JOURNAL OF BACTERIOLOGY, Mar. 1991, p. 1729–1737
0021-9193/91/051729-09$02.00/0

Vol. 173, No. 5

Isolation, Cloning, Mapping, and Nucleotide Sequencing of the Gene Encoding Flavodoxin in *Escherichia coli*

CHARLES OSBORNE, LEE-MAY CHEN, AND ROWENA G. MATTHEWS*

Biophysics Research Division and Department of Biological Chemistry, The University of Michigan, Ann Arbor, Michigan 48109

Received 31 October 1990/Accepted 19 December 1990

The flavodoxins constitute a highly conserved family of small, acidic electron transfer proteins with flavin mononucleotide prosthetic groups. They are found in prokaryotes and in red and green algae, where they provide electrons at low potentials for the reduction of nitrogen by nitrogenase, for the light-dependent reduction of $NADP^+$ in photosynthesis, and for the reduction of sulfite. Proteins with the physical characteristics of flavodoxins have been implicated in the reductive activation of pyruvate formate-lyase and cobalamin-dependent methionine synthase in *Escherichia coli*. We have purified flavodoxin to homogeneity from *E. coli*, determined its N-terminal amino acid sequence, and used this sequence to construct a 64-fold degenerate oligonucleotide probe for the flavodoxin gene. Because the phenotype of a flavodoxin mutant is not known, we used this degenerate probe to screen the phages of the Kohara library and identified two phages, with inserts mapping at ~16 min, that hybridized to the probe. The flavodoxin gene, designated *fldA*, was subcloned from the DNA in the overlap region of these two clones. The deduced amino acid sequence, determined by nucleotide sequencing of the flavodoxin gene, shows strong homology with flavodoxins from nitrogen-fixing bacteria and cyanobacteria. The *fldA* gene maps at 15.9 min on the *E. coli* chromosome and is transcribed in a counterclockwise direction.

Flavodoxins are a family of small, acidic electron transfer proteins with flavin mononucleotide (FMN) prosthetic groups. They differ from the majority of other flavoproteins in that they are unable to be reduced by hydride transfer, and hence they do not react directly with pyridine nucleotides or carbohydrate substrates but are only able to accept and donate electrons. Thus, they are classified as obligate electron transferases (19). Flavodoxins have been identified in prokaryotes and in red and green algae but not in higher plants or animals. In nitrogen-fixing bacteria such as *Azotobacter vinelandii* and *Klebsiella pneumoniae*, flavodoxin provides the electrons needed for the reduction of dinitrogen to ammonia by nitrogenase, serving as an electron donor to the Fe protein of nitrogenase (18; reference 4 and references cited therein). In cyanobacteria, flavodoxin sometimes serves as the electron donor for the ferredoxin-$NADP^+$ oxidoreductase, providing reducing equivalents for the light-dependent reduction of $NADP^+$ associated with photosynthesis. During growth in the presence of adequate iron in the medium, reduction of ferredoxin-$NADP^+$ oxidoreductase is normally accomplished by the 2Fe-2S electron transfer protein ferredoxin, but on growth in iron-deficient media, flavodoxin synthesis is induced and flavodoxin substitutes for ferredoxin (20). The flavodoxin gene from *Anacystis nidulans* has been cloned and sequenced and shown to be regulated at the transcriptional level by the level of iron in the medium (16). In anaerobic bacteria such as *Clostridium pasteurianum*, flavodoxins can substitute for 8Fe-8S ferredoxins in the phosphoroclastic oxidation of pyruvate as follows: $H^+ + CH_3COCOO^- + HPO_4^{2-} \longrightarrow CH_3COOPO_3^{2-} + CO_2 + H_2$. The clastic system includes the enzymes pyruvate ferredoxin-oxidoreductase, phosphotransacetylase, and hydrogenase and requires coenzyme A as a cofactor; ferredoxin or flavodoxin is required to mediate electron transfer between pyruvate-ferredoxin oxidoreductase and hydrogenase (20). Flavodoxin or ferredoxin plays an analogous role in dissimilatory sulfate reduction, mediating the transfer of electrons between hydrogen and sulfite (20).

In *Escherichia coli*, flavodoxin is constitutively synthesized (31). It has been shown to participate in the activation of pyruvate formate-lyase associated with the shift from aerobic to anaerobic conditions during growth on glucose. Ferredoxin cannot substitute for flavodoxin in this reaction (14). The activation of pyruvate formate-lyase is accomplished by an iron-dependent activator protein (enzyme II), reduced flavodoxin, adenosylmethionine (AdoMet), and the effector pyruvate. During this reaction, AdoMet is reductively processed to yield methionine, adenine, and 5′-deoxyribose, and it has been suggested that transient adenosylation of enzyme II is required for its function in the activation of pyruvate formate-lyase (13). According to this hypothesis, flavodoxin mediates electron transfer to the iron of enzyme II and the reaction is coupled to an adenosyl transfer from AdoMet to the iron center. The adenosyl-enzyme II complex may function in a manner similar to that of adenosylcobalamin, abstracting a hydrogen atom from an amino acid side chain to generate the protein radical associated with the active form of pyruvate formate-lyase, with concomitant formation of 5′-deoxyadenosine and a reduced metal center (12).

A protein with the physical properties of a flavodoxin, designated the F protein, was shown to provide electrons for the reductive activation of cobalamin-dependent methionine synthase in *E. coli* (8). In this article, we demonstrate by nucleotide sequence analysis that this protein is a member of the flavodoxin family. Reductive activation of cobalamin-dependent methionine synthase also requires AdoMet (17) and is coupled to the transfer of the methyl group from AdoMet to the reduced cobalamin cofactor (2, 8, 30). Thus, the two reactions involving flavodoxin in *E. coli* both appear

* Corresponding author.

to involve reduction of a protein-AdoMet complex with concomitant heterolytic cleavage of a C-S bond of AdoMet.

It would be interesting to study the role of flavodoxin in in vivo metabolism in *E. coli* and to ascertain whether this protein plays an essential role in the activation of pyruvate formate-lyase and methionine synthase in vivo. It is possible that flavodoxin is also involved in other reactions that require the generation of an enzyme-bound radical such as the radicals of the aerobic iron-containing ribonucleotide reductase (27) or the anaerobic AdoMet-linked ribonucleotide reductase (6). Such studies have been hampered by the fact that the gene encoding flavodoxin in *E. coli* had not been identified and no mutants lacking flavodoxin are known. Furthermore, the phenotype of a mutant lacking flavodoxin remains to be characterized, so that phenotypic selection of mutant species is not possible at present.

In the present study, we have purified flavodoxin from *E. coli*, monitoring its purification by its ability to provide electrons for the reductive activation of methionine synthase. We have determined the amino acid sequence at the N terminus of the protein by Edman degradation and used this information to construct an oligonucleotide probe for the flavodoxin gene. The flavodoxin gene was located in two phages from the Kohara library of genomic DNA from *E. coli* K-12 (28) by Southern blotting and subcloned from a region of DNA that was present in the inserts of both these phages. The *fld* gene encodes a protein with significant homologies to flavodoxins from other prokaryotes and maps at 15.9 min on the *E. coli* chromosome.

