results in reduced expression compared to transcripts with AUG or GUG initiation codons (Reddy *et al.*, 1985). Failure to detect EntD may also indicate that the protein is chemically labile; functional instability for EntD has been reported (Greenwood & Luke, 1976, 1980). Other components of the Ent synthetase complex, EntF, EntB(G) and possibly EntE, are readily observed; as part of this complex, EntD would be expected to be present in amounts comparable to these proteins. Interestingly, the putative EntD protein sequence contains four consecutive basic amino acids (Arg-Lys-Arg-Lys; residues 53–56 of the larger EntD) and outer-membrane protease OmpT (Rupprecht *et al.*, 1984), also known as protease VII (Sugimura & Higashi, 1988), cleaves between adjacent basic amino acids (Sugimura & Nishihara, 1988). If this protease attacked EntD in cell lysates, this could at least partially explain the difficulties in identifying EntD. Attempts to observe EntD in *ompT* strains were unsuccessful (M. C. Ammerlaan, unpublished observations) but this basic region could still be a cleavage site for other proteases. The most promising opportunity for observing EntD may involve the production of antibody to EntD peptides derived from the nucleotide sequence. These antibodies could then be used to identify and locate the EntD protein within the cell.

Synthesis of all Ent synthetase components, including EntD, is repressed by inclusion of iron

in the growth medium (Greenwood & Luke, 1981). No known iron regulatory sites were found in the *fepA-entD* intergenic region but two Fur boxes apparently exist upstream of *fepA* (Pettis *et al.*, 1988). Here, iron regulation of *entD* mRNA was most obvious when RNA was isolated from a strain harbouring a plasmid (pMS101) that contained both *entD* and *fepA* and its regulatory region. This indicates that transcription of *entD* normally originates from the *fepA* promoter, in accord with the suggestion of Pettis *et al.* (1988). That multicopy plasmids containing only *entD* bacterial DNA are able to complement AN90 is consistent with results obtained with other *ent* genes. The genes *entC*, *E B(G)* and *A* are cotranscribed from a promoter upstream of *entC* (Elkins & Earhart, 1988; Liu *et al.*, 1989; Nahlik *et al.*, 1989) and *entF* is the second gene in a polycistronic message initiated upstream of *fes* (Pettis & McIntosh, 1987). Nonetheless, *entE*, *entB(G)*, *entA* and *entF* are active in complementation tests when individually cloned on multicopy plasmids. The sequencing data suggest that a secondary promoter, albeit a weak one because of unusual spacing between the -35 and -10 regions, exists upstream of *entD*. That a normal promoter is unnecessary for expression of *ent* genes when they are on a multicopy plasmid would also explain why Tn5 mutations in *fepA* are not polar on *entD* in complementation tests.

The significance of the sequence TGATAAGCT upstream of *entD* is unclear; surprisingly, the only other genes reported to be headed by this sequence encode outer-membrane proteins. If this sequence were involved in iron regulation, this could explain the slightly enhanced effects of iron starvation observed in cells with plasmid pPC12. However, the sequence is present on plasmid pPC17LZ and no iron regulation for fusion-protein-directed β -galactosidase activity was detected.

The IS186 sequence found downstream of *entD* is of interest for several reasons: (i) three IS186 sequences are present in *E. coli* RRI (Kothary *et al.*, 1985) and the location of one in AB1515 and its derivatives is now known; (ii) by being flanked on one side by a transposon, the *Ent* gene cluster gains similarity with the aerobactin system, which is flanked on both sides by IS1 (Lawlor & Payne, 1984; McDougall & Neilands, 1984; Perez-Casal & Crosa, 1984); and (iii) should the IS186 sequence be defective, it would suggest a basis for the unusually high frequency and variety of deletion mutations observed in this portion of the genome (Elish *et al.*, 1988).

We thank Drs I. Molineux and J. R. Walker for providing plasmids and strains.

This work was supported by Public Health Service grant AI-22203 from the National Institute of Allergy and Infectious Diseases.

