Hankinson, O. (1983) *Somatic Cell Genet. 9*, 497-514.
Hankinson, O., Anderson, R. D., Birren, B. W., Sander, F., Negishi, M., & Nebert, D. W. (1985) *J. Biol. Chem. 260*, 1790-1795.
Hannah, R. R., Lund, J., Poellinger, L., Gillner, M., & Gustafsson, J.-A. (1986) *Eur. J. Biochem. 156*, 237-242.
Henry, E. C., Rucci, G., & Gasiewicz, T. A. (1989) *Biochemistry 28*, 6430-6440.
Johnson, P. F., & McKnight, S. L. (1989) *Annu. Rev. Biochem. 58*, 799-839.
Jones, P. B. C., Galeazzi, D. R., Fisher, J. M., & Whitlock, J. P., Jr. (1985) *Science 227*, 1499-1502.
Karenlampi S. O., Legraverend, C., Gudas, J. M., Carremanzana, N., & Hankinson, O. (1988) *J. Biol. Chem. 263*, 10111-10117.
Kumar, V., & Chambon, P. (1988) *Cell 55*, 145-156.
Laemmli, U. K. (1970) *Nature 227*, 680-685.
Landers, J. P., Piskorska-Pliszczynska, J., Zacharewski, T., Bunce, N. J., & Safe, S. (1989) *J. Biol. Chem. 264*, 18463-18471.
Lin, S.-Y., & Riggs, A. D. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 947-951.
Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Miller, A. G., Israel, D. I. & Whitlock, J. P., Jr. (1983) *J. Biol. Chem. 258*, 3523-3527.
Nakabeppu, Y., Ryder, K., & Nathans, D. (1988) *Cell 55*, 907-915.
Nemoto, T., Mason, G. G. F., Wilhelmsson, A., Cuthill, S., Hapgood, J., Gustafsson, J.-A., & Poellinger, L. (1990) *J. Biol. Chem. 265*, 2269-2277.
Perdew, G. H. (1988) *J. Biol. Chem. 263*, 13082-13805.
Poland, A., & Knutson, J. C. (1982) *Annu. Rev. Pharmacol. Toxicol. 2*, 517-554.
Poland, A., & Glover, E. (1987) *Biochem. Biophys. Res. Commun. 146*, 1439-1449.
Poland, A., & Glover, E. (1988) *Arch. Biochem. Biophys. 261*, 103-111.
Poland, A., Glover, E., Ebetino, F. H., & Kende, A. S. (1986) *J. Biol. Chem. 261*, 6352-6365.
Prokipcak, R. D., & Okey, A. B. (1988) *Arch. Biochem. Biophys. 267*, 811-828.
Roux, E., Strubin, M., Hagenbuchle, O., & Wellauer, P. K. (1989) *Genes Dev. 3*, 1613-1624.
Shen, E. S., & Whitlock, J. P., Jr. (1989) *J. Biol. Chem. 264*, 17754-17758.
Tsai, S. Y., Carlstedt-Duke, J., Weigel, N. L., Dahlmann, K., Gustafsson, J.-A., Tsai, M. H., & O'Malley, B. W. (1988) *Cell 55*, 361-369.
Whitlock, J. P., Jr. (1990) *Annu. Rev. Pharmacol. Toxicol. 30*, 251-277.
Wilhelmsson, A., Wikstrom, A.-C., & Poellinger, L. (1986) *J. Biol. Chem. 261*, 13456-13463.
Wilhelmsson, A., Cuthill, S., Denis, M., Wikstrom, A.-C., Gustafsson, J.-A., & Poellinger, L. (1990) *EMBO J. 9*, 69-76.
Winter, R. B., & von Hipple, P. H. (1981) *Biochemistry 20*, 6948-6960.
Yamamoto, K. R. (1985) *Annu. Rev. Genet. 19*, 209-252.

Biosynthesis of the *Escherichia coli* Siderophore Enterobactin: Sequence of the *entF* Gene, Expression and Purification of EntF, and Analysis of Covalent Phosphopantetheine[†,‡]

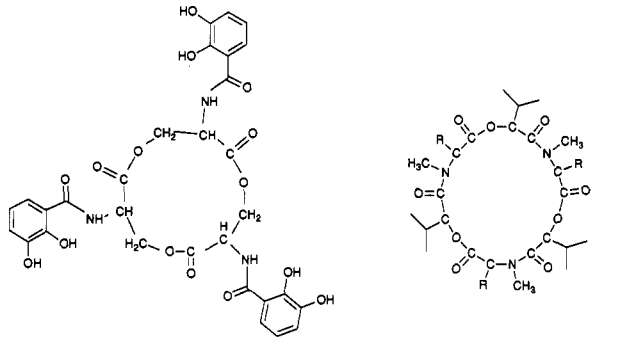
Frank Rusnak, Masahiro Sakaitani, Dale Drueckhammer,[§] Janice Reichert, and Christopher T. Walsh*
Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115
Received September 11, 1990; Revised Manuscript Received December 12, 1990

ABSTRACT: The sequence of the *entF* gene which codes for the serine activating enzyme in enterobactin biosynthesis is reported. The gene encodes a protein with a calculated molecular weight of 142006 and shares homologies with the small subunits of gramicidin S synthetase and tyrocidine synthetase. We have subcloned and overexpressed *entF* in a multicopy plasmid and attempted to demonstrate L-serine-dependent ATP-[^{32}P]PP$_i$ exchange activity and its participation in enterobactin biosynthesis, but the overexpressed enzyme appears to be essentially inactive in crude extract. A partial purification of active EntF from wild-type *Escherichia coli*, however, has confirmed the expected activities of EntF. In a search for possible causes for the low level of activity of the overexpressed enzyme, we have discovered that EntF contains a covalently bound phosphopantetheine cofactor.

Under conditions of iron deprivation, many microorganisms synthesize and secrete low molecular weight compounds termed siderophores which bind ferric ions with high affinity and are used to supply iron for metabolic processes. *Escherichia coli* and other enteric bacteria synthesize the catechol-containing siderophore enterobactin, a macrocyclic trilactone comprised of three molecules each of 2,3-dihydroxybenzoate (2,3-DHB) and L-serine (Figure 1A). Much is now known about the intricate metabolic machinery of *E. coli* involved in enterobactin biosynthesis, uptake, and processing. Thus, the *ent* genes encode the biosynthetic enzymes (Young et al., 1971; Luke & Gibson, 1971; Woodrow et al., 1975a,b), the *fep* genes are required for the uptake of the ferric enterobactin complex (Pierce et al., 1983; Pierce & Earhart, 1986;

[†] Supported by NIH Grant GM20011 (C.T.W.) and NIH Postdoctoral Fellowships GM12806 (F.R.), GM12014 (D.D.), and GM13804 (J.R.).
[‡] The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J05325.
* Author to whom correspondence should be addressed.
[§] Present address: Department of Chemistry, Stanford University, Stanford, CA 94305.

0006-2960/91/0430-2916$02.50/0



A. Enterobactin

B. Enniatin A: R = -CH(CH$_3$)CH$_2$CH$_3$
Enniatin B: R = -CH(CH$_3$)$_2$
Enniatin C: R = -CH$_2$CH(CH$_3$)$_2$

FIGURE 1: Structures of enterobactin (A) and the enniatin family of depsipeptides (B) are shown. Each of these cyclic molecules contain three amide and three ester bonds.

Lundrigan & Kadner, 1986; Ozenberger et al., 1987; Elkins & Earhart, 1989), and the *fes* gene product, an esterase, hydrolyzes the macrocycle for iron release (Langman et al., 1972). These *ent*, *fep*, and *fes* genes are clustered in a 24-kilobase region around minute 13 on the *E. coli* chromosome, and they are grouped in distinct transcriptional units, each under control of iron and the gene regulatory Fur protein (Coderre & Earhart, 1984; Nahlik et al., 1987; Pettis et al., 1987, 1988; Elkins & Earhart, 1988, 1989; Ozenberger et al., 1989).

Our interest in the biosynthesis of enterobactin initially focused on the reactions which convert chorismic acid to 2,3-dihydroxybenzoate. As a result, we have subcloned, expressed, and purified the three proteins which bring about this process and have initiated mechanistic studies on each of these enzymes: isochorismate synthase, the product of the *entC* gene, which interconverts chorismate and isochorismate (Liu et al., 1990); isochorismatase (EntB), which hydrolyzes the enol pyruvyl ether side chain of isochorismate to yield 2,3-dihydro-2,3-dihydroxybenzoate (Rusnak et al., 1990); and the dehydrogenase product of the *entA* gene, which produces the catechol product in an oxidation step utilizing NAD$^+$ (Liu et al., 1989; Sakaitani et al., 1990).

In the second half of the pathway, the macrocycle is assembled with the formation of three amide and three ester bonds to make the head-to-tail serine trilactone with each serine N-acylated by 2,3-dihydroxybenzoate (Figure 1A). Prior studies identified four *ent* genes for this process, *entD*, *entE*, *entF*, and *entG*, and in fact, the proteins which correspond to these genes have been partially resolved (Bryce & Brot, 1971; Greenwood & Luke, 1976, 1980; Woodrow et al., 1979). These early investigations revealed that the EntE and EntF proteins activated the carboxylate groups of 2,3-DHB and L-serine to the corresponding acyl- and aminoacyl-adenylates, respectively, and that a thiol group on EntF apparently reacted further to bring about the covalent attachment of serine as a thioester. Furthermore, it was postulated that the amino group of this thioester-linked serine reacted with the acyladenylate of 2,3-DHB to form (2,3-dihydroxybenzoyl)serine which remained covalently attached to EntF (Bryce & Brot, 1971). Recently, the EntE protein has also been subcloned, expressed, and purified in these laboratories, and it was verified that EntE does indeed activate 2,3-DHB to the acyladenylate (Rusnak et al., 1990). The roles of EntD and EntG remain obscure although they play an essential role in the final assembly process. The recent determination of the DNA sequence of the *entD* gene has suggested that EntD may be membrane-associated (Armstrong et al., 1989; Coderre & Earhart, 1989). Furthermore, the *entG* mutation has been mapped to the 3′ region of the *entB* gene, suggesting possibly that this activity may reside on the EntB polypeptide (Nahlik et al., 1987). In fact, antibodies raised against EntB inhibit enterobactin synthesis in vitro (Staab & Earhart, 1990).

There are several noteworthy similarities between enterobactin biosynthesis and the assembly of cyclic peptide antibiotics such as gramicidin S and tyrocidine as well as cyclic lactones and depsipeptides (Roskoski et al., 1970a,b; Gevers et al., 1968, 1969; Zocher et al., 1982; Keller, 1987). For example, the depsipeptide family of enniatins resemble enterobactin in that they also contain three amide and three ester bonds in a cyclic structure (Figure 1B) (Zocher et al., 1982). More importantly, the assembly of these cyclic compounds is known to progress via a nonribosomal protein template mechanism with the initial formation of aminoacyladenylates and the subsequent transfer of the activated amino acid to a thiol group on the enzyme.

