JOURNAL OF BACTERIOLOGY, Sept. 1989, p. 4728–4735
0021-9193/89/094728-08$02.00/0
Copyright © 1989, American Society for Microbiology

Vol. 171, No. 9

Isolation, Nucleotide Sequence, and Preliminary Characterization of the *Escherichia coli* K-12 *hemA* Gene

ELIZABETH VERKAMP* AND BARRY K. CHELM†

Department of Energy Plant Research Laboratory and Department of Microbiology and Public Health, Michigan State University, East Lansing, Michigan 48824-1312

Received 20 March 1989/Accepted 7 June 1989

The *Escherichia coli hemA* gene, essential for the synthesis of 5-aminolevulinic acid (ALA), was isolated and sequenced. The following criteria identified the cloned gene as *hemA*. (i) The gene complemented a *hemA* mutation of *E. coli*. (ii) The gene was localized to approximately 26.7 min on the *E. coli* chromosomal linkage map, consistent with the location of the mapped *hemA* locus. Furthermore, DNA sequence analysis established that the cloned gene lay directly upstream of *prfA*, which encodes polypeptide chain release factor 1. (iii) Deletion of the gene resulted in a concomitant requirement for ALA. The *hemA* gene directed the synthesis of a 46-kilodalton polypeptide in maxicell experiments, as predicted by the coding sequence. The DNA and deduced amino acid sequences of the *E. coli hemA* gene displayed no detectable similarity to the ALA synthase sequences which have been characterized from a variety of organisms, but are very similar to the cloned *Salmonella typhimurium hemA* sequences (T. Elliott, personal communication). Results of S1 nuclease protection experiments showed that the *hemA* mRNA appeared to have two different 5′ ends and that a nonoverlapping divergent transcript was present upstream of the putative distal *hemA* transcriptional start site.

The prosthetic groups of cytochromes, catalase, peroxidases, siroheme, and vitamin B_{12}, as well as active components of chlorophyll and bacteriochlorophyll molecules in photosynthetic organisms, are formed from tetrapyrrole derivatives. The first committed precursor of the common tetrapyrrole biosynthetic pathway is 5-aminolevulinic acid (ALA), which can be synthesized via two dissimilar routes. In one pathway, the condensation of succinyl-coenzyme A and glycine to form ALA is catalyzed by ALA synthase (EC 2.3.1.37). Genes which encode ALA synthase have been isolated from a number of procaryotic and eucaryotic organisms (6, 9, 10, 17, 24, 27, 30, 44, 46, 49), and based on sequence similarity, the enzyme displays significant evolutionary conservation (6, 32). A second pathway, in which ALA is derived from the intact 5-carbon skeleton of glutamate in a multienzymatic process requiring $tRNA^{Glu}$, exists in plant chloroplasts, *Chlamydomonas reinhardtii*, *Euglena gracilis*, cyanobacteria, and an anaerobic archaebacterium, *Methanobacterium thermoautotrophicum* (for review, see reference 22). To our knowledge, genes which encode the component enzymes of this pathway have not been cloned.

The focus of this investigation is the molecular characterization of ALA biosynthesis in *Escherichia coli*. Mutations of *E. coli* which cause a defect in ALA synthesis and therefore result in ALA auxotrophy under nonfermentative growth conditions have been isolated and mapped to two different loci on the linkage map, the *hemA* and *popC* loci at approximately 27 and 4 min, respectively (37, 43). *hemA* mutants were characterized initially as respiration deficient, small-colony variants that lack catalase and cytochromes due to their inability to synthesize heme (42). We describe in this paper the isolation and DNA sequencing of the *hemA* gene from *E. coli* and the preliminary characterization of its product. We also report a preliminary investigation into the organization of the transcriptional unit of the *hemA* gene with the eventual aim of studying the regulation of heme biosynthesis in *E. coli*.

* Corresponding author.
† Deceased 2 September 1987.