MATERIALS AND METHODS

Materials. The phagemid vector Bluescript II SK$^+$ (pBS2SK$^+$) was purchased from Stratagene (La Jolla, Calif.), and the vectors pTZ18 and pTZ19 were from Bio-Rad.

Media and buffer solutions. Luria broth-agarose for plates was prepared as described by Sambrook et al. (23). Where indicated, the Luria broth medium was supplemented with ampicillin (40 μg/ml) and/or X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside; 0.8 mg per plate) with IPTG (isopropyl-β-D-galactopyranoside; 0.8 mg per plate). KPB is potassium phosphate buffer, pH 7.2. Prehybridization and hybridization buffers were prepared for plaque hybridization as described by Sambrook et al. (23).

Assay of flavodoxin. Flavodoxin can be assayed for its ability to transfer electrons to cobalamin-dependent methionine synthase in the presence of NADPH, AdoMet, and NADPH-flavodoxin oxidoreductase (8). The assays contained 100 mM KPB, 1 mM homocysteine, 250 μM (6*R*,*S*)-[*methyl*-^{14}C]CH_3-H_4folate (2,000 dpm/nmol), 38 μM AdoMet, 80 pM methionine synthase, 100 pM NADPH-flavodoxin oxidoreductase, and 0.4 to 4 pM flavodoxin in a total volume of 1 ml. The concentration of flavodoxin was estimated from the A_{464} by using an extinction coefficient of 8,420 M^{-1} cm^{-1} (8). Cobalamin-dependent methionine synthase was purified as described by Frasca et al. (7), and NADPH flavodoxin oxidoreductase was purified from a fraction of the flavodoxin preparation (see below) by adaptation of the method of Fujii and Huennekens (8).

Tubes containing all the components of the assay mixture except NADPH were equilibrated in a 37°C bath. NADPH was added to initiate the assay, and the assay solutions were incubated for 10 min and then quenched by heating to 98°C for 2 min. After cooling the tubes on ice, the contents of each tube were passed over AG1×8 chloride minicolumns (a 13-cm pasteur pipette with a glass wool plug, filled to a 5-cm height with resin). After the columns were loaded, they were rinsed with 2 ml of glass-distilled water. The eluate from loading and rinsing was collected in a scintillation vial, 10 ml of aqueous scintillation fluid was added, and the contents were mixed and counted. Radiolabeled CH_3-H_4folate is retained on the column, while the labeled methionine product of the transmethylation catalyzed by active methionine synthase is eluted.

Purification of flavodoxin from *E. coli*. Flavodoxin was purified from packed, washed, and frozen cells of *E. coli* B (ATCC 11303) obtained from Grain Processing (Muscatine, Iowa). These cells had been grown aerobically in Kornberg minimal medium, supplemented with 0.15 μM cyanocobalamin, and harvested at three-quarters exponential phase. The cells were disrupted by sonication in buffer containing phenylmethanesulfonyl fluoride (400 mg/liter). Batch chromatography on DEAE-52 was used to separate fractions containing flavodoxin, NADPH-flavodoxin oxidoreductase, and methionine synthase. The fraction containing flavodoxin was further purified by column chromatography on DEAE-52 and hydroxylapatite and finally by fast protein liquid chromatography on a MonoQ HR 16/10 column. Details of the purification procedure are available on request. Flavodoxin fractions can be located by their orange color and evaluated for purity by visible absorbance spectroscopy. Pure fractions of flavodoxin show maxima at 464 and 369 nm, with an $A_{369/464}$ of 0.85 (8). Identification of the purified flavodoxin was verified by enzymatic assay. Selected fractions of flavodoxin were subjected to polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate and staining with Coomassie blue dye to evaluate their purity.

Determination of the N-terminal amino acid sequence of flavodoxin. Amino acid sequence analysis on a homogeneous sample of flavodoxin from *E. coli* was performed by using an Applied Biosystems model 470A gas phase sequencer. A 1-nmol portion of reduced alkylated protein was precipitated with cold acetonitrile and redissolved in 88% formic acid. High-performance liquid chromatographic analysis of sequencer-derived phenylthiohydantoin-amino acids was performed by using a Beckman model 332 gradient liquid chromatograph equipped with a Hewlett-Packard 3390A integrator, employing a modification of the procedure of Tarr (29).

Identification of the flavodoxin gene in the Kohara library. The Kohara phage library of *E. coli* genomic DNA was screened by plaque hybridization by using a procedure adapted from the protocol described by Silhavy et al. (24). A 500-μl volume of cells of strain MM294 suspended in 10 mM $MgSO_4$ to an A_{600} of 2 to 3 was mixed with 15 ml of 0.7% top agarose prewarmed to 37°C, and the suspension was plated onto 135-mm-diameter Luria broth plates. After the cells were allowed to grow for 1 h at 37°C, 0.5-μl aliquots of each Kohara lysate were spotted in a grid pattern. A total of 120 lysates from the 476-element Kohara library were spotted on each 135-mm plate. Lysis was apparent after 4 h, when clear plaques were seen. The plates were refrigerated for 1 h to harden the top agarose before nitrocellulose lifts were prepared. Plaque hybridization was performed overnight at 55°C by using 250 ng of a 5′ end-labeled 20-mer oligonucleotide with 64-fold degeneracy as the probe, and the filters were washed five to six times with 6× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) at room temperature and then for 30 min with 6× SSC at 44°C. The wet nitrocellulose filters were wrapped in Saran Wrap, and fluorograms were prepared by exposure of Kodak XAR5 film at −80°C.

```
(Met)•Ala•Ile•Thr•Gly•Ile•Phe•Gly•Ser•Asp•Thr•Gly•Asn•Thr

(ATG)•GCT•ATC•ACC•GGT•ATC•TT
       C   T   C   T
       A
       G
```

FIG. 1. *E. coli* flavodoxin N-terminal amino acid sequence and oligonucleotide probe.

Purification of phage DNA, subcloning, and nucleotide sequencing. λ phage DNA was purified from Kohara clones 172 and 173 by using the protocol described by Chisholm (3). Subcloning was accomplished by using standard protocols. Double- or single-stranded plasmid DNA was isolated according to protocols provided by Stratagene. DNA sequencing by the Sanger dideoxy chain termination method with α-^{35}S-dATP was carried out with Sequenase from United States Biochemicals (Cleveland, Ohio) according to the manufacturer's instructions.

Nucleotide sequence accession number. The sequence of *fldA* and its flanking regions has been submitted to GenBank and has been assigned accession number M59426.