In this study, we focus on the *entF* gene, which resides in an iron-regulated transcript downstream from the *fes* gene product. Earlier DNA sequencing of this region proceeded approximately 250 base pairs into the 5′ end of *entF*, estimated to encode a polypeptide of ≈160 kDa by SDS–PAGE gel electrophoresis (Pettis & McIntosh, 1987). We report here the DNA sequence of *entF* (Figure 2), the last of the known *ent* genes to be sequenced. We have also subcloned the *entF* gene behind the T7 promoter, expressed the protein in *E. coli*, purified the protein to ≈80% homogeneity, and confirmed its identity by N-terminal sequence determination. EntF purified from an iron-depleted *E. coli* K12 strain catalyzes the L-serine-dependent exchange of [^{32}P]pyrophosphate into ATP and also complements the *E. coli entF* mutant MK1 (McIntosh, personal communication) in an in vitro assay measuring enterobactin production. In efforts to further characterize possible catalytic functions of EntF, we have determined that EntF incorporates β-[^{3}H]alanine and contains a covalently attached phosphopantetheine group.

Materials and Methods

Materials. α-[^{35}S]Thio-dATP (500 Ci/mmol), disodium hydrogen [^{32}P]pyrophosphate (60 mCi/mmol), and β-[3-^{3}H]alanine (87 Ci/mmol) were from NEN DuPont, Wilmington, DE. [3-^{14}C]Serine (56 Ci/mol) was from Amersham, Arlington Heights, IL. Adenosine 5′-triphosphate, β-alanine, dithiothreitol, and alkaline phosphatase were purchased from Sigma Chemical Co., St. Louis, MO. The compound 2,3-dihydroxybenzoate was from Aldrich (Milwaukee, WI). Restriction enzymes and bacteriophage cloning vectors M13mp18 and M13mp19 were from New England Biolabs, Beverly, MA, and were used according to the manufacturers' instructions. T4 DNA ligase and ligase 10× buffer were from IBI, New Haven, CT. Custom oligonucleotides were synthesized by Dr. Alex Nussbaum, Harvard Medical School. Silica gel H plates were purchased from Alltech Associates, Inc. Applied Science Labs.

Growth Media and Conditions. LB broth (1% bactotryptone, 1% NaCl, and 0.5% bacto yeast extract) and LB agar (LB broth plus 1.5% bactoagar) were used for culture growth and maintenance. When necessary, the antibiotics ampicillin and tetracycline were supplied at 50 and 12.5 μg/mL, respectively.

For the partial purification of EntF from wild-type *E. coli*, *E. coli* K12 were grown in low-iron media: (per liter) Na_2HPO_4, 8.6 g; KH_2PO_4, 5.6 g; $(NH_4)_2SO_4$, 1.0 g; $MgSO_4 \cdot 7H_2O$, 0.1 g; $Ca(NO_3)_2 \cdot 4H_2O$, 14.3 mg; thiamin hydro-

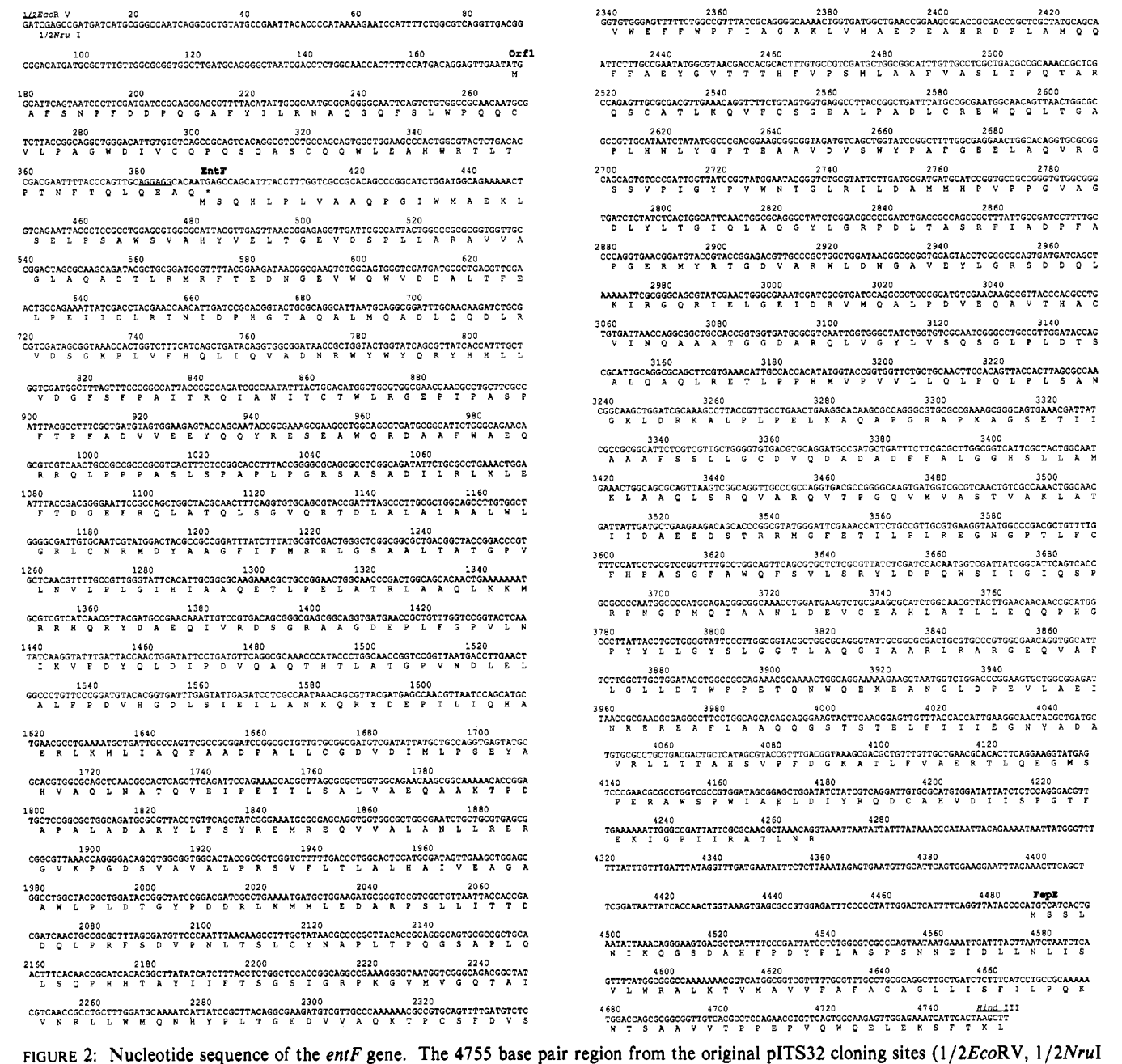


FIGURE 2: Nucleotide sequence of the *entF* gene. The 4755 base pair region from the original pITS32 cloning sites (1/2*Eco*RV, 1/2*Nru*I to *Hin*dIII) is shown. The amino acid assignments occur under the middle base of each codon in the single-letter code. The Shine–Dalgarno sequence of *entF* is underlined. An additional open reading frame is found about 200 base pairs downstream of *entF* and represents the amino terminus of *fepE* (McIntosh, personal communication).

chloride, 0.2 mg; glucose, 10 g; 2,2′-dipyridyl, 31.2 mg (Brot et al., 1966).

In vivo labeling of pantetheine-containing proteins with β-[^{3}H]alanine was carried out by inoculating 5 mL of Dex-E-B1-met media (Vogel & Bonner, 1956) containing 0.4% α-D-glucose, 10^{-4}% thiamine, and 0.002% methionine with a culture of *E. coli* SJ16 and incubating at 37 °C for 12–14 h after which the faintly turbid β-alanine-starved culture was diluted 100-fold into 1 mL of Dex-E-B1-met media containing β-[^{3}H]alanine at a final concentration and specific activity of 0.55 μM and 87.2 Ci/mmol, respectively. The cells were grown for an additional 16 h at 37 °C before harvesting.

Large-Scale Preparation and Purification of β-[^{3}H]-Alanine-Labeled EntF. A 5-mL culture of *E. coli* SJ16, starved for β-alanine as described above, was diluted 100-fold into 4 mL of Dex-E-B1-met media containing β-[^{3}H]alanine and grown at 37 °C for 16 h after which it was used to inoculate two flasks each containing 200 mL of Dex-E-B1-met media containing β-[^{3}H]alanine at a final concentration and specific activity of 0.55 μM and 43.6 Ci/mmol, respectively. The cells were grown for 6 h at 37 °C after which 2,2′-dipyridyl was added to a final concentration of 200 μM and the cells were incubated at 37 °C for an additional 18 h. The cells were harvested by centrifugation and washed once with 100 mM Tris-HCl, pH 8.0, and crude extract and ammonium sulfate fractionations were prepared as described below for the purification of EntF from *E. coli* K12 cells. The 50% $(NH_4)_2SO_4$ pellet was dissolved in 1 mL of buffer containing 50 mM Tris-HCl, pH 8, and 5 mM dithiothreitol and desalted by passage through a NAP-10 column equilibrated with the same buffer. The sample was centrifuged to remove precipitated protein and then applied in aliquots to a Mono-Q HR5/5 (Pharmacia) column. After the column was washed with 10 mL of 50 mM Tris-HCl, pH 8, containing 5 mM dithiothreitol,

the protein was eluted from the column with a 30-mL gradient from 0 to 1 M KCl in the same buffer. Fractions were electrophoresed on 13% SDS–PAGE gels as previously described (Laemmli, 1970). The gels were soaked in Amplify (Amersham), then dried, and autoradiographed. Radiolabeled proteins, EntF included, could clearly be seen after an overnight exposure (Kodak X-OMAT AR film). The fractions containing radiolabeled EntF were combined, concentrated on a PM30 Centricon (Amicon), and chromatographed on a 2.6 cm × 60 cm gel filtration column of Sephadex G200 with a flow rate of 2.25 mL/h. Radiolabeled EntF eluted from the column at about 150 mL. After gel filtration, EntF (20 μCi) was the only protein which contained β-[^{3}H]alanine as judged by autoradiography of fractions electrophoresed on SDS–PAGE gels.

Identification of 4′-Phosphopantetheine Released from EntF. (*a*) *Hydrolysis.* The hydrolysis of 4′-phosphopantetheine was performed on the partially purified enzyme mixture containing EntF as the only radiolabeled protein. The enzyme mixture (10 μL; 20 000 dpm) was heated at 70 °C for 1 h in 200 μL of 50 mM NaOH. The solution was neutralized by the addition of 100 μL of 100 mM HCl. After dilution with 700 μL of 200 mM Tris-HCl, pH 6.8, the solution was concentrated on a PM 30 Centricon (Amicon). The enzyme-containing fraction was washed once with 200 μM Tris-HCl, pH 6.8. The radiolabel was found to have been quantitatively removed from the enzyme by this procedure.