MATERIALS AND METHODS

Bacteriological procedures. *E. coli* strains were grown routinely on LB medium or M9 minimal medium (15). Carbon sources were added as indicated at the following concentrations: glucose, 0.010 M; acetate, 0.020 M; succinate, 0.016 M. Minimal medium was supplemented as required with L-amino acids at 40 μg/ml and thiamine at 1 μg/ml. Strains SASX41B and EV61 (Table 1) were grown routinely on medium supplemented with ALA (50 μg/ml). Antibiotics were incorporated into LB medium when indicated at the following concentrations: ampicillin, 80 μg/ml; kanamycin, 50 μg/ml; and tetracycline, 5 μg/ml. Anaerobic conditions were achieved by sparging the cultures with N_2 in closed, stoppered bottles for 30 min and incubating at 37°C in GasPak jars (BBL Microbiology Systems), and aerobic cultures were grown in flasks with vigorous shaking. Cells were harvested at the mid-log phase of growth.

For ALA synthase assays, the appropriate *E. coli* strains were inoculated into 400 ml of M9 glucose, succinate, or acetate medium with 4 ml of a mid-log-phase culture grown in the same medium and then incubated aerobically or anaerobically. The RNA samples used in S1 protection experiments were prepared from 500-ml cultures grown under aerobic or anaerobic conditions in M9 medium containing glucose or succinate.

P1 transductions were performed by methods described by Miller (35).

A method described by Winans et al. (48) was used to construct an *E. coli* strain with a deletion/insertion mutation in the *hemA* gene. The entire *hemA* gene and flanking sequences were present on a 6.0-kilobase-pair (kbp) *Hin*dIII fragment that was cloned into pUC19 to create pMR57 (Fig.

TABLE 1. *E. coli* strains

Strain	Genotype	Source or reference
CR63	*supD60 lamB63* λ⁻	5
HB101	*hsdS20* (r_B^- m_B^-) *recA13 leu-6 thi-1 ara-14 proA2 lacY1 galK2 rpsL20* (Sm^r) *xyl-5 mtl-1 supE44* λ⁻	11
DH5	*endA1 recA1 hsdR17* (r_K^- m_K^+) *supE44 thi-1 gyrA96 relA1*? λ⁻	19
JC7623	*thr-1 ara-14 leuB6* Δ(*gpt-proA*)*62 lacY1 tsx-33 supE44 galK2* Rac⁻ *hisG4 rfbD1 mgl-51 rpsL31 kdgK51 xyl-5 mtl-1 argE3 thi-1 recB21 recC22 sbcB15 sbcC201* λ⁻	5, 48
JM103	*hsdR4* Δ(*lac-pro*) (F' *traD36 proAB lacI*^{q}*Z*Δ*M15 thi strA supE endA sbcB*)	33
ORN125	*zcg*::Tn*10 thr-1 leuB thi-1* Δ(*argF-lac*)*U169 malA1 xyl-7 ara-13 mtl-2 gal-6 rpsL fhuA2 supE44*	45
SASX41B	HfrPO2A *hemA41 metB1 relA1*	B. Bachmann; CGSC 4806
EV61	Same as JC7623, except *hemA*	This study
MC1024	*araD139* Δ(*ara-leu*)*7697* Δ(*lacZ*)*M15 galU galK strA recA56 srl*::Tn*10*	13
EV10	Same as SASX41B, except *recA56 srl*::Tn*10*	This study
JK268	*trpE trpA dadR1 purB* (same as JK266)	18

1). Plasmid pMR61 is a derivative of pMR57 in which a 278-base-pair (bp) *Bgl*II fragment within the coding region of the *hemA* gene has been deleted and replaced by a 1.3-kbp *Bam*HI kanamycin resistance cassette derived from plasmid pRL161 (designated kanamycin cassette C.K1 in reference 16). The *nptII* gene is oriented in the same direction as the putative *hemA* open reading frame (ORFA in Fig. 1). Approximately 2 μg of this plasmid was digested with *Hin*dIII and transformed into *E. coli* JC7623 (Table 1). Transformants were selected on LB containing kanamycin and ALA and tested for the absence of the ampicillin resistance marker present on vector sequences and for ALA auxotrophy.