RESULTS AND DISCUSSION

Purification and properties of flavodoxin and determination of its N-terminal amino acid sequence. A homogeneous preparation of flavodoxin was obtained by the procedures described in Materials and Methods and subjected to N-terminal amino acid analysis by Edman degradation. This sequence and the oligonucleotide probe that was prepared on the basis of this sequence are shown in Fig. 1. Because alanine is frequently observed as the second amino acid in proteins from *E. coli* (9) and because peptide bonds between an N-terminal methionine and alanine are cleaved by methionine aminopeptidase (21), we reasoned that the DNA sequence encoding the flavodoxin gene would begin with an N-terminal methionine and synthesized our probe accordingly. Both threonine and glycine are encoded by four different codons, ACN and GGN, respectively. However, codon preferences in *E. coli* strongly favor C or U in the third position of the codons for both of these amino acids (15), and we reduced the degeneracy of our probe sequence by guessing that the flavodoxin gene would conform to the observed codon preferences. The isoleucine codon AUA is also rarely used in *E. coli* (15), and we reduced the degeneracy of our probe still further by assuming that isoleucine would be encoded only by AUU or AUC. The resulting 64-fold degenerate 20-mer oligonucleotide was called ECF1-7.

Purified homogeneous flavodoxin was shown to mediate the transfer of reducing equivalents from NADPH-flavodoxin oxidoreductase to methionine synthase in vitro, in agreement with earlier observations of Fujii and Huennekens (8). The purified protein was located on the gene-protein index of Neidhardt as described by Phillips et al. (22), on which it was assigned the α-numeric A19.0 and migrated at 121 × 45 mm. The isoelectric point of 4.64 determined by isoelectric focusing of the denatured protein indicates that flavodoxin is one of the most acidic proteins in *E. coli*. The estimated molecular weight of the protein from its mobility during polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate was 19,000, in good agreement with the molecular weight of 19,400 reported by Fujii and Huennekens (8) for flavodoxin.

Identification of the flavodoxin gene in the Kohara library. To our knowledge, degenerate probes have not previously been used to identify genes in the Kohara library. The more commonly used procedure (see, for example, the work of Baird and Georgopoulos [1]) is to clone the gene using the mini-Mu plasmid in vivo cloning technique (10) and then to map the gene using DNA hybridization to identify the gene in the Kohara library. This method requires the availability of mutant strains with a phenotype that can be complemented by a gene on the mini-Mu plasmid. When mutants are not available, a reverse genetics approach in which the gene is identified by DNA hybridization in an appropriate library is required. Use of the Kohara library for this purpose is ideal because identification of a transducing phage from the Kohara library containing the desired gene simultaneously maps and clones the gene. Furthermore, since an extensive restriction map of each phage insert is available, a subcloning strategy is immediately suggested.

Plaque hybridizations were performed with the 476 cosmid clones of the Kohara library by using the 5′-end-labeled oligonucleotide probe ECF1-7. Only two clones, 172 and 173, hybridized with the radiolabeled oligonucleotide probe. These two clones contain inserts of DNA mapping at about 16 min on the *E. coli* chromosome, and the left end of the insert in 173 overlaps with the right end of the insert in 172. We presumed that the flavodoxin gene was located in the region of the overlap. Southern blotting of an *Eco*RI digest of the DNA from Kohara clone 173 established that a 5.8-kb *Eco*RI fragment contained sequence complementary to the labeled oligonucleotide probe ECF1-7.

Subcloning and sequencing of the flavodoxin gene. DNA of ~5.8 kb was prepared by *Eco*RI digestion of DNA from Kohara clone 173 and ligated into the *Eco*RI site of the phagemid vector pBS2SK$^+$. *E. coli* DH5αF′ was transformed with the ligation mixture, and white colonies were selected on Luria broth plates containing ampicillin, X-Gal, and IPTG. Plasmid DNA was isolated from cells grown from 18 of the resultant colonies and subjected to Southern blotting with the 5′-end-labeled oligonucleotide ECF1-7. Figure 2A shows the results of electrophoretic analysis of the minipreps of plasmid DNA. It can be seen that significantly less DNA was obtained from colonies 1, 2, 5, 8, 9, 10, 11, and 18 than from the other colonies, even though the cell density of the inoculates from which DNA was isolated was approximately the same for all minipreps. In Fig. 2B, a radioautogram of the Southern blot is shown, and it can be seen that the N terminus of the flavodoxin gene is present only in those colonies that produced reduced amounts of plasmid DNA. There are two *Eco*RI fragments of the same 5.8-kb size in clone 173 from the Kohara library, and these were introduced into DH5αF′ with approximately equal frequency. Only one of the two fragments, that encoding the flavodoxin gene, led to restricted production of plasmid DNA. Analysis of the orientations of the inserts in pBS2SK$^+$ established that the fragment which does not encode flavodoxin was inserted into the phagemid vector in random orientations, while all eight clones that encode the N terminus of flavodoxin contained inserts in the same orientation. As will be shown, this orientation places the flavodoxin gene so that it is oriented in the direction opposite to the orientation of the *lac* promoter, and induction of synthesis of single-stranded DNA by infection with helper phage leads to production of the DNA strand with a sequence complementary to that of the mRNA encoding flavodoxin (the sense strand).

Phagemids with the *Eco*RI fragment in the opposite orientation were obtained by cutting out the insert by digestion

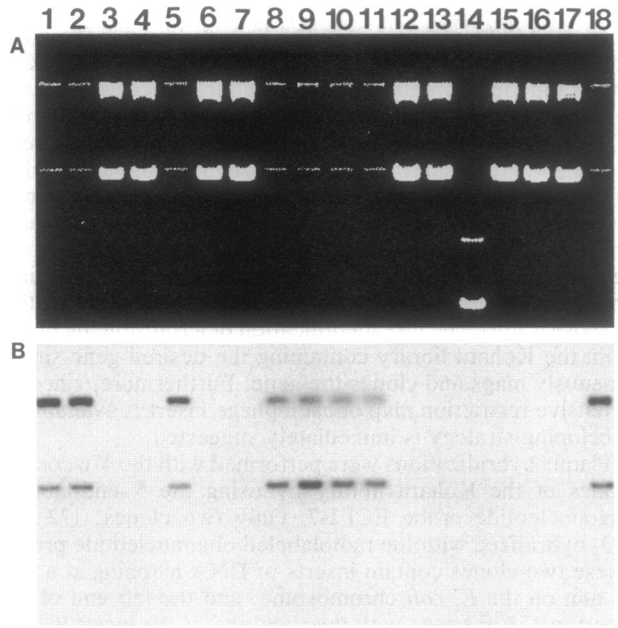

FIG. 2. The DNA from the region of the flavodoxin gene leads to apparent copy restriction of the plasmid on which it is located. DNA from Kohara phage 173 was isolated and cleaved with *Eco*RI, and the fragments were separated on an agarose gel. Fragments with an apparent size of 6 kb were isolated from the gel by electroelution and ligated into a pBS2SK$^+$ vector that had been cleaved with *Eco*RI. The restriction map published by Tabata et al. (28) indicates that two ~6-kb fragments of DNA are generated by *Eco*RI cleavage of DNA from this phage. Cells of strain DH5αF′ were transformed with the ligated DNA and selected for resistance to ampicillin. DNA minipreps were prepared from approximately equal numbers of cells derived from ampicillin-resistant colonies. (A) High-contrast photograph of an agarose gel loaded with samples of 18 DNA minipreps and visualized under UV light after staining with ethidium bromide. The miniprep in lane 14 was from cells transformed with the vector lacking an insert. (B) Southern blots of a nitrocellulose lift from the gel shown in panel A. 5′-end-labeled oligonucleotide probe ECF1-7 was used to probe the DNA. The probe hybridized to those colonies that displayed apparent copy restriction in panel A.