(*b*) *Chromatography.* The fraction containing 4′-[^{3}H]-phosphopantetheine was desalted by chromatography on a 0.5 cm × 3.5 cm DEAE-cellulose (DE 52, Whatman) column equilibrated with 100 mM ammonium acetate in 10% 1-propanol. After the column was washed with 10 mL of water, the 4′-[^{3}H]phosphopantetheine was eluted from the column with 10 mL of 0.1 M ammonium acetate in 10% 1-propanol. The ^{3}H-containing fractions were combined and lyophilized. A sample of the residue was then dissolved in 10 μL of water and analyzed by thin-layer chromatography.

(*c*) *Treatment with Alkaline Phosphatase.* After chromatography, the remaining ^{3}H-labeled sample was treated with 0.3 units of alkaline phosphatase (bacterial; suspension in 2.5 M ammonium sulfate solution) in 97 μL of 100 mM Tris-HCl, pH 8.0. The solution was incubated at 37 °C for 1 h and then lyophilized, and the residue was extracted with 50 μL of ethanol. A sample of the extract was analyzed by thin-layer chromatography.

(*d*) *Thin-Layer Chromatography.* The 4′-[^{3}H]phosphopantetheine [from (b) and (c) above] and pantetheine derived from EntF were identified by thin-layer chromatography on silca gel H plates (20 cm × 20 cm, 250 μm; activated by heating at 110 °C for 30 min prior to use) developed to 14 cm in 1-butanol/acetic acid/water (5:2:4). Sections (0.5 cm) were scraped into scintillation vials containing 5 mL of scintillation fluid and counted to locate the position of radiolabeled compounds. Comparison was then made with known R_f values (Jackowski & Rock, 1981) for 4′-phosphopantetheine and pantetheine.

Bacterial Strains and Plasmids. E. coli strain XL1Blue [*rec*A1, *end*A1, *gyr*A96, *thi*, *hsd*R17 (r_{k^-}, m_{k^-}), *sup*E44, *rel*A1, λ$^-$, (*lac*), {F′, *pro*AB, *lacI*^{q}, lacZΔM15, Tn10 (tet^R)}] from Stratagene was used to maintain all M13 recombinant clones for nucleotide sequencing. The *panD* mutant SJ16 (Jackowski & Rock, 1981) was a gift from Dr. Eugene Kennedy, Harvard Medical School. *E. coli* MK1 (*entF*, *mini-kan*-1), a derivative of AB1515 (Pettis & McIntosh, 1987), and pITS32, a recombinant plasmid harboring the *entF* gene cloned into the *Nru*I/*Hin*dIII sites of pBR328, were provided by Dr. M. A. McIntosh of the University of Missouri—Columbia. The plasmid pJS4700 containing the *entF* gene under the control of the T7 promoter was provided by Drs. Janet Staab and Charles Earhart of the University of Texas at Austin. JM109(DE3), a λ-lysogen containing the gene for bacteriophage T7 RNA polymerase under control of the *lac* promoter, was purchased from Promega, Madison, WI, and is similar to the strain BL21(DE3) as developed by Studier and Moffatt (1986).

Production of Sequencing Clones. The recombinant clones mFRF1 and mFRF2, which contain the 5′ region of the *entF* gene in opposite orientations in M13mp18, were constructed by ligating the 1.9-kb *Bam*HI insert from pITS32 into *Bam*HI-digested M13mp18. Ligation of the 3.1-kb *Bam*HI–*Hin*dIII fragment from pITS32 containing the 3′ end of the *entF* gene into M13mp19 created the recombinant clone mFRF4. The recombinant clone mFRF5 containing the 3′end of *entF* in the opposite orientation from mFRF4 was prepared in the following manner. The plasmid pITS32 was digested with *Hin*dIII and purified by phenol/chloroform extraction followed by ethanol precipitation. The overhanging 5′ ends of *Hin*dIII-digested pITS32 were filled in with 4 units of Klenow in the presence of 2.5 mM aliquots of the four deoxynucleotide triphosphates according to Manniatis et al. (1982). After phosphorylated *Eco*RI linkers were added, the entire reaction was digested with *Eco*RI and *Bam*HI and electrophoresed in low melting point agarose. The 3.1-kb fragment containing the 3′ region of the *entF* gene was cut out from the agarose gel, purified by using an Elutip column (Schleicher & Schuell) and ligated into *Eco*RI–*Bam*HI-digested M13mp19.

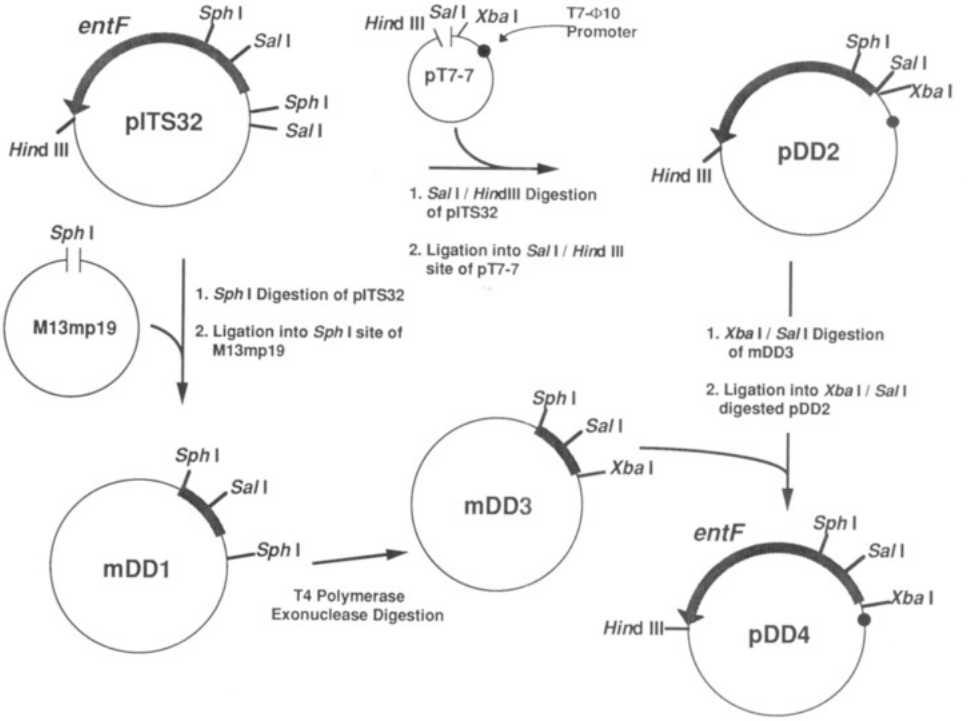
The clone mFRF1EH, generated by ligating the 1.5-kb *Eco*RI/*Hpa*I fragment of pITS32 into *Eco*RI/*Sma*I-digested M13mp19, was sequenced with a synthetic oligonucleotide primer to verify continuity across the *Bam*HI restriction site common to mFRF1, mFRF2, mFRF4, and mFRF5.

Sets of overlapping deletion clones from mFRF1, mFRF2, mFRF4, and mFRF5 were obtained according to the method of Dale et al. (1985) using the Cyclone I biosystem kit from IBI. When required, oligonucleotides were synthesized for use as sequencing primers. DNA sequence generated from these clones was analyzed by using the program DNAstar (DNAstar, Inc., Madison, WI) running on a Compaq Deskpro 386 personal computer. The NBIF protein data bank (version 23.0) and GenBank DNA sequence library (version 62.0) were searched for sequence homologies with the protein and DNA sequences of EntF by using the program Eugene (Baylor College of Medicine, Houston, TX) running on a Sun 3/60 computer at Harvard Medical School.

DNA Sequencing. The nucleotide sequence of both strands was obtained by using the chain termination method of Sanger (Sanger et al., 1977) according to the instructions provided in the Sequenase kit (United States Biochemical Corp., Cleveland, OH) using α-[^{35}S]thio-dATP as the label. The reaction mixture was electrophoresed on 6% polyacrylamide gels. After the gels were dried, the acrylamide was subjected to autoradiography with Kodak X-OMAT AR film. Compressions and sequence mismatches between opposite strands were resolved by using dITP in place of dGTP according to the instructions provided in the Sequenase kit.

Subcloning of the entF Gene into pT7-7. Subcloning of the *entF* gene was carried out in three steps (Scheme I). First (step a, below), the *Sph*I fragment containing the 5′ portion of *entF* was cloned into M13mp19 to give mDD1 which was

Scheme I: Subcloning Strategy To Generate the Overexpression Vector pDD4 for EntF[a]



[a] The 5′ *Sph*I fragment of *entF* was cloned into the *Sph*I site of M13mp19 to give the clone mDD1. After linearization of single-stranded template of mDD1, >750 base pairs were removed up to the Shine–Dalgarno sequence for *entF* using the 3′–5′ exonuclease activity of T4 DNA polymerase to yield after religation the clone mDD3, which contains a new *Xba*I site at the 5′ end of *entF*. The *Xba*I/*Sal*I 5′ fragment of *entF* was then cloned into the *Xba*I/*Sal*I site of pDD2, which contains the 3′ end of *entF* cloned into pT7-7 behind the T7-Φ10 promoter.

used for the T4 exonuclease digestion, eventually producing the clone mDD3 which has >750 base pairs of DNA removed at the 5′ end of *entF* and contains a unique *Xba*I site just upstream of the Shine–Dalgarno sequence of *entF*. At the same time, the *Sal*I/*Hin*dIII fragment containing the 3′ end of *entF* was cloned into the *Sal*I/*Hin*dIII sites of pT7-7 and the vector pDD2 isolated (step b, below). Lastly, after digestion of the double-stranded form of mDD3 with *Xba*I/*Sal*I, the 0.8-kb fragment containing the 5′ end of *entF* was ligated into *Xba*I/*Sal*I-digested pDD2, to give the complete *entF* gene in the vector pDD4 (step c, below). The relevant steps in this process are as follows:

(*a*) *Deletion of Uncoding DNA at the 5′ End of entF.* As shown in Scheme I, the plasmid pITS32 was digested with *Sph*I, and the 2.0-kb fragment containing the 5′ end of *entF* was purified in low melting point agarose. The fragment was excised, melted, mixed with *Sph*I-digested M13mp19, T4 DNA ligase, and ATP, and ligated at 22 °C for 24 h. The ligation mixture was transformed into *E. coli* XL1-Blue, and clear plaques were isolated. Preparation of single-stranded DNA followed by sequencing showed the presence of the insert in the proper orientation.