REFERENCES

- ADLER, H. I., FISCHER, W. D., COHEN, A. & HARDIGREE, A. A. (1967). Miniature *Escherichia coli* cells deficient in DNA. *Proceedings of the National Academy of Sciences of the United States of America* **57**, 321-326.
- ARNOW, L. E. (1937). Colorimetric determinations of the components of 3,4-dihydroxyphenylalanine-tyrosine mixtures. *Journal of Biological Chemistry* **118**, 531-537.
- BAGG, A. & NEILANDS, J. B. (1987). Molecular mechanism of regulation of siderophore-mediated iron assimilation. *Microbiological Reviews* **51**, 509-518.
- BERG, C. M., SHAW, K. J., VENDER, H. & BROUCHA-MANKIEWICZ, M. (1979). Physiological characterization of polar Tn5-induced isoleucine-valine auxotrophs in *Escherichia coli* K-12: evidence for an internal promoter in the *ilvOGEDA* operon. *Genetics* **93**, 309-310.
- BIGGEN, M. D., GIBSON, T. J. & HONG, G. F. (1983). Buffer gradient gels and ³⁵S label as an aid to rapid DNA sequence determination. *Proceedings of the National Academy of Sciences of the United States of America* **80**, 3963-3965.
- BIRNBOIM, H. C. & DOLY, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Research* **7**, 1513-1523.
- DEBRUIJN, F. J. & LUPSKI, J. R. (1984). The use of transposon Tn5 mutagenesis in the rapid generation of correlated physical and genetic maps of DNA segments cloned into multicopy plasmids - a review. *Gene* **27**, 131-149.
- CASADABAN, M. J., CHOU, J. & COHEN, S. N. (1980). *In vitro* gene fusions that join an enzymatically active beta-galactosidase segment to amino-terminal fragments of exogenous proteins: *Escherichia coli* plasmid vectors for the detection and cloning of translational initiation signals. *Journal of Bacteriology* **143**, 971-980.
- CHONG, P., HUI, I., LOO, T. & GILLAM, S. (1985). Structural analysis of a new GC-specific insertion element IS186. *FEBS Letters* **192**, 47-52.
- CODERRE, P. E. & EARTHART, C. F. (1984). Characterization of a plasmid carrying the *Escherichia coli* K-12 *entD*, *fepA*, *fes* and *entF* genes. *FEMS Microbiology Letters* **25**, 111-116.
- COHEN, S. N., CHANG, A. C. Y. & HSU, L. (1972). Nonchromosomal antibiotic resistance in bacteria:

- genetic transformation of *Escherichia coli* by R-factor DNA. *Proceedings of the National Academy of Sciences of the United States of America* **69**, 2110–2114.
- DALE, R. M. K., MCCLURE, B. A. & HORCHINS, J. P. (1985). A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondrial 18s rDNA. *Plasmid* **13**, 31–40.
- DUNN, J. J. & STUDIER, F. W. (1983). Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 genetic elements. *Journal of Molecular Biology* **166**, 477–535.
- EARHART, C. F. (1987). Ferrienterobactin transport in *Escherichia coli*. In *Iron Transport in Microbes, Plants, and Animals*, pp. 67–84. Edited by G. Winkelmann, D. vander Helm & J. B. Neilands. Weinheim: VCH Verlagsgesellschaft.
- ELISH, M. E., PIERCE, J. R. & EARHART, C. F. (1988). Biochemical analysis of spontaneous *fepA* mutants of *Escherichia coli*. *Journal of General Microbiology* **134**, 1355–1364.
- ELKINS, M. E. & EARHART, C. F. (1988). An *Escherichia coli* enterobactin cluster gene with sequence homology to *trpE* and *pabB*. *FEMS Microbiology Letters* **56**, 35–40.
- FAIRBANKS, G., STECK, T. L. & WALLACE, D. F. H. (1971). Electrophoretic analysis of the major polypeptides of the human erythrocyte membrane. *Biochemistry* **10**, 2606–2617.
- FLEMING, T. P., NAHLIK, M. S., NEILANDS, J. B. & MCINTOSH, M. A. (1985). Physical and genetic characterization of cloned enterobactin genomic sequences from *Escherichia coli* K-12. *Gene* **34**, 47–54.
- VON GABAIN, A., BELASCO, J. G., SCHOTTEL, J. L., CHANG, A. C. Y. & COHEN, S. N. (1983). Decay of mRNA in *Escherichia coli*: investigation of the fate of specific segments of transcripts. *Proceedings of the National Academy of Sciences of the United States of America* **80**, 653–657.
- GOLD, L. (1988). Posttranscriptional regulatory mechanisms in *Escherichia coli*. *Annual Review of Biochemistry* **57**, 199–233.
- GREENWOOD, K. T. & LUKE, R. K. J. (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. *Biochimica et biophysica acta* **454**, 285–297.
- GREENWOOD, K. T. & LUKE, R. K. J. (1980). Studies on the enzymatic synthesis of enterochelin in *Escherichia coli* K-12, *Salmonella typhimurium* and *Klebsiella pneumoniae*. *Biochimica et biophysica acta* **614**, 185–195.
- GREENWOOD, K. T. & LUKE, R. K. J. (1981). Regulation of enterochelin synthetase in *Escherichia coli* K-12. *Biochimica et biophysica acta* **660**, 371–374.
- HOWARD-FLANDERS, P., SIMSON, E. & THERIOT, L. (1964). A locus that controls filament formation and sensitivity to radiation in *Escherichia coli* K-12. *Genetics* **49**, 237–246.
- KONIGSBERG, W. & GODSON, G. N. (1983). Evidence for use of rare codons in *dnaG* gene and other regulatory genes of *Escherichia coli*. *Proceedings of the National Academy of Sciences of the United States of America* **80**, 687–691.
- KOTHARY, R. K., JONES, D. & CANDIDO, E. P. M. (1985). IS186: an *Escherichia coli* insertion element isolated from a cDNA library. *Journal of Bacteriology* **164**, 957–959.
- LAEMMLI, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature, London* **227**, 680–685.
- LAWLOR, K. M. & PAYNE, S. M. (1984). Aerobactin genes in *Shigella* spp. *Journal of Bacteriology* **160**, 266–272.
- LIU, J., DUNCAN, K. & WALSH, C. T. (1989). Nucleotide sequence of a cluster of *Escherichia coli* enterobactin biosynthetic genes: identification of *entA* and purification of its product 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase. *Journal of Bacteriology* **171**, 791–798.
- LUGTENBERG, B., MEIJERS, J., PETERS, R., VAN DER HOEK, P. & VAN ALPHEN, L. (1975). Electrophoretic resolution of the 'major outer membrane protein' of *Escherichia coli* K-12 into four bands. *FEBS Letters* **58**, 254–258.
- LUNDRIGAN, M. D. & KADNER, R. J. (1986). Nucleotide sequence of the gene for the ferrienterochelin receptor *FepA* in *Escherichia coli*: homology among outer membrane receptors which interact with TonB. *Journal of Biological Chemistry* **261**, 10797–10801.
- MACKIE, G. A. (1981). Nucleotide sequence of the gene for ribosomal protein S20 and its flanking regions. *Journal of Biological Chemistry* **256**, 8177–8182.
- MANIATIS, T., FRITSCH, E. F. & SAMBROOK, J. (1982). *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- MCDONNELL, M. W., SIMON, M. N. & STUDIER, F. W. (1977). Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. *Journal of Molecular Biology* **110**, 119–146.
- MCDUGALL, S. & NEILANDS, J. B. (1984). Plasmid- and chromosome-coded aerobactin synthesis in enteric bacteria: insertion sequences flank operon in plasmid-mediated systems. *Journal of Bacteriology* **159**, 300–305.
- MEAGHER, R. B., TAIT, R. C., BETLACH, M. & BOYER, H. W. (1977). Protein expression in *E. coli* minicells by recombinant plasmids. *Cell* **10**, 521–536.
- MELTON, D. A., KRIEG, P. A., REBAGLIATI, M. R., MANIATIS, T., ZINN, K. & GREEN, M. R. (1984). Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. *Nucleic Acids Research* **12**, 7035–7056.
- MILLER, J. H. (1972). *Experiments in Molecular Genetics*, p. 352. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- NAHLIK, M. S., BRICKMAN, T. J., OZENBERGER, B. A. & MCINTOSH, M. A. (1989). Nucleotide sequence and transcriptional organization of the *Escherichia coli* enterobactin biosynthesis cistrons *entB* and *entA*. *Journal of Bacteriology* **171**, 784–790.
- NEWBURY, S. F., SMITH, N. H., ROBINSON, E. C., HILES, I. D. & HIGGINS, C. F. (1987a). Stabilization of translationally active mRNA by prokaryotic REP sequences. *Cell* **48**, 297–310.
- NEWBURY, S. F., SMITH, N. H. & HIGGINS, C. F. (1987b). Differential mRNA stability controls rela-