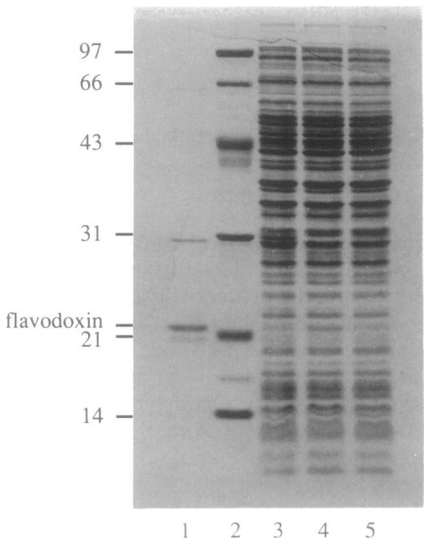

FIG. 3. Photograph of an analysis of total cellular proteins of *E. coli* by one-dimensional polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate. The proteins were visualized by staining with Coomassie blue dye. Lane 1, ~2 μg of purified flavodoxin from *E. coli*; lane 2, low-molecular-weight markers with the weights (10^3) indicated on the vertical axis; lanes 3 through 5, protein extracts from cells of strain DH5αF′ transformed with plasmids 3, 8, and 14, respectively, from Fig. 2. As demonstrated by Southern blotting and gel electrophoresis, plasmid 3 consists of pBS2SK$^+$ with an insert that does not contain the flavodoxin gene, plasmid 8 contains an insert encoding the flavodoxin gene, and plasmid 14 does not contain an insert. The two major bands at about 21,000 associated with the purified flavodoxin in lane 2 represent holoenzyme (containing noncovalently bound FMN even under denaturing conditions) and apoenzyme (lacking noncovalently bound FMN), with the apoenzyme migrating with an apparently lower molecular weight. The band at 31,000 is an impurity in this sample of flavodoxin.

with *Eco*RI and then religating. Of seven daughter plasmids analyzed, two contained the *Eco*RI insert in the opposite orientation. All daughter plasmids showed marked copy restriction as judged by the yield of plasmid DNA isolated from transformed cells. When single-stranded DNA production was initiated in the phagemids containing reverse inserts by the addition of helper phage, the cells lysed and no single-stranded DNA was obtained.

Protein extracts from colonies 8 (flavodoxin insert), 14 (no insert), and 3 (insert with no flavodoxin) were subjected to analysis by polyacrylamide gel electrophoresis in the presence of sodium dodecyl sulfate and were compared with a sample of purified flavodoxin from *E. coli*. The extract from colony 8, but not those from colonies 3 or 14, showed increased concentrations of a protein with the same mobility as purified flavodoxin (Fig. 3).

The plasmid pRM*Eco*RI was isolated from cells derived from colony 8 and was digested with *Eco*RV and then religated. The ligation mixture was used to transform strain DH5αF′, and transformants were selected on Luria broth plates supplemented with ampicillin, X-Gal, and IPTG. White colonies were isolated and subjected to restriction analysis. A typical plasmid was designated pRM*Eco*RV3 and is shown in Fig. 4. Southern blotting established that the ECF1-7 oligonucleotide hybridized to pRM*Eco*RV3.

Because very low yields of plasmid DNA were obtained from both the phagemids containing the *Eco*RI and *Eco*RV3 fragments, it was difficult to obtain enough DNA for double-stranded sequencing. Adequate quantities of single-stranded DNA could be obtained from pRM*Eco*RI, but only one of the two strands was synthesized, as noted above. By pooling several minipreps of pRM*Eco*RV3 DNA, and by using the M13 −20 primer, we were able to obtain sequence data from the 5′ end of the *Eco*RV3 fragment, and we used this to synthesize a nondegenerate 17-mer oligonucleotide sequencing primer. This in turn was used to sequence the single-stranded DNA from pRM*Eco*RI by primer extension. The sequencing strategy for the sense strand is outlined in Fig. 5.

Sequencing the antisense strand required an alternate approach. The pRM*Eco*RV3 phagemid was digested with *Hin*dIII and then religated, as shown in Fig. 4. This procedure results in the excision of the upstream portion of the flavodoxin gene. The resulting phagemid was designated pRMδ3′FLD. Yields of plasmid DNA from strain DH5αF′ transformed with this vector were approximately equal to the yields from cells transformed with pBS2SK$^+$ lacking an insert. A phagemid containing the upstream portion of the flavodoxin gene was prepared by subjecting DNA from