Single-stranded mDD1 was primed with the 25-mer oligonucleotide (5′-ATCCTCTAGAGTCGACCTGCAGGGG-3′) and linearized with *Xba*I at 50 °C for 1.5 h. After heating to 65 °C for 10 min to inactivate *Xba*I, the reaction was cooled to room temperature, and 4 nmol of the 28-mer oligonucleotide (5′-TCTAGAGGAGGCACAATGAGCCAGCATT-3′), corresponding to the first 13 base pairs of the *entF* gene, 9 upstream residues, and 6 overhanging residues of an *Xba*I site, was added. This was heated to 65 °C for 10 min and allowed to cool slowly to 35 °C over 30 min. Exonuclease digestion of single-stranded DNA in the 3′ to 5′ direction was carried out using T4 DNA polymerase (20 units) at 37 °C for 30 min in the presence of 0.2 mM dTTP. Deletion of single-stranded DNA occurred up to the region of double-stranded DNA, at which point the exonuclease activity was compensated by the polymerase activity in the presence of dTTP, the nucleotide corresponding to the first base of the double-stranded region. T4 polymerase was then inactivated by heating for 10 min at 65 °C, and the reaction mixture was allowed to cool slowly to room temperature over 2 h. Ligation was accomplished by using T4 DNA ligase plus ATP to 1 mM at 24 °C for 1 h. The ligation mixture was transformed into *E. coli* XL1-Blue, and clear plaques were isolated. These were screened first on an agarose gel to check for deletion of ≈750 base pairs, and further checked by DNA sequencing. The proper clone, mDD3, with ≈750 base pairs deleted and an *Xba*I site inserted into the front of the Shine–Dalgarno sequence of *entF* was used subsequently in step c described below.

(*b*) *Cloning of the 3′ End of entF.* The plasmid pITS32 was digested with *Hin*dIII and *Sal*I and, the 3.5-kb 3′ fragment of *entF* was purified in low melting point agarose. After excision, the gel fragment was melted and mixed with *Hin*dIII/*Sal*I-digested pT7-7 and ligated as described in (a) for 24 h at 22 °C. The ligation mixture was transformed into *E. coli* XL1-Blue, and ampicillin-resistant colonies were further screened by plasmid preparation and restriction digest to identify the correct clone, pDD2 (Scheme I).

(*c*) *Cloning of the 5′ End of the entF into pDD2.* Double-stranded mDD3 was isolated and digested with *Xba*I/*Sal*I and the 0.8-kb 5′ *entF* fragment purified on LMP agarose. The fragment was excised, melted, and mixed with *Xba*I/*Sal*I-digested pDD2 and ligated as described above in (a) for 18 h at 22 °C. The ligation mixture was transformed into *E. coli* XL1-Blue, and ampicillin-resistant colonies were isolated and screened by plasmid preparation and restriction digest. The correct clone containing the entire *entF* gene with a new *Xba*I site cloned into the vector pT7-7 is called pDD4 (Scheme I).

Purification of EntF from E. coli K12. (*a*) *Crude Cell Extract.* Four 1.5-L cultures of *E. coli* K12 were grown in low-iron media containing 0.2 mM 2,2′-dipyridyl (Brot et al., 1966). Eighteen hours later, the cells (6 g) were harvested by centrifugation, washed with 100 mM Tris-HCl, pH 8.0, resuspended in 24 mL of buffer containing 10 mM $MgCl_2$, 5.0 mM dithiothreitol, and 100 mM Tris-HCl, pH 8.0, and lysed by two passages through a French press operating at 12 000 psi at the orifice. Cell debris was removed by centrifugation (20 min, 10000*g*), and nucleic acids were precipitated by the addition of one-fifth volume of 2% protamine sulfate to yield crude extract.

(*b*) *Ammonium Sulfate Fractionation.* To the crude extract was added ammonium sulfate slowly with stirring on ice to give a final concentration of 50% saturation. The precipitate, containing EntF, was collected by centrifugation (20 min, 10000*g*) and dissolved in 8.0 mL of 25 mM Tris-HCl buffer, pH 8.0, containing 10 mM $MgCl_2$ and 5 mM dithiothreitol (buffer A).

(*c*) *Anion-Exchange Chromatography.* A protein fraction after ammonium sulfate fractionation was dialyzed against buffer A and applied to a prepacked Mono-Q HR16/10 column (Pharmacia) at 4 °C and washed with 100 mL of buffer A. EntF eluted from the column with a linear 300-mL gradient of 0–1 M NaCl in buffer A at about 0.4 M NaCl. The flow rate under these conditions was 6.0 mL/min; 6.0-mL fractions were taken and assayed for protein by measuring the absorbance at 280 nm. Fractions of the major peaks which eluted were assayed for the formation of (2,3-dihydroxybenzoyl)serine (DBS)-containing compounds (vide infra). A single peak which eluted after the void volume was found to synthesize DBS derivatives in the presence of 2,3-dihydroxybenzoate. The fractions from this peak were pooled for the next purification step.

(*d*) *Gel Filtration Chromatography.* The fractions active in the formation of DBS derivatives after anion-exchange chromatography were next applied to a prepacked Superose 6 prep-grade (Pharmacia) column for gel filtration chromatography at 4 °C. The protein was eluted with buffer A at a flow rate of 0.6 mL/min; 6.0-mL fractions were collected. The fractions containing DBS synthase activity were pooled for the next purification step.

(*e*) *Hydrophobic Chromatography.* The pooled fractions after gel filtration chromatography were dialyzed against buffer A containing 1.7 M $(NH_4)_2SO_4$ and applied to a prepacked phenyl-Superose HR 5/5 (Pharmacia) column at 4 °C which was washed first with 10 mL of the same buffer and then with a linear gradient of decreasing $(NH_4)_2SO_4$ concentration from 1.7 to 0 M. EntF eluted from the column in the wash at about 1.7 M. The flow rate under these conditions was 0.5 mL/min.

(*f*) *Analysis of Proteins.* Polyacrylamide gel electrophoresis under denaturing conditions was performed on 7.5% gels as previously described (Laemmli, 1970). Molecular weight standards were from Bethesda Research Laboratories, Bethesda, MD. Protein concentrations were determined by using the Bradford method (Bradford, 1976) using bovine serum albumin as a standard.

(*g*) *Storage of Active Enzyme.* It was found that under dilute conditions (i.e., after anion-exchange chromatography) and/or storage at 4 °C, purified fractions of EntF lost activity within a few days. In order to stabilize the enzyme, the enzyme samples were concentrated with an Amicon filtration cell on a PM-30 membrane followed by dialysis in Spectra/Por 2 dialysis tubing (American Scientific Products Inc.) against buffer A containing 50% glycerol over 20 h. The resulting enzyme solution had an increased stability over a few weeks when stored at −20 °C.

(*h*) *Amino-Terminal Sequence Determination.* Twelve residues of the amino terminus of EntF were determined according to the Edman degradation procedure by Dr. William Lane at the Harvard Microchemistry Facility, Harvard University, Cambridge, MA.

Purification of EntF from E. coli JM109(DE3)/pDD4. (*a*) *Crude Cell Extract.* Four 1.5-L cultures of *E. coli* JM109(DE3)/pDD4 were grown in LB broth containing ampicillin at 37 °C until the OD_{595} reached 0.7, at which point IPTG was added to a final concentration of 1 mM. Three hours later, the cells (6 g) were harvested by centrifugation, and crude extract was prepared by the same procedure described for preparation of *E. coli* K12 crude cell extract.

Protein fractions from each purification step were assayed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis to determine the presence of EntF.

(*b*) *Gel Filtration Chromatography.* The crude cell extract was applied to a 2.5 by 80 cm column of G-150 at 4 °C. The protein was eluted with buffer A at a flow rate of 20 mL/h.

(*c*) *Anion-Exchange Chromatography.* The fractions containing EntF after gel filtration chromatography were applied to a 2.5 by 30 cm column of DEAE-Sepharose CL-6B at 4 °C and washed with 100 mL of buffer A. EntF eluted from the column with a linear 1000-mL gradient of 0–1 M KCl in buffer A. The flow rate under these conditions was 20 mL/h.

(*d*) *Affinity Chromatography.* The pooled fractions after anion-exchange chromatography were dialyzed against buffer A and then applied to a 1.5 by 10 cm column of Dyemetrex Blue A–agarose at 4 °C which was washed first with 40 mL of the same buffer and then with a linear gradient of potassium chloride concentration from 0 to 1.0 M.

(*e*) *Anion-Exchange Chromatography.* A protein fraction after affinity chromatography was dialyzed against buffer A and applied to a prepacked Mono-Q HR5/10 column (Pharmacia) at 4 °C and washed with 8 mL of buffer A. EntF eluted from the column with a linear 30-mL gradient of 0–1 M NaCl in buffer A. The flow rate under these conditions was 0.5 mL/min.

Assays for Ent F. (*a*) *ATP-[^{32}P]Pyrophosphate Exchange Assay.* The exchange of [^{32}P]pyrophosphate into ATP was carried out as previously described (Bryce & Brot, 1972; Greenwood & Luke, 1976) with minor modifications. The assay mixture contained 25 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$, 5.0 mM dithiothreitol, 5.0 mM ATP, 2.0 mM disodium hydrogen [^{32}P]pyrophosphate, 1.0 mM L-serine, and enough enzyme to give detectable counts in a final volume of 0.1 mL. After incubation at 37 °C for 15 min, the reaction was quenched by the addition of 0.4 mL of 0.5 M perchloric acid; 0.4 mL of 0.1 M tetrasodium pyrophosphate was added, and ATP was adsorbed by adding 0.2 mL of a 4% charcoal suspension in water. After being washed 4 times with 1.0 mL of water each time, the charcoal was resuspended in 1.0 mL of water and transferred to a scintillation vial. The Eppendorf tube was washed with an additional 1.0 mL of water, and this was also transferred to the scintillation vial. To the combined 2.0 mL of charcoal suspension was added 5.0 mL of scintillation fluid prior to counting.

(*b*) *Synthesis of (Dihydroxybenzoyl)serine Derivatives.* A 1.0-L culture of MK1 [*entF*$^-$ strain, a derivative of AB1515 (Pettis & McIntosh, 1987)] was grown in LB media in the presence of 2,2′-dipyridyl (0.2 mM) at 37 °C until the absorbance at 595 nm reached 1.2. The cells (1 g) were har-

vested by centrifugation, washed in 100 mM Tris-HCl, pH 8.0, resuspended in 4 mL of buffer containing 10 mM $MgCl_2$, 5.0 mM dithiothreitol, and 100 mM Tris-HCl, pH 8.0, and lysed by two passages through a French pressure cell operating at 12000 psi at the orifice. Cell debris was removed by centrifugation (20 min, 10000*g*), and nucleic acids were precipitated by the addition of one-fifth volume of 2% protamine sulfate. The supernatant obtained after centrifugation (20 min, 10000*g*) was dialyzed in Spectra/Por 2 dialysis tubing against buffer A containing 50% glycerol over 20 h. The resulting solution was used to assay for the synthesis of DBS-containing compounds. The incubation mixture employed to measure the synthesis of DBS-containing compounds contained, in a total volume of 0.2 mL, 25 mM Tris-HCl, pH 8.0, 10 mM $MgCl_2$, 5.0 mM dithiothreitol, 5.0 mM ATP, 0.5 mM 2,3-dihydroxybenzoate, 0.5 mM [3-^{14}C]serine (0.5 Ci/mol), fractions containing EntF, and crude extract from low-iron-induced MK1 (100–200 ng of protein). After incubation at 37 °C for 30 min, the reaction was stopped by the addition of 0.8 mL of 0.05 M HCl. DBS-containing compounds were extracted with 4.0 mL of ethyl acetate, washed with 1.0 mL of 0.05 M HCl, and dried over anhydrous Na_2SO_4. A 2.0-mL sample of this was transferred to a scintillation vial, and 5.0 mL of scintillation fluid was added prior to counting.