- tive gene expression within a polycistronic operon. *Cell* **51**, 1131–1143.
- PEREZ-CASAL, J. F. & CROSA, J. H. (1984). Aerobactin iron uptake sequences in plasmid ColV-K30 are flanked by inverted IS1-like elements and replication regions. *Journal of Bacteriology* **160**, 256–265.
- PETTIS, G. S. & MCINTOSH, M. A. (1987). Molecular characterization of the *Escherichia coli* enterobactin cistron *entF* and its coupled expression to the *fes* gene. *Journal of Bacteriology* **169**, 4154–4162.
- PETTIS, G. S., BRICKMAN, T. J. & MCINTOSH, M. A. (1988). Transcriptional mapping and nucleotide sequence of the *Escherichia coli* *fepA-fes* enterobactin region. *Journal of Biological Chemistry* **263**, 18857–18863.
- POULIS, M. I., SHAW, D. C., CAMPBELL, H. D. & YOUNG, I. G. (1981). *In vitro* synthesis of the respiratory NADH dehydrogenase of *Escherichia coli*. Role of UUG as initiation codon. *Biochemistry* **20**, 4178–4185.
- REDDY, P., PETERKOFKY, A. & MCKENNEY, K. (1985). Translational efficiency of the *Escherichia coli* adenylate cyclase gene: mutating the UUG initiation codon to GUG or AUG results in increased gene expression. *Proceedings of the National Academy of Sciences of the United States of America* **82**, 5656–5660.
- ROY, A., HAZIZA, C. & DANCHIN, A. (1983). Regulation of adenylate cyclase synthesis in *Escherichia coli*: nucleotide sequence of the control region. *EMBO Journal* **2**, 791–797.
- RUPPRECHT, K. R., GORDON, G., LUNDRIGAN, M. D., GAYDA, R. C., MARKOVITZ, A. & EARHART, C. F. (1984). *ompT*: *Escherichia coli* K-12 structural gene for protein a (3b). *Journal of Bacteriology* **153**, 1104–1106.
- SANGER, F., NICKLEN, S. & COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74**, 5463–5468.
- SCHWYN, B. & NEILANDS, J. B. (1987). Universal chemical assay for the detection and determination of siderophores. *Analytical Biochemistry* **160**, 47–56.
- SHARP, P. M. & LI, W. H. (1986). Codon usage in regulatory genes in *Escherichia coli* does not reflect selection for 'rare' codons. *Nucleic Acids Research* **14**, 7737–7749.
- SHINE, J. & DALGARNO, L. (1974). The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proceedings of the National Academy of Sciences of the United States of America* **71**, 1342–1346.
- STUDIER, F. W. & MOFFATT, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of Molecular Biology* **189**, 113–130.
- SUGIMURA, K. & HIGASHI, N. (1988). A novel outer-membrane-associated protease in *Escherichia coli*. *Journal of Bacteriology* **170**, 3650–3654.
- SUGIMURA, K. & NISHIHARA, N. (1988). Purification, characterization and primary structure of *Escherichia coli* protease VII with specificity for paired basic residues: identity of protease VII and OmpT. *Journal of Bacteriology* **170**, 5625–5632.
- TINOCO, I., BORER, P. N., DENGLE, B., LEVINE, M. D., UHLENBECK, O. C., CROTHERS, D. M. & GRALLA, J. (1973). Improved estimation of secondary structure in ribonucleic acids. *Nature New Biology* **246**, 40–41.
- WHITE, B. A. & BANCROFT, F. C. (1982). Cytoplasmic dot hybridization – simple analysis of relative mRNA levels in multiple small cell or tissue samples. *Journal of Biological Chemistry* **257**, 8569–8572.
- WOODROW, G. C., YOUNG, I. G. & GIBSON, F. (1979). Biosynthesis of enterochelin in *Escherichia coli* K-12: separation of the polypeptides coded for by the *entD*, *E*, *F* and *G* genes. *Biochimica et biophysica acta* **582**, 145–153.
- YANISCH-PERRON, C., VIEIRA, J. & MESSING, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp8 and pUC19 vectors. *Gene* **33**, 103–119.
- YIN, K. C., BLINKOWA, A. & WALKER, J. R. (1986). Nucleotide sequence of the *Escherichia coli* gene *dnaZX*. *Nucleic Acids Research* **14**, 6541–6549.