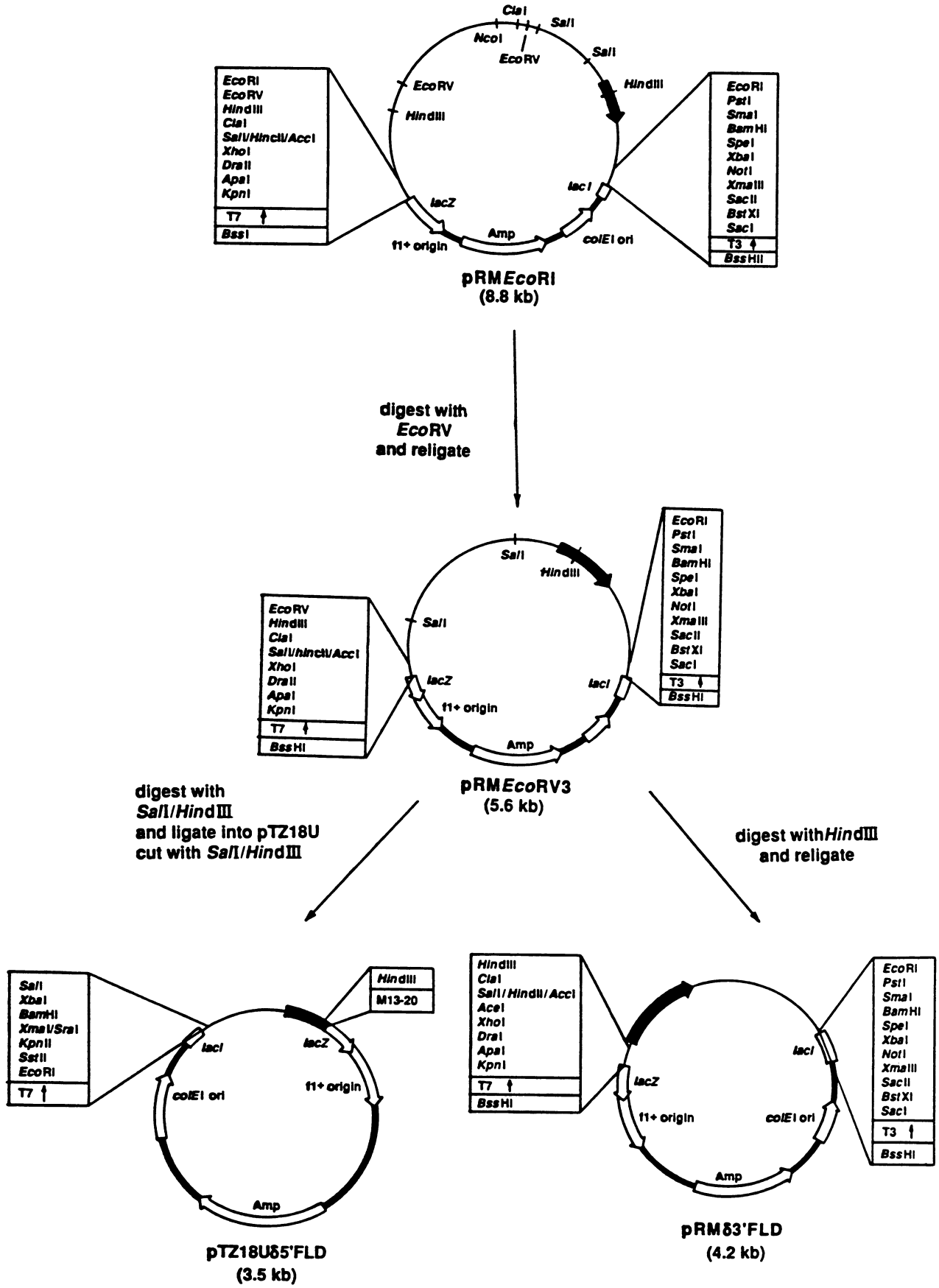

FIG. 4. Restriction map of pRM*Eco*RI and construction of plasmids pRM*Eco*RV3, pTZ18Uδ5'FLD, and pRMδ3'FLD. Details of the construction are given in Materials and Methods.

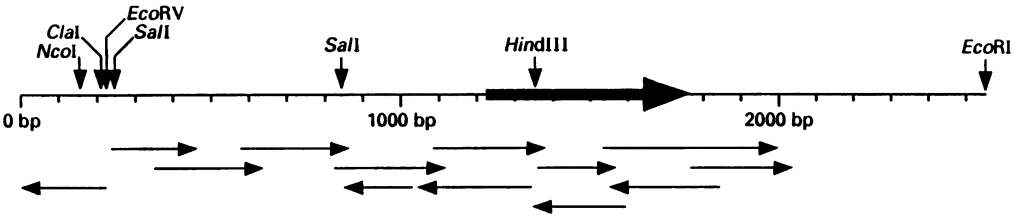

FIG. 5. Sequencing strategy for the flavodoxin gene and surrounding DNA. The entire ~2.5-kb *Eco*RV-*Eco*RI insert of pRM*Eco*RV3 is shown, as well as 228 nucleotides from the 3′ end of *Eco*RV1 that are contiguous to the *Eco*RV site. The lengths of the arrows are drawn to scale. The large arrow indicates the direction of transcription of the *fld* gene. For details, refer to Materials and Methods.

pRM*Eco*RV to digestion with *Hin*dIII and *Sal*I and ligating the restricted DNA into the vector pTZ18U that had previously been cut with *Hin*dIII and *Sal*I. The resulting phagemid was designated pTZ18Uδ5′FLD. Yields of plasmid DNA from cells transformed with this vector were also approximately equal to yields from cells transformed with the vector lacking an insert. Minipreps of double-stranded DNA were prepared from the phagemids containing the upstream and downstream portions of the flavodoxin gene and were sequenced on the antisense strand by primer extension. The strategy is indicated in Fig. 5.

The observation of copy restriction in plasmids containing the intact flavodoxin gene on a *Sal*I-*Eco*RI insert, but not in plasmids containing the *Sal*I-*Hin*dIII half of the same insert or the *Hin*dIII-*Eco*RI half, strongly suggests that the observed reduction in plasmid DNA content per cell is not a property of the cloned DNA per se but rather results from expression of the flavodoxin gene. These observations suggest that the *E. coli* flavodoxin gene is a "toxic gene" and were particularly surprising because expression of the flavodoxin gene from *A. nidulans* in strain DH5αF′ did not lead to copy restriction of the pBS2SK$^+$ plasmid containing this flavodoxin insert (data not shown), even though high levels of expression of the *A. nidulans* flavodoxin were observed.

```
TCGACCGGGACGACCGAAAACTAATCCGCTTTCGCGCGATGAACAGCTGCGTATTAATAA

                                                100
ACGCAACCAGCTAAAACGCGACAAAGTACGTGGCCTTAAGCGTGTCGAACTGAAGCTGAA

CGCGGAAGCTGTCGAGGCGCTGAACGAGCTGGCGGAGTCGCGCAATATGAGCCGTAGCGA

                                    200
ACTGATCGAAGAGATGTTAATGCAGCAACTGGCGGCGCTGCGTAGTCAGGGCATCGTTTA

                                                            300
ATTTCCACTTTCATGTAGCACAGTGTCGAGTCCTGCTCGTTTGCTGATTCCGCAAGACTG

CCTGTTCTGCTATGATTGCCTTTATCCGTGGGCAATTTTCCACCCCCATTTCAATAAGTT

                                                        400
TCAAGAGGTTATTTCACTC ATG GCT ATC ACT GGC ATC TTT TTC GGC AGC GAC ACC
                    Met Ala Ile Thr Gly Ile Phe Phe Gly Ser Asp Thr

GGT AAT ACC GAA AAT ATC GCA AAA ATG ATT CAA AAA CAG CTT GGT AAA GAC
Gly Asn Thr Glu Asn Ile Ala Lys Met Ile Gln Lys Gln Leu Gly Lys Asp

                                                    500
GTT GCC GAT GTC CAT GAC ATT GCA AAA AGC AGC AAA GAA GAT CTG GAA GCT
Val Ala Asp Val His Asp Ile Ala Lys Ser Ser Lys Glu Asp Leu Glu Ala

TAT GAC ATT CTG CTG CTG GGC ATC CCA ACC TGG TAT TAC GGC GAA GCG CAG
Tyr Asp Ile Leu Leu Leu Gly Ile Pro Thr Trp Tyr Tyr Gly Glu Ala Gln

                                                600
TGT GAC TGG GAT GAC TTC TTC CCG ACT CTC GAA GAG ATT GAT TTC AAC GGC
Cys Asp Trp Asp Asp Phe Phe Pro Thr Leu Glu Glu Ile Asp Phe Asn Gly
```

```
AAA CTG GTT GCG CTG TTT GGT TGT GGT GAC CAG GAA GAT TAC GCC GAA TAT
Lys Leu Val Ala Leu Phe Gly Cys Gly Asp Gln Glu Asp Tyr Ala Glu Tyr

                                                700
TTC TGC GAC GCA TTG GGC ACC ATC CGC GAC ATC ATT GAA CCG CGC GGT GCA
Phe Cys Asp Ala Leu Gly Thr Ile Arg Asp Ile Ile Glu Pro Arg Gly Ala

ACC ATC GTT GGT CAC TGG CCA ACT GCG GGC TAT CAT TTC GAA GCA TCA AAA
Thr Ile Val Gly His Trp Pro Thr Ala Gly Tyr His Phe Glu Ala Ser Lys

                                                800
GGT CTG GCA GAT GAC GAC CAC TTT GTC GGT CTG GCT ATC GAC GAA GAC CGT
Gly Leu Ala Asp Asp Asp His Phe Val Gly Leu Ala Ile Asp Glu Asp Arg

CAG CCG GAA CTG ACC GCT GAA CGT GTA GAA AAA TGG GTT AAA CAG ATT TCT
Gln Pro Glu Leu Thr Ala Glu Arg Val Glu Lys Trp Val Lys Gln Ile Ser

                                                900
GAA GAG TTG CAT CTC GAC GAA ATT CTC AAT GCC TGA TGT GATGCGGCGTAGACT
Glu Glu Leu His Leu Asp Glu Ile Leu Asn Ala  *   *

CATGTCTACGCCGTATTAATAGATAATGCCAATCAAAATAATTGCTACAAATTTGTAACTT
```