(*c*) *Detection of Enzyme–Substrate Complex.* The methods for observing substrate complexes with EntF were identical with those described previously for monitoring the formation of analogous complexes with EntE (Rusnak et al., 1989), the activating enzyme for 2,3-dihydroxybenzoate. Acid precipitation of enzyme reaction mixtures was performed as described previously (Rusnak et al., 1989).

Results

Nucleotide Sequence of entF. The *entF* mutant, defective in the biosynthesis of enterobactin but not 2,3-dihydroxybenzoate, was mapped near minute 13 on the *E. coli* chromosome in a transcript just downstream of the *fes* gene (Luke & Gibson, 1971; Codere & Earhart, 1984). This mutant was shown to be defective in a L-serine-dependent ATP–[^{32}P]PP_i exchange assay (Greenwood & Luke, 1976, 1980; Woodrow et al., 1979). Using in vivo complementation studies, Pettis and McIntosh isolated the plasmid pITS32 which complemented the *entF* mutant and succeeded in obtaining a partial sequence of the 5′ region of the gene (Pettis & McIntosh, 1987).

By use of the Dale deletion method to generate a series of overlapping clones of both strands of the DNA from four clones derived from pITS32, the complete *entF* gene has now been sequenced. One continuous open reading frame was found which encodes a polypeptide of 1293 amino acids with a predicted molecular weight of 142006, which agrees well with a molecular weight of 130K–140K estimated from SDS–PAGE gels of the partially purified EntF protein (vida infra). A consensus ribosome binding site is situated five bases upstream of the translation initiation site (Shine & Dalgarno, 1974). A sequence of 33 base pairs immediately following the translation termination site for EntF has the potential to form a stem and loop structure and may represent a ρ-independent transcriptional terminator. Further downstream of this terminator, another open reading frame which represents the gene for *fepE* extends beyond the *Hin*dIII cloning site (McIntosh, personal communication).

A nucleotide binding motif (Walker et al., 1982; Wierenga et al., 1986; Higgins et al., 1986) which might represent the ATP binding site for EntF was not found nor were there any sequence homologies with the purported "HIGH" ATP binding domain of the aminoacyl-tRNA synthetases (Burbaum et al., 1990). In addition, no obvious primary sequence repeats were found.

```
EntF    VDVSWYPAFGEELAQVRGSSVPIGYPVWNTGLRILDAMMHPVPPGVAGDLYLTGIQLAQG
GrsA    ATTWVATK.....ETIG.H-----A-IQ--QIY-V-ENLQLKSV-E--E-CIG-EG--R-
TycA    ATIWEA-S.....N-LSVQ-----K-IQ--HIY-VNEDLQLL-TADE-E-CIG-VG--R-

EntF    YLGRPDLTASRFIADPFAPGERMYRTGDVARWLDNGAVEYLGRSDDQLKIRGQRIELGEI
GrsA    -WK--E--SQK-VDN--V---KL-K---Q----SD-NI-----I-N-V----H-V--E-V
TycA    -WN------EK-VDN--V---K------L-K--TD-TI-F---I-H-V----H-------

EntF    DRVMQALPDVEQAVTHACVINQAAATGGDARQLVGYLVSQSGLPLDTSALQAQLRETLPP
GrsA    ESILLKHMYISETAVSVHKDH-EQPY.....-CA-F--EKHI--EQ..-RQFSS-E--T
TycA    ES-LL-HEHITE--VI-REDQH-GQY.....-CA-YI--QEATPAQ..-RDYAAQK--A

EntF    HMVPVVLLQLPQLPLSANGKLDRKALPLPELKA...QAPGRAPKAGSETIIAAAFSSLLG
GrsA    Y-I-SYFI--DKM--TS---I---Q--E-D-TFGMRVD.YE--RNEI-ETLVTIWQDV--
TycA    Y-L-SYFVK-DKM--TP-D-I------E-D-T-NQS--AYHP-RTET-S-LVSIWQNV--

EntF    CDVQDADADFFALGGHSLLAMKLAAQL
GrsA    IEKIGIKDN-Y----D-IK-IQV--R-
TycA    IEKIGIRDN-YS---D-IQ-IQVV-R-
```

FIGURE 3: Sequence alignment of EntF with GrsA and TycA. The dashes represent identical amino acid identities with EntF. The dots represent gaps between residues. Residues 756–1112 of EntF are compared with residues 332–583 of GrsA and residues 319–572 of TycA.

Comparison of EntF with Other Nonribosomal Peptide Synthetases. A search of the protein sequence of EntF against the NBIF protein sequence databank (version 23.0) revealed homology with the light chain of tyrocidine synthase (TycA) which activates and racemizes phenylalanine in the biosynthesis of the cyclic decapeptide tyrocidine (Weckermann et al., 1988). A stretch of 267 amino acids roughly in the middle of the EntF polypeptide is 38% homologous with the amino terminus of TycA, and this homology increases even further when conservative replacements are considered (Figure 3). Further investigation indicated a similar degree of homology between EntF and the gramicidin S synthetase subunit GrsA, which also carries out the activation and racemization of phenylalanine in the biosynthesis of gramicidin S (Krätzschmar et al., 1989). This homology lends further support to the notion that EntF is the serine activating enzyme.

Subcloning and Overexpression of EntF from pDD4 and Purification from E. coli K12. In order to study the EntF protein separate from the other components required for enterobactin biosynthesis, the *entF* gene was cloned behind the T7 promoter and expressed in *E. coli*. Since it was thought that using the available restriction sites flanking *entF* would lead to a construction with less than optimum expression, a method to remove noncoding DNA at the 5′ end of *entF* up to the Shine–Dalgarno sequence *and* insert a unique restriction site (Xba*I*) at this locus was undertaken using the exonuclease activity of T4 DNA polymerase as previously described to provide for better expression of EntF (Liu et al., 1990; Rusnak et al., 1990). This method, as outlined under Materials and Methods and in Scheme I, resulted in the clone pDD4. The construction of pDD4 was accomplished by subcloning the amino and carboxy termini of *entF* separately since attempts at subcloning the complete *entF* gene into M13mp19 for the deletion procedure were unsuccessful. After transformation of pDD4 into the *E. coli* strain JM109(DE3), the EntF protein was overexpressed to about 5% of the total cellular protein (Figure 4A, lanes 1 and 2) with an estimated molecular weight of 130K–140K. This strain was subsequently chosen to isolate and examine the activity of EntF. Two assays were employed to follow EntF during purification: the L-serine-dependent ATP–[^{32}P]PP_i exchange reaction and the synthesis of enterobactin using an in vitro complementation assay with crude extract from the *entF*$^-$ strain MK1. However, when attempts to purify EntF from crude extracts of JM109(DE3)/pDD4 were undertaken, these activities could not be found (Table I), even though the protein was apparently being expressed. The fact that this overexpressed protein was indeed EntF was

Table I: Copurification from *E. coli* K12 of Activities for the Synthesis of (2,3-Dihydroxybenzoyl)serine (2,3-DBS)-Containing Compounds and L-Serine-Dependent ATP-[^{32}P]PP_i Exchange[a]

		synthesis of 2,3-DBS compounds				L-serine-dependent exchange			
step	protein (mg)	act. + 2,3-DHB (dpm)	act. − 2,3-DHB (dpm)	sp act. (dpm min^{-1} mg^{-1})	purification (*x*-fold)	act. + L-serine (dpm)	act. − L-serine (dpm)	sp act. (dpm min^{-1} mg^{-1})	purification (*x*-fold)
Source: *E. coli* K12									
crude extract	253	1230	580	123	1	10800	2000	16700	1
50% $(NH_4)_2SO_4$ supernatant	135	450	298	49	0.4	2400	2200	1100	0.07
50% $(NH_4)_2SO_4$ ppt	83	1390	600	181	1.5	48800	6300	97100	5.8
Mono-Q anion-exchange chromatography	24	1570	470	832	6.8	47000	2000	340500	20.4
Superose 6 gel chromatography	6.4	890	540	625	5.1	18600	700	314000	18.8
phenyl-Superose chromatography	1.9	1170	490	1324	10.8	36100	900	690100	41.3
Source: JM109(DE3)/pDD4									
crude extract	1163	7341	6823	22		5600	4900	1200	
50% $(NH_4)_2SO_4$ supernatant	273	295	277	3		1700	1700	0	
50% $(NH_4)_2SO_4$ ppt	590	9135	8733	15		9300	6900	4000	

[a] The corresponding activities of extracts prepared from the overproducing strain JM109(DE3)/pDD4 are shown for comparison.

verified by obtaining an amino-terminal sequence of the protein after it had been substantially purified via anion-exchange chromatography, gel filtration chromatography, Blue A dye ligand chromatography, and a final anion-exchange chromatography step; the first 12 residues were identical with those predicted from the DNA sequence of Figure 2. Since the *entF* gene was pieced together in two fragments to form pDD4, it was thought that a problem may have arisen during cloning which would lead to inactive enzyme being produced in vivo. Subsequently, pDD4 was found to complement an *entF* mutant in vivo in the laboratory of Dr. Mark McIntosh at the University of Missouri—Columbia (McIntosh, personal communication) using a complementation test described previously (Fleming et al., 1983), indicating pDD4 encodes an active enzyme in vivo. In addition, EntF expressed from JM109(DE3)/pJS4700, in which the *entF* gene was cloned into the vector pGEM3 in a single cloning step, was as inactive in crude extracts as EntF expressed from pDD4 (data not shown).

Since the EntF enzyme expressed from pDD4 appeared to be largely inactive in vitro, and because we were concerned about possible posttranslational modification issues, partial purification of EntF from wild-type *E. coli* was attempted. As can be seen in Table I, the purported activities of EntF, namely, the in vitro complementation assay with crude extract from an *entF* strain in the synthesis of (2,3-dihydroxybenzoyl)serine derivatives and the L-serine-dependent ATP-[^{32}P]PP_i exchange activities, are obvious even in crude extracts. The ATP-[^{32}P]PP_i exchange is not dependent upon (2,3-dihydroxybenzoyl)serine. After three chromatography steps, EntF had been purified almost 11-fold in the biosynthesis of 2,3-DHB conjugates and over 40-fold in the L-serine-dependent ATP-[^{32}P]PP_i exchange reaction, yielding a sample in which the EntF polypeptide is clearly visible in Coomassie-stained SDS-PAGE gels (Figure 4A, lanes 3 and 4). The ethyl acetate extract of the enzyme incubation mixture was chromatographed on reverse-phase TLC plates to verify that (2,3-dihydroxybenzoyl)serine conjugates were being formed (data not shown). The activity of the enzyme isolated from wild-type *E. coli* is in marked contrast to the activity of EntF obtained from the overexpressing plasmid pDD4 (Table I).