FIG. 6. Sequence of the *fld* gene and the predicted amino acid sequence. The nucleotide sequence is numbered from the *Sal*I site immediately upstream of the flavodoxin gene (Fig. 5). The deduced amino acid sequence is numbered with alanine as the first residue; this residue is N terminal in the mature protein. The first 14 residues from the N terminus were confirmed by Edman degradation of the purified flavodoxin protein. A putative Shine-Dalgarno sequence upstream of the flavodoxin gene is underlined.

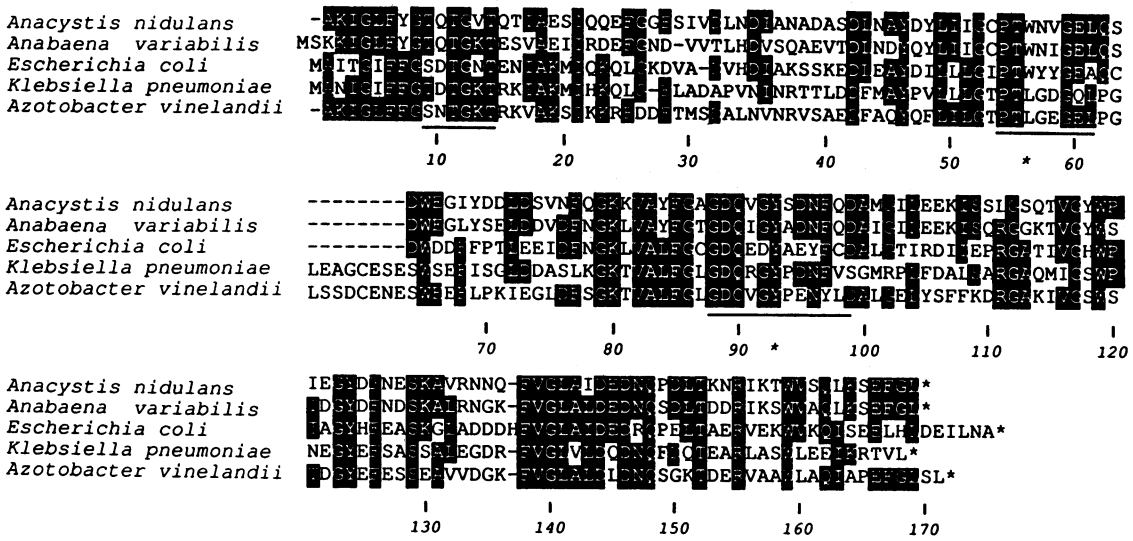

FIG. 7. Comparison of the sequences of five long flavodoxins from various prokaryotes. Identities in the sequences of three or more flavodoxins are highlighted on a black background. Asterisks indicate the residues that are known to flank the flavin in the X-ray structure of the *Anacystis nidulans* flavodoxin. Residues composing the FMN-binding site in the *A. nidulans* structure have been underlined. The numbering system is shown for the mature *E. coli* flavodoxin, in which alanine is the N-terminal residue.

We are currently attempting to construct an *E. coli* strain that lacks a flavodoxin gene in order to determine whether this protein plays a previously unrecognized essential function that might account for the intolerance of cells to even low levels of overexpression of the flavodoxin gene from *E. coli*. For this construction, we will use the method developed in S. R. Kushner's laboratory for the construction of strains with chromosomal mutations in essential genes (11).

The sequence of the *Eco*RV3 fragment 3′ to the *Sal*I site is shown in Fig. 6. An open reading frame with a deduced amino acid sequence showing strong homology to that of other flavodoxin genes was found between bp 380 and 907. All of the DNA encoding this open reading frame has been sequenced on both strands. The deduced amino acid sequence at the 5′ end of this open reading frame begins with a methionyl residue, and the next 14 residues are identical to those obtained by Edman degradation of the purified flavodoxin from *E. coli*. The *Hin*dIII site in the flavodoxin gene is located at bp 511.

Inspection of the nucleotide sequence immediately upstream of the flavodoxin gene in *E. coli* suggests the presence of a Shine-Dalgarno sequence, GAGG, 15 residues upstream of the start site of translation. Starting at position 311, the sequence TATGAT shows moderately good agreement with the consensus sequence for the −10 promoter, TATAAT. Twenty-one residues upstream from this sequence, at position 290, is the sequence CCGCAA, with only rather poor agreement with the consensus sequence TTGACA. There is an open reading frame upstream of the flavodoxin gene and in the same reading frame. This open reading frame extends from the *Sal*I site to bp 238. No significant homologies were found for this sequence in GenBank.

The deduced molecular weight of *E. coli* flavodoxin is 19,736, and after removal of the N-terminal methionine, this weight would be reduced to 19,606, in excellent agreement with the molecular weight of 19,000 determined by two-dimensional gel electrophoresis of the purified protein. The calculated isoelectric point of 4.02 is even more acidic than the isoelectric point of 4.64 estimated from the position of the denatured protein after isoelectric focusing and may indicate that the estimated isoelectric point is somewhat unreliable for very acidic proteins. The calculated amino acid composition is in excellent agreement with the composition determined by amino acid analysis of the purified protein (data not shown).