Previously, Bryce & Brot were able to isolate a [^{14}C]seryl-EntF complex by gel filtration chromatography and showed that this complex was stable to dilute acid, a hallmark of thioesters. With EntF purified from *E. coli* K12, we repeated these experiments and found that there was very little difference in the amount of enzyme which was labeled with [^{14}C]serine regardless of whether ATP was present. This was true whether the [^{14}C]seryl-EntF complex was isolated by gel filtration chromatography on Sephadex G25 or by precipitation in dilute trichloroacetic acid at 0 °C. Thus, even though EntF isolated from wild-type *E. coli* has exchange and enterobactin activity, the failure to detect significant loading of the enzyme with [^{14}C]serine suggests that a majority of the enzyme molecules may be inactive.

Labeling of the EntF Polypeptide with β-[^{3}H]Alanine. One reason for the apparent lack of activity in crude extracts of pDD4 may be that the overexpressed EntF was not properly posttranslationally modified. The fact that the synthetases for gramicidin S and tyrocidine have both been shown to contain 4′-phosphopantetheine as a covalent cofactor (Laland et al., 1972; Lee & Lipmann, 1974; Kleinkauf & von Döhren, 1981) and that EntF shares sequence homologies with subunits of these synthetases suggested that 4′-phosphopantetheine may also be a necessary cofactor in enterobactin biosynthesis. With the *E. coli* strain SJ16, a *panD* mutant defective in the biosynthesis of β-alanine, the EntF polypeptide was shown to incorporate β-alanine, suggesting that it contains a covalently bound 4′-phosphopantheine group. Thus, when extracts of whole cells of SJ16, grown in minimal medium without iron supplementation in the presence of β-[^{3}H]alanine, were electrophoresed on sodium dodecyl sulfate–polyacrylamide gels, several peptides were intensely labeled. Lanes 1a and 1b of Figure 4B show a 13% PAGE gel which resolves four ^{3}H-labeled peptides of molecular mass ≈140, ≈35, ≈20, and ≈15 kDa. The ≈15-kDa band represents acylated acyl carrier protein (acyl-ACP) while the ≈20-kDa band represents ACP; these have been reported to contain almost 90% of the total radioactivity in radiolabeling experiments with β-[^{3}H]alanine (Alberts & Vagelos, 1966; Rock & Cronan, 1979; Rock, 1982; Guerra et al., 1988). The identity of the 35-kDa-labeled peptide is unknown, but it may represent the peptide of molecular mass 44 kDa identified previously by SDS-PAGE (Rock, 1982). The 140-kDa-labeled peptide comigrates with EntF. If SJ16 is grown in the same media but in the presence of 2 μM $FeSO_4$, the labeled EntF protein is considerably less intense (Figure 4B, lanes 2a and 2b) and in some cases undetectable, as EntF is transcriptionally regulated by iron and the Fur protein (Pettis & McIntosh, 1987). These results

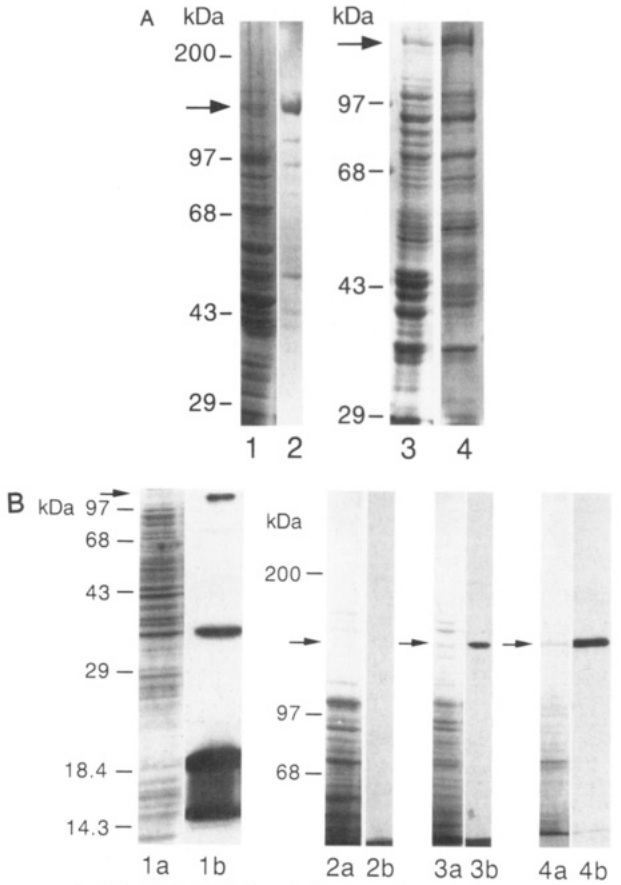


FIGURE 4: (A) SDS–PAGE analysis: lane 1, crude homogenate from JM109(DE3)/pDD4; lane 2, after purification by chromatography on a Mono Q column as described under Materials and Methods (7.5% gel, Coomassie blue stained); lane 3, crude homogenate from K12; lane 4, after purification by Superose 6 gel filtration chromatography as described under Materials and Methods (10% gel, Coomassie blue stained). (B) SDS–PAGE analysis: lane 1a, crude homogenate from SJ16 containing β-[^{3}H]alanine-labeled proteins; lane 1b, autoradiograph (13% gel, Coomassie blue stained); lane 2a, crude homogenate from SJ16 containing β-[^{3}H]alanine-labeled proteins, grown in the presence of 2 μM $FeSO_4$; lane 3a, crude homogenate from SJ16 containing β-[^{3}H]alanine-labeled proteins; lane 4a, after purification by Sephadex G200 gel filtration chromatography as described under Materials and Methods; lanes 2b, 3b, and 4b, autoradiographs of lanes 2a, 3a, and 4a, respectively (7.5% gel, Coomassie blue stained).

strongly suggest that EntF contains 4′-phosphopantetheine or a similar derivative as a covalently bound cofactor.

In order to show that this cofactor is indeed 4′-phosphopantetheine, we purified β-[^{3}H]alanine-labeled EntF from a large-scale (400 mL) culture by ammonium sulfate fractionation, anion-exchange chromatography, and gel filtration chromatography (Figure 4B, lanes 3a,b and 4a,b). Although β-[^{3}H]alanine-labeled EntF was not homogeneous, it was the only radiolabeled protein observable on autoradiographs of SDS–PAGE gels. Hydrolysis of β-[^{3}H]alanine-labeled EntF provided a ^{3}H-labeled sample which was analyzed by silica gel thin-layer chromatography (Jackowski & Rock, 1981). As a control, a sample of β-[^{3}H]alanine-labeled ACP, which also contains 4′-[^{3}H]phosphopantetheine (Alberts & Vagelos, 1966), was subjected to identical hydrolysis conditions. The mobilities of the ^{3}H-labeled compounds derived from the samples on silica gel H (Figure 5A,B) were very similar. Further, treatment of the 4′-[^{3}H]phosphopantetheine derived from EntF with alkaline phosphatase provided a compound which comigrated with pantetheine on silica gel (Figure 5C).

4′-Phosphopantetheine is always found attached to a serine residue in a phosphodiester linkage, and the sequence surrounding the serine residue is usually highly conserved. Several sequences of 4′-phosphopantetheine-containing peptides, which illustrate the conserved residues around the serine attachment site, are compared in Figure 6. Two potential sites in EntF are also shown, although neither shows the striking homologies evident in the other peptides. Whether one or both of these represent 4′-phosphopantetheine binding sites for EntF remains to be shown.

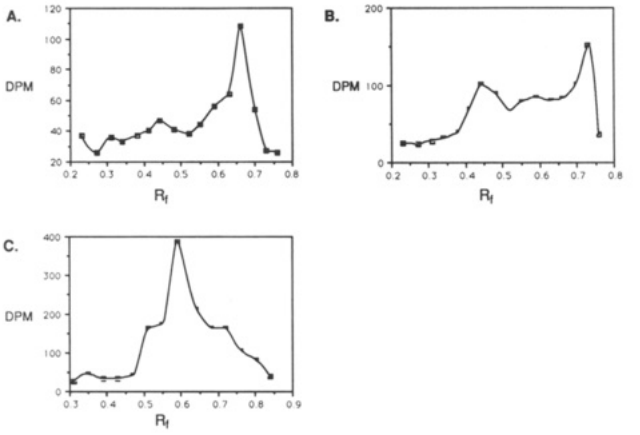


FIGURE 5: (A) Graph of disintegrations per minute vs mobility of the purified radiolabeled compounds derived from base hydrolysis of acyl carrier protein. The peak corresponding to 4′-[^{3}H]phosphopantetheine is located at R_f 0.43 while the more intense band at R_f 0.67 corresponds to pantetheine. (B) Graph of disintegrations per minute vs mobility of the purified radiolabeled compounds derived from base hydrolysis of EntF. 4′-[^{3}H]phosphopantetheine was located at R_f 0.43 while the band for pantetheine was observed at R_f 0.72 (mobilities of the radiolabeled compounds were found to be affected by minor contaminents). (C) Graph of disintegrations per minute vs mobility of the radiolabeled compound from the alkaline phosphatase treatment of 4′-[^{3}H]phosphopantetheine derived from EntF. The major band at R_f 0.59 corresponds to pantetheine.

Field Mustard ACP	A	D	L	G	A	D	S	L	D	T	.	V	E	I	V	M
Spinach ACP	S	K	L	G	A	D	S	L	D	T	.	V	E	I	V	M
Barley ACP	S	E	L	G	A	D	S	L	D	T	.	V	E	I	V	M
E. coli ACP	E	D	L	G	A	D	S	L	D	T	.	V	E	L	I	M
Neurospora crassa ACP	N	D	L	G	L	D	S	L	D	T	.	V	E	V	V	M
Rape ACP	V	D	L	G	A	D	S	L	D	T	.	V	E	I	V	M
Citrate Lyase	I	Q	L	E	I	D	S	I	V	K	.	Q	E	F	G	A
Granaticin PK Synthase	E	E	L	G	Y	D	S	L	A	L	.	M	E	S	A	S
Tetracenomycin PK Synthase	Q	D	L	G	Y	D	S	I	A	L	.	L	E	I	S	A
EntF (107-122)	Q	D	L	R	V	D	S	G	K	P	L	V	F	H	Q	L
EntF (1000-1015)	F	A	L	G	G	H	S	L	L	A	M	K	L	A	A	Q

FIGURE 6: Alignments of the sequences surrounding the phosphopantetheine binding site of several acyl carrier proteins (ACP's) and the two potential phosphopantetheine binding sites of EntF. The boxed regions are residues which are strictly conserved between the majority ACP's. The dots represent gaps between residues. The references are the following: ACP from field mustard *Brassica campestris* (Rose et al., 1987); ACP from spinach *Spinacia oleracea* (Scherer & Knauf, 1987); ACP from barley *Hordeum vulgare* (Hoj & Svendsen, 1983); ACP from *Escherichia coli* (Jackowski & Rock, 1987); ACP from *Neurospora crassa* (Brody & Mikolajczyk, 1988); ACP from rape *Brassica napus* (Safford et al., 1988); citrate lyase from *Klebsiella pneumoniae* (Beyreuther et al., 1978); ACP of granaticin-producing polyketide synthase from *Streptomyces violaceoruber* (Sherman et al., 1989); ACP of tetracenomycin C producing polyketide synthase from *Streptomyces glaucescens* (Bibb et al., 1989).

Discussion

Herein we report the nucleotide sequence of the *entF* gene, and as a result, all the known enterobactin biosynthetic genes have now been sequenced. The encoded polypeptide shares homologies with the light chains of gramicidin S (GrsA) and tyrocidine (TycA) synthetases, two enzymes involved in the nonribosomal biosynthesis of cyclic peptides in *Bacillus*. This homology indicates a common ancestry for EntF, GrsA, and

TycA and now places EntF within the family of enzymes involved in the biosynthesis of cyclic peptides, lactones, and depsipeptides through protein template mechanisms. Since EntF, GrsA, and TycA activate amino acids (L-serine in the case of EntF and D- or L-phenylalanine for both GrsA and TycA) to the amino acyladenylates, one might expect that this homology represents either an ATP or an amino acyladenylate binding site on the enzyme. In fact, similar homologies are found in other enzymes which activate carboxylates such as EntE, which forms 2,3-DHB–AMP in enterobactin biosynthesis (Staab et al., 1989), 4-coumaryl:CoA ligase, which converts 4-coumarate to 4-coumaryladenylate before forming the CoA thioester (Lozoya et al., 1988), and luciferin monooxygenase, which also proceeds through an acyladenylate intermediate (De Wet et al., 1987). We have looked carefully within these homologous regions for a consensus nucleotide binding domain typically found in protein kinases and nicotinamide utilizing dehydrogenases (Walker et al., 1982; Wierenga et al., 1986; Higgins et al., 1986) but none was apparent. We also have not found sequences which are known to interact with ATP or the amino acyladenylate of the aminoacyl-tRNA synthetases (Burbaum et al., 1989). These homologies present in EntF and the other activating enzymes might represent, then, a new structural motif for binding ATP and stabilizing the acyladenylate intermediates. In addition, four cysteines conserved between GrsA and TycA (Krätzschmar et al., 1989) that might be important in tethering phenylalanine are not conserved in EntF.

In an attempt to dissect the function of EntF in the enterobactin biosynthetic machinery, we have overexpressed it (pDD4 vector) and found that it was unable either to carry out L-serine activation or to participate in enterobactin biosynthesis. As a result, we have resorted to a partial purification of the peptide from wild-type *E. coli* and verified that it does indeed activate L-serine at least transiently via an L-seryl-AMP intermediate as indicated by the L-serine-dependent ATP–[^{32}P]PP$_i$ exchange and that it does participate in the formation of (2,3-dihydroxybenzoyl)serine conjugates (Table I). However, we were unable to detect measurable loading of the partially purified wild-type enzyme with [^{14}C]serine. This suggests that the majority of EntF molecules have become inactive during purification. Similar behavior has been observed in a previous purification of the multienzyme cyclosporin synthetase which was able to catalyze the ATP–[^{32}P]PP$_i$ exchange reaction of all constitutive amino acids of cyclosporin A, but was unable to synthesize the complete molecule (Zocher et al., 1986). This lability, which appears to be a hallmark of this class of amino acid activating enzymes, has only recently been overcome in the case of cyclosporin synthetase (Lawen et al., 1990). The inactivity of overexpressed EntF from pDD4 even in crude extract may be due to exceptional lability or possibly insufficient posttranslational modification of the overexpressed enzyme with phosphopantetheine. As we have demonstrated, EntF contains a phosphopantetheine cofactor, and two serine residues are potential sites for its attachment, although neither site shares the striking homology seen in other phosphopantetheine-containing peptides (Figure 6). Thus, if insufficient posttranslational processing of EntF is the cause for inactivity, then one wonders whether these sequences contribute to an inefficient loading of cofactor to EntF by an as yet unidentified enzyme. These two serines as well as the surrounding residues are targets for mutagenesis to evaluate the effects of β-[^{3}H]alanine incorporation.

It is likely that EntF not only catalyzes L-serine activation but also, in the presence of 2,3-DHB and EntE, will generate the amide bond linkage present 3 times in the cyclic (2,3-dihydroxybenzoyl)serine trimer enterobactin. Experiments supporting this course have already been reported (Bryce & Brot, 1971). However, in that study, the enzyme was only partially resolved with gel filtration and anion-exchange chromatographies, and the stoichiometry of loading was not determined. To date, we have not detected linear or cyclic (dihydroxybenzoyl)serine (DBS) species on incubation of pure EntE, EntB, EntF (≈80% homogeneous from pDD4), ATP, 2,3-DHB, and L-serine, but the EntF peptide was inactive in catalyzing ATP–[^{32}P]PP$_i$ exchange at this stage. However, the coupling and exchange reactions may not proceed until EntD is present to serve as a template, membrane-associated for vectorial export of the cyclic trilactone enterobactin product so as not to exacerbate the already low intracellular iron concentration. We have already overproduced and purified EntE and EntB (Rusnak et al., 1989, 1990) so that EntD is the sole component yet missing to attempt a clean four-protein in vitro reconstitution.

Assuming for the moment that, as precedented, an EntE–EntF complex is the key one in transferring the activated 2,3-DHB moiety from 2,3-DHB–AMP in the active site of EntE to the seryl moiety at the active site of EntF, then several questions arise. Since DBS-dependent ATP–[^{32}P]PP$_i$ exchange was not detected, the simplest formulation is amide bond formation in a *direct* transfer from DHB–AMP to the α-amino group of seryl-AMP bound at EntF to yield DHB–seryl-AMP, still carboxyl-activated for subsequent ester bond formation. If amide formation occurred between the acyl-AMP and aminoacyl-AMP species, there would be no role for the thiol group of the phosphopantetheine prosthetic group.

On the other hand, the thiol group of phosphopantetheine could be involved *prior* to amide bond formation, and the acyl group either of seryl-AMP or of DHB-AMP could be transferred to the phosphopantetheine thiol nucleophile. The conversion of seryl-AMP to seryl-*S*-EntF has precedent in the other members of this general class of catalysts, e.g., gramicidin S synthetase, tyrocidine synthetase, enniatin synthetase, and, most recently and spectacularly, cyclosporin synthetase (Lawen & Zocher, 1990), which activates 11 amino acids; these are large polypeptides of 100–800 kDa with multiple capacities for amino acid activation and covalent tethering. It is reasonable that the 142-kDa EntF polypeptide follows similar form and function. Amide bond formation from EntE–DHB–AMP would then yield a charged DBS–phosphopantetheinyl–EntF. As yet, though, in none of the multienzyme complexes noted has it actually been proven that the phosphopantetheine thiolate is the enzyme thiol to which the activated amino acid is transferred, and indeed phosphopantethein stoichiometries are undetermined. Enzyme cysteines may instead activate the amino acids in nonribosomal amide bond assembly.

Alternatively, the phosphopantetheine thiol could, in enterobactin assembly, accept the activated 2,3-DHB group from DHB–AMP–EntE, i.e., engage in acyl transfer rather than aminoacyl transfer. This is in a sense more attractive since acyl transfer is the role of phosphopantetheine in fatty acid synthetase (Wakil, 1989), where the growing acyl chain cycles between phosphopantetheine and a cysteine thiol to permit the alternate loading and elongation processes. Furthermore, the part of phosphopantetheine in ester bond formation has thus far not been documented. Of note in this regard, EntE shows homology with the enzyme 4-coumarate:coenzyme A ligase, which forms 4-coumaryl-CoA through the reactive acyladenylate intermediate. It is tempting to speculate that this

homology represents a recognition site for, in the case of 4-coumarate:CoA ligase, the pantetheine stretch of coenzyme A, and in the case of EntE, the identical molecular arm of phosphopantetheine which is tethered to EntF. One will need fully active, purified EntF to discern between and among loading of [^{14}C]serine and/or 2,3-[^{3}H]DHB groups onto EntF and their regiochemical location and state of elongation, a problem not yet generically solved in any member of this class of multienzyme complexes.

Even less obvious in outline are the latter steps in presumed EntF template-mediated assembly of cyclic enterobactin, e.g., how the second and third DBS moieties are loaded and assembled. The EntF amino acid sequence reveals no obvious repeat structures which could alert one to this possibility. Also unclear is how the three ester bonds are assembled. As the enniatin synthetase molecular logic becomes unraveled (Zocher et al., 1982), parallel logic is likely to apply. One particular clue which may direct future approaches toward understanding enterobactin ester bond formation and eventual cyclization is another serine residue found in EntF, serine-1138, which lies within the conserved pentapeptide sequence G-X_1-S-X_2-G found in lipases, thioesterases, and the acetyltransferase subunit of FAS, where X_1 is Y or H and X_2 is usually L, M, F, or Q (Mikkelsen et al., 1985; Krätzschmar et al., 1989; Brady et al., 1990; Winkler et al., 1990). The sequence LLGYSLG in EntF is identical with the sequence around the active-site serine of triacylglycerol lipases (Wion et al., 1987; Brady et al., 1990; Winkler et al., 1990). Although its role in EntF is not understood at this time, one recalls the role this serine plays in transacylation in fatty acid synthetase, transferring acetyl and malonyl units from the CoA conjugates to either cysteine or pantetheine thiols, as the respective *O*-seryl enzyme intermediates (Mikkelsen et al., 1985). Interestingly, the gene *grsT*, discovered fortuitously during the cloning of the gramicidin synthetase genes, is highly homologous to thioesterases of FAS (Krätschmar et al., 1989). In EntF, it may be that (dihydroxybenzoyl)seryl-*O*-seryl-EntF is a cognate intermediate, set up for elongation transesterification steps.

Acknowledgments

We thank Prof. Mark A. McIntosh for running the in vivo complementation assays with pDD4 and for providing the plasmid pITS32 and the *E. coli* strain MK1. We thank Professor Charles Earhart and Dr. Janet Staab for providing the plasmid pJS4700. We gratefully acknowledge Prof. Eugene Kennedy, Dr. Anthony Fischl, Dr. Otto Geiger, and Dr. Herman Spaink for providing the *E. coli* strain SJ16 and samples of purified [^{3}H]acyl carrier protein and for much helpful advice.

References

Alberts, A. W., & Vagelos, P. R. (1966) *J. Biol. Chem. 241*, 5201–5204.

Beyreuther, K., Bohmer, H., & Dimroth, P. (1978) *Eur. J. Biochem. 87*, 101–110.