A comparison of the sequence of flavodoxin from *E. coli* with the sequences of flavodoxins from other prokaryotic organisms is shown in Fig. 7. Two types of flavodoxins have been isolated and sequenced in prokaryotes—the so-called short and long flavodoxins (20). Short flavodoxins are found in *Clostridium beijerinckii*, in *Megasphaera elsdenii*, and in *Desulfovibrio vulgaris* and typically have molecular weights of 14,500 to 16,300. Long flavodoxins have been found in *Anabaena variabilis*, *A. vinelandii*, *Chondrus crispus*, and *K. pneumoniae* and have molecular weights between 19,000 and 23,000. The flavodoxin from *E. coli* is a long flavodoxin, as judged both by its deduced molecular weight of 19,736 and by its close homology with other long flavodoxins. The homologies with other long flavodoxins include conservation of residues that form the flavin-binding site; these residues are underlined in Fig. 7. One of these conserved residues, tyrosine 93 in the *E. coli* deduced amino acid sequence, has been shown to flank the isoalloxazine ring system of the FMN prosthetic group in the *A. nidulans* enzyme (26). This residue is conserved in all five of the long-flavodoxin sequences shown in Fig. 7. The other residue flanking the isoalloxazine ring system in the *A. nidulans* flavodoxin is a tryptophan, and this residue is conserved in the *A. variabilis* and *E. coli* sequences (it is W56 in the deduced *E. coli* sequence) and is replaced by leucine in the *K. pneumoniae* and *A. vinelandii* sequences. Of particular interest is the presence of two consecutive tyrosine residues in positions 57 and 58 of the *E. coli* sequence. Position 57 is immediately adjacent to W56, which is one of the residues flanking the flavin, and in all other species, the homologous position is occupied by a glycine or an asparagine. The presence of two bulky aromatic residues in this part of the protein may affect

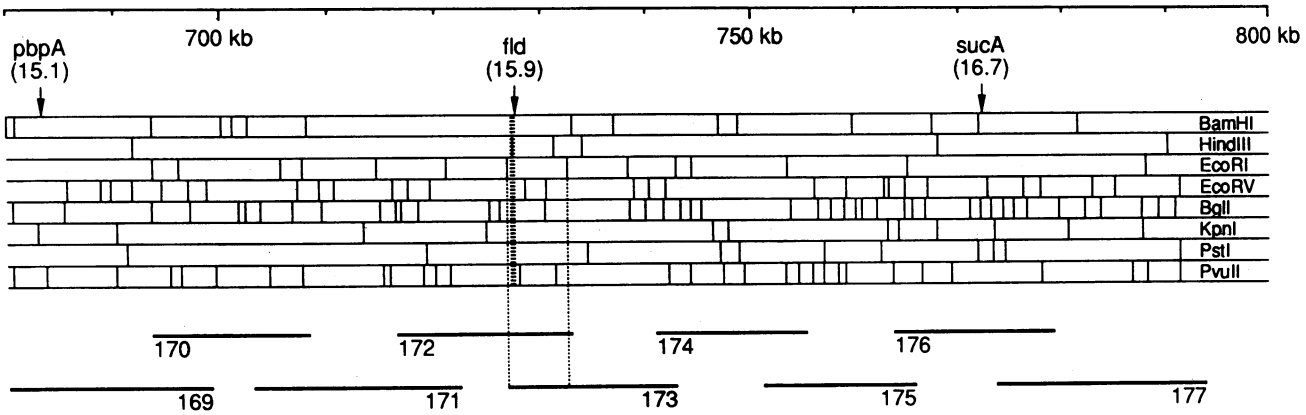

FIG. 8. Location of the flavodoxin gene in the Kohara *E. coli* genomic library. Location of the gene in the inserts common to phage 172 and 173 was established by Southern blotting with a probe to the N terminus of the flavodoxin gene. Location of the gene within that overlap region was established by restriction mapping and confirmation of the restriction sites by sequencing.

the ease with which the loop surrounding the flavin is able to rotate to provide hydrogen bonding stabilization of the flavin semiquinone (25) and may be responsible for the rather low potential for the oxidized flavin/flavin semiquinone couple in the *E. coli* enzyme (31).

From a comparison of the sequences of the *E. coli*, *A. variabilis*, and *A. vinelandii* flavodoxins, we observe differences in the distribution of charged residues that are expected to affect the charge distribution of the protein and hence to influence the recognition of flavodoxin by the proteins with which it interacts. As noted by Drummond (5), the nitrogen-fixing bacteria contain a cluster of basically charged residues in helix 1 of the protein, immediately adjacent to and overlapping with the conserved sequence GXXTGXT that forms the binding site for the phosphate residue of the FMN prosthetic group. The residues composing helix 1 of the *A. nidulans* flavodoxin correspond to residues 12 through 29 in the sequence of the mature *E. coli* protein. In the *A. nidulans* protein, this helix lies on the surface of the molecule, with its N terminus lying closest to the edge of the flavin that is exposed to solvent. In the residues corresponding to this helix in the flavodoxins of *K. pneumoniae* and *A. vinelandii*, there is a net charge of +4, whereas the net charge of the same region in the *E. coli* protein is +1. The net charge of this region is −3 in the *A. nidulans* protein and −2 in the *A. variabilis* protein. Immediately adjacent to tyrosine 93, one of the residues that flanks the isoalloxazine ring system in the long flavodoxins, are two negatively charged residues in positions 91 and 92 of the *E. coli* deduced sequence. These two negative charges are unique to the *E. coli* sequence and might be expected to alter the electrostatic potential at the flavin edge significantly. In the *A. nidulans* flavodoxin three-dimensional structure, the α-carbon of G93, which is homologous to D92 in the *E. coli* sequence, is only 0.87 nm from the 8-methyl group of the flavin. Two other residues, E145 and R147, are predicted to lie close to the flavin edge by comparison with the *A. nidulans* structure and differ in charge from the homologous residues of other long flavodoxins. Such differences in the pattern of charge distribution on the surface of the protein adjacent to the flavin edge may be responsible for the different specificities of these flavodoxins for their donor and acceptor partners that underlie their different physiological roles in their respective organisms.

Our data can be used to locate the gene encoding flavodoxin within the restriction map of the *E. coli* chromosome published by Tabata et al. (28). We have designated the gene encoding flavodoxin *fldA*. The *fldA* gene is present within the overlap region between Kohara transducing phages 172 and 173 and is located astride the leftmost *Hin*dIII site in phage 173, as shown in Fig. 8. We have confirmed the location of this gene in an *Eco*RI fragment derived from phage 173 that is ~5.8 kb in length and contains two *Eco*RV sites.

ACKNOWLEDGMENTS

This work was partially supported by NIH grant GM24908. Lee-may Chen was supported in part by a Medical Student Summer Research Fellowship with funds from NIH training grant HL 07690.

We thank Martha Ludwig and Cathy Luchinsky for helpful discussions during the course of this work.