Bibb, M. J., Biró, S., Motamdei, H., Collins, J. F., & Hutchinson, C. R. (1989) *EMBO J. 8*, 2727–2736.

Brady, L., Brzozowski, A. M., Derewenda, Z. S., Dodson, E., Dodson, G., Tolley, S., Turkenburg, J. P., Christiansen, L., Huge-Jensen, B., Norskov, L., Thim, L., & Menge, U. (1990) *Nature 343*, 767–770.

Brody, S., & Mikolajczyk, S. (1988) *Eur. J. Biochem. 173*, 353–359.

Brot, N., Goodwin, J., & Fales, H. (1966) *Biochem. Biophys. Res. Commun. 25*, 454–461.

Bryce, G. F., & Brot, N. (1972) *Biochemistry 11*, 1708–1715.

Burbaum, J. J., Starzyk, R. M., & Schimmel, P. (1990) *Proteins: Struct., Funct., Genet. 7*, 99–111.

Coderre, P. E., & Earhart, C. F. (1984) *FEMS Microbiol. Lett. 25*, 111–116.

Dale, R. M. K., McClure, B. A., & Hutchins, J. P. (1985) *Plasmid 13*, 31–40.

De Wet, J. R., Wood, K. V., DeLuca, M., Helinski, D. R., & Subramani, S. (1987) *Mol. Cell Biol. 7*, 725–737.

Elkins, M. F., & Earhart, C. F. (1988) *FEMS Microbiol. Lett. 56*, 35–40.

Elkins, M. F., & Earhart, C. F. (1989) *J. Bacteriol. 171*, 5443–5451.

Fleming, T. P., Nahlik, M. S., & McIntosh, M. A. (1983) *J. Bacteriol. 156*, 1171–1177.

Gevers, W., Kleinkauf, H., & Lipmann, F. (1968) *Proc. Natl. Acad. Sci. U.S.A. 60*, 269–276.

Gevers, W., Kleinkauf, H., & Lipmann, F. (1969) *Proc. Natl. Acad. Sci. U.S.A. 63*, 1335–1342.

Greenwood, K. T., & Luke, R. K. J. (1976) *Biochim. Biophys. Acta 454*, 285–297.

Greenwood, K. T., & Luke, R. K. J. (1980) *Biochim. Biophys. Acta 614*, 185–195.

Guerra, D. J., Dziewanowska, K., Ohlrogge, J. B., & Beremand, P. D. (1988) *J. Biol. Chem. 263*, 4386–4391.

Higgins, C. F., Hiles, I. D., Salmond, G. P. C., Gill, D. R., Downie, J. A., Evans, I. J., Holland, I. B., Gray, L., Buckel, S. D., Bell, A. W., & Hermodson, M. A. (1986) *Nature 323*, 448–450.

Hoj, P. B., & Svendsen, I. (1983) *Carlsberg Res. Commun. 48*, 285–305.

Jackowski, S., & Rock, C. O. (1981) *J. Bacteriol. 148*, 926–932.

Jackowski, S., & Rock, C. O. (1987) *J. Bacteriol. 169*, 1469–1473.

Keller, U. (1987) *J. Biol. Chem. 262*, 5852–5856.

Kleinkauf, H., & Von Döhren, H. (1981) *Curr. Top. Microbiol. Immunol. 91*, 129–177.

Krätzschmar, J., Krause, M., & Marahiel, M. A. (1989) *J. Bacteriol. 171*, 5422–5429.

Laland, S. G., Frøyshov, Ø., Gilhuus-Moe, C., & Zimmer, T. L. (1972) *Nature (London), New Biol. 239*, 43–44.

Langman, L., Young, I. G., Frost, G. E., Rosenberg, H., & Gibson, F. (1972) *J. Bacteriol. 112*, 1142–1149.

Lawen, A., & Zocher, R. (1990) *J. Biol. Chem. 265*, 11355–11360.

Lee, S. G., & Lipmann, F. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 607–611.

Liu, J., Duncan, K., & Walsh, C. T. (1989) *J. Bacteriol. 171*, 791–798.

Liu, J., Quinn, N., Berchtold, G. A., & Walsh, C. T. (1990) *Biochemistry 29*, 1417–1425.

Lozoya, E., Hoffmann, H., Douglas, C., Schulz, W., Scheel, D., & Hahlbrock, K. (1988) *Eur. J. Biochem. 176*, 661–667.

Luke, R. K. J., & Gibson, F. (1971) *J. Bacteriol. 107*, 557–562.

Lundrigan, M. D., & Kadner, R. J. (1986) *J. Biol. Chem. 261*, 10797–10801.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) in *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Mikkelsen, J., Højrup, P., Rasmussen, M. M., Roepstorff, P., & Knudsen, J. (1985) *Biochem. J. 227*, 21–27.

Nahlik, M. S., Fleming, T. P., & McIntosh, M. A. (1987) *J. Bacteriol. 169*, 4163–4170.

Ozenberger, B. A., Nahlik, M. S., & McIntosh, M. A. (1987) *J. Bacteriol. 169*, 3638–3646.
Ozenberger, B. A., Brickman, T. J., & McIntosh, M. A. (1989) *J. Bacteriol. 171*, 775–783.
Pettis, G. S., & McIntosh, M. A. (1987) *J. Bacteriol. 169*, 4154–4162.
Pettis, G. S., Brickman, T. J., & McIntosh, M. A. (1988) *J. Biol. Chem. 263*, 18857–18863.
Pierce, J. R., & Earhart, C. F. (1986) *J. Bacteriol. 166*, 930–936.
Pierce, J. R., Pickett, C. L., & Earhart, C. F. (1983) *J. Bacteriol. 155*, 330–336.
Rock, C. O. (1982) *J. Bacteriol. 152*, 1298–1300.
Rose, R. E., DeJesus, C. E., Moylan, S. L., Ridge, N. P., Scherer, D. E., & Knauf, V. C. (1987) *Nucleic Acids Res. 15*, 7197.
Roskoski, R., Gevers, W., Kleinkauf, H., & Lipmann, F. (1970a) *Biochemistry 9*, 4839–4845.
Roskoski, R., Kleinkauf, H., Gevers, W., & Lipmann, F. (1970b) *Biochemistry 9*, 4846–4851.
Rusnak, F., Faraci, W. S., & Walsh, C. T. (1989) *Biochemistry 28*, 6827–6835.
Rusnak, F., Liu, J., Quinn, N., Berchtold, G. A., & Walsh, C. T. (1990) *Biochemistry 29*, 1425–1435.
Safford, R., Windust, J. H. C., Lucas, C., de Silva, J., James, C. M., Hellyer, A., Smith, C. G., Slabas, A. R., & Hughes, S. G. (1988) *Eur. J. Biochem. 174*, 287–295.
Sakaitani, M., Rusnak, F., Quinn, N., Frigo, T. B., Tu, C., Berchtold, G. A., & Walsh, C. T. (1990) *Biochemistry 29*, 6789–6798.
Sanger, F., Nicklens, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.
Scherer, D. E., & Knauf, V. C. (1987) *Plant Mol. Biol. 9*, 127–134.
Sherman, D. H., Malpartida, F., Bibb, M. J., Kieser, H. M., Bibb, M. J., & Hopwood, D. A. (1989) *EMBO J. 8*, 2717–2725.
Shine, J., & Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 1342–1346.
Staab, J. F., & Earhart, C. F. (1990) *J. Bacteriol. 11*, 6403–6410.
Staab, J. F., Elkins, M. F., & Earhart, C. F. (1989) *FEMS Microbiol. Lett. 59*, 15–20.
Studier, F. W., & Moffatt, B. A. (1986) *J. Mol. Biol. 189*, 113–130.
Vogel, H. J., & Bonner, D. M. (1956) *J. Biol. Chem. 218*, 97.
Walker, J. E., Saraste, M., Runswick, M., & Gay, N. J. (1982) *EMBO J. 1*, 945–951.
Weckermann, R., Fürbass, & Marahiel, M. A. (1988) *Nucleic Acids Res. 16*, 11841.
Wierenga, R. K., Terpstra, P., & Hol, W. G. J. (1986) *J. Mol. Biol. 187*, 101–107.
Winkler, F. K., D'Arcy, A., & Hunziker, W. (1990) *Nature 343*, 771–774.
Wion, K. L., Kirchgessner, T. G., Lusis, A. J., Schotz, M. C., & Lawn, R. M. (1987) *Science 235*, 1638–1641.
Woodrow, G. C., Young, I. G., & Gibson, F. (1975a) *J. Bacteriol. 124*, 1–6.
Woodrow, G. C., Young, I. G., & Gibson, F. (1975b) *Biochim. Biophys. Acta 582*, 145–153.
Young, I. G., Langman, L., Luke, R. K. J., & Gibson, F. (1971) *J. Bacteriol. 106*, 51–57.
Zocher, R., Keller, U., & Kleinkauf, H. (1982) *Biochemistry 21*, 43–48.

Chemical Synthesis of Oligodeoxynucleotide Dumbbells

Gary W. Ashley* and Diana M. Kushlan
Department of Chemistry, Northwestern University, Evanston, Illinois 60208
Received September 4, 1990; Revised Manuscript Received November 5, 1990

ABSTRACT: The chemical synthesis of DNA dumbbells is investigated by using two sequences, *cyclo*-d(GCG-T_4-CGCCGC-T_4-GCG) and *cyclo*-d(TTCC-T_4-GGAATTCC-T_4-GGAA). This method readily and inexpensively yields multimicromole quantities of circular DNA, allowing detailed structural and physical studies to be carried out. Linear oligomers of sequence d(GCG-T_4-CGCCGC-T_4-GCG) having either 3′- or 5′-phosphates were cyclized in 40% and 67% isolated yield, respectively, by using 1-[3-(dimethylamino)propyl]-3-ethylcarbodiimide. Formation of the circular product is confirmed by a 28 °C increase in the optical melting temperature, anomalously rapid electrophoretic migration, sequential nuclear Overhauser enhancements between protons of G1 and G20, and observed nuclear coupling between the ligated phosphorus and protons of both G1 and G20. *cyclo*-d(TTCC-T_4-GGAATTCC-T_4-GGCC) was synthesized from the corresponding linear 3′-phosphate in 80% yield by the same procedure. Chemical ligation is most effective for 3′-phosphorylated nick sites flanked by two purine bases.

DNA structures containing unpaired bases play significant roles in biological processes. Such noncanonical structures occur as cruciform junctions in supercoiled DNA (Lilley, 1989), near control regions in the chromosome (Muller & Fitch, 1982), and in bulged regions during replication (Yager & von Hipple, 1987). A major interest so far has been the thermodynamics of such conformations (Senior et al., 1988), although their unusual drug binding properties are coming under intense investigation (Williams & Goldberg, 1988; Guo et al., 1990).

Proper understanding of the link between noncanonical DNA structure and biological function requires structural

* To whom correspondence should be addressed.

0006-2960/91/0430-2927$02.50/0