REFERENCES

1. **Baird, L., and C. Georgopoulos.** 1990. Identification, cloning, and characterization of the *Escherichia coli sohA* gene, a suppressor of the *htrA* (*degP*) null phenotype. J. Bacteriol. **172**:1587–1594.
2. **Banerjee, R. V., S. R. Harder, S. W. Ragsdale, and R. G. Matthews.** 1990. Mechanism of reductive activation of cobalamin-dependent methionine synthase: an electron paramagnetic resonance spectroelectrochemical study. Biochemistry **29**: 1129–1135.
3. **Chisholm, D.** 1989. A convenient moderate-scale procedure for obtaining DNA from bacteriophage lambda. Biotechniques **7**:21–23.
4. **Deistung, J., and R. N. F. Thorneley.** 1986. Electron transfer to nitrogenase. Characterization of flavodoxin from *Azotobacter chroococcum* and comparison of its redox potentials with those of flavodoxins from *Azotobacter vinelandii* and *Klebsiella pneumoniae* (*nifF*-gene product). Biochem. J. **239**:69–75.
5. **Drummond, M. H.** 1986. Structure predictions and surface charge of nitrogenase flavodoxins from *Klebsiella pneumoniae* and *Azotobacter vinelandii*. Eur. J. Biochem. **159**:549–553.
6. **Eliasson, R., M. Fontecave, H. Jornvall, M. Krook, E. Pontis, and P. Reichard.** 1990. The anaerobic ribonucleoside triphosphate reductase from *Escherichia coli* requires S-adenosylmethionine as a cofactor. Proc. Natl. Acad. Sci. USA **87**:3314–3318.
7. **Frasca, V., R. V. Banerjee, W. R. Dunham, R. H. Sands, and R. G. Matthews.** 1988. Cobalamin-dependent methionine synthase from *Escherichia coli* B: electron paramagnetic resonance spectra of the inactive form and the active methylated form of the enzyme. Biochemistry **27**:8458–8465.
8. **Fujii, K., and F. M. Huennekens.** 1974. Activation of methionine

synthetase by a reduced triphosphopyridine nucleotide-dependent flavoprotein system. J. Biol. Chem. **249**:6745–6753.
9. **Gold, L., and G. Stormo.** 1987. Translational initiation, p. 1302–1307. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol. 2. American Society for Microbiology, Washington, D.C.
10. **Groisman, E. A., and M. J. Casadaban.** 1986. Mini-Mu bacteriophage with plasmid replicons for in vivo cloning and *lac* gene fusing. J. Bacteriol. **168**:357–364.
11. **Hamilton, C. M., M. Aldea, B. K. Washburn, P. Babitzke, and S. R. Kushner.** 1989. New method for generating deletions and gene replacements in *Escherichia coli*. J. Bacteriol. **171**:4617–4622.
12. **Knappe, J., F. A. Neugebauer, H. P. Blaschkowski, and M. Ganzler.** 1984. Posttranslational activation introduces a free radical into pyruvate formate-lyase. Proc. Natl. Acad. Sci. USA **81**:1332–1335.
13. **Knappe, J., and T. Schmitt.** 1976. A novel reaction of S-adenosyl-L-methionine correlated with the activation of pyruvate-formate lyase. Biochem. Biophys. Res. Commun. **71**:1110–1117.
14. **Knoelle, H.-E., and J. Knappe.** 1974. *Escherichia coli* ferredoxin, an iron-sulfur protein of the adrenodoxin type. Eur. J. Biochem. **50**:245–252.
15. **Konigsberg, W., and G. N. Godson.** 1983. Evidence for use of rare codons in the *dnaG* gene and other regulatory genes of *Escherichia coli*. Proc. Natl. Acad. Sci. USA **80**:687–691.
16. **Laudenbach, D. E., M. E. Reith, and N. A. Straus.** 1988. Isolation, sequence analysis, and transcriptional studies of the flavodoxin gene from *Anacystis nidulans* R2. J. Bacteriol. **170**:258–265.
17. **Mangum, J. H., and K. G. Scrimgeour.** 1962. Cofactor requirements and intermediates in methionine biosynthesis. Fed. Proc. **21**:242.
18. **Martin, A. E., B. K. Burgess, S. E. Ilsmaa, C. T. Smartt, M. R. Jacobson, and D. R. Dean.** 1989. Construction and characterization of an *Azotobacter vinelandii* strain with mutations in the genes encoding flavodoxin and ferredoxin I. J. Bacteriol. **171**:3162–3167.
19. **Massey, V., and P. Hemmerich.** 1980. Active-site probes of flavoproteins. Biochem. J. **8**:246–257.
20. **Mayhew, S. G., and M. L. Ludwig.** 1975. Flavodoxins and electron-transferring flavoproteins, p. 57–118. *In* P. D. Boyer (ed.), The enzymes, vol. 12B. Academic Press, Inc., New York.
21. **Miller, C. G.** 1987. Protein degradation and proteolytic modification, p. 680–691. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol. 1. American Society for Microbiology, Washington, D.C.
22. **Phillips, T. A., V. Vaughn, P. L. Bloch, and F. C. Neidhardt.** 1987. Gene-protein index of *Escherichia coli* K-12, edition 2, p. 919–966. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol. 2. American Society for Microbiology, Washington, D.C.
23. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
24. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions, p. 43–46. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. **Smith, W. W., R. M. Burnett, G. D. Darling, and M. L. Ludwig.** 1977. Structure of the semiquinone form of flavodoxin from *Clostridium M. P.* J. Mol. Biol. **117**:195–225.
26. **Smith, W. W., K. A. Pattridge, M. L. Ludwig, G. A. Petsko, D. Tsernoglou, M. Tanaka, and K. T. Yasanobu.** 1983. Structure of oxidized flavodoxin from *Anacystis nidulans*. J. Mol. Biol. **165**:737–753.
27. **Stubbe, J.** 1989. Protein radical involvement in biological catalysis? Annu. Rev. Biochem. **58**:257–285.
28. **Tabata, S., A. Higashitani, M. Takanami, K. Akiyama, Y. Kohara, Y. Nishimura, A. Nishimura, S. Yasuda, and Y. Hirota.** 1989. Construction of an ordered cosmid collection of the *Escherichia coli* K-12 W3110 chromosome. J. Bacteriol. **171**:1214–1218.
29. **Tarr, G. E.** 1986. Manual Edman sequencing system, p. 155–194. *In* J. E. Shively (ed.), Microcharacterization of polypeptides: a practical manual. Humana Press, Clifton, N.J.
30. **Taylor, R. T., and H. Weissbach.** 1969. *Escherichia coli* B N^5-methyltetrahydrofolate-homocysteine methyltransferase: sequential formation of bound methylcobalamin with S-adenosyl-L-methionine and N^5-methyltetrahydrofolate. Arch. Biochem. Biophys. **129**:728–744.
31. **Vetter, H., Jr., and J. Knappe.** 1971. Flavodoxin and ferredoxin of *Escherichia coli*. Hoppe-Seyler's Z. Physiol. Chem. **352**:433–446.