Nucleic acid methods. Southern hybridizations, small-scale and large-scale isolations of plasmid DNA, and genomic DNA preparations were performed according to described procedures (3). RNA was isolated by a CsCl gradient method as described before (2).

*Bal*31 exonuclease was used to generate a series of overlapping deletions spanning the DNA fragment to be sequenced (36). The appropriate DNA fragments were then cloned into M13mp18 or M13mp19 (50) with the *Bal*31-deleted end inserted nearest the primer-binding site in the *lacZ* coding sequence. Single-stranded DNA template was prepared in JM103 (33, 34) and sequenced by the method of Sanger et al. (41). The nucleotide sequence of both strands was determined at least once. Materials and enzymes used for sequencing were purchased from either Bethesda Research Laboratories or United States Biochemical Corporation and used according to the supplier's recommendations. End-labeled probes used in S1 protocols were sequenced by the method of Maxam and Gilbert (31).

The method of Berk and Sharp (8) as modified by Adams and Chelm (2) was used to map the 5' end of the *hemA* mRNA. For initial mapping, a 447-bp *Sal*I-*Bst*NI fragment that includes 144 bp of the *hemA* coding sequence and 303 bp of upstream sequence was isolated from a polyacrylamide gel, treated with calf intestinal alkaline phosphatase, and then 5'-end labeled with [γ-^{32}P]ATP by using T4 polynucleotide kinase. Single-stranded labeled fragments were separated on an 8% denaturing polyacrylamide gel, isolated, and hybridized to RNA by the procedures referenced above.

For better resolution of the 5' end of the *hemA* transcript, a *hemA*-specific single-stranded DNA probe was synthesized by a primer extension method described by Adams and

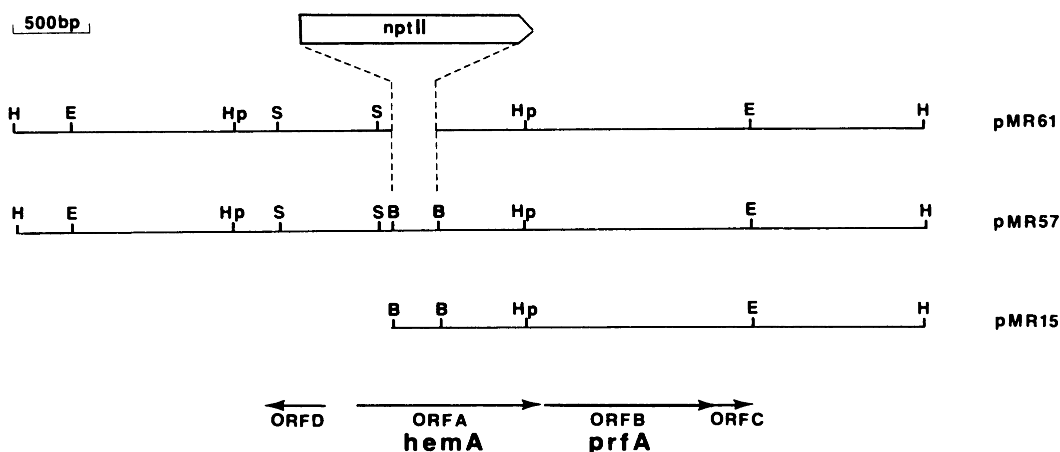

FIG. 1. Restriction endonuclease map of plasmid subclones of the isolated *E. coli hemA* region. The plasmid pMR15 is a subclone of cosmid pMRC2; the genomic DNA insert is carried on the cosmid vector pV35. The *Bgl*II site delineates one boundary of recombinant DNA present on the parent cosmid. The entire *hemA* and *prfA* genes and flanking sequences are carried on a 6.0-kbp *Hin*dIII DNA fragment that was cloned from pMRC5 into the unique *Hin*dIII site of pUC19 to yield plasmid pMR57. The location and direction of transcription of each ORF is indicated by a solid arrow; the DNA sequence was determined for the region delineated by these arrows. A 278-bp *Bgl*II fragment within the *hemA* coding region was deleted and replaced by an *nptII* cassette to create a plasmid, pMR61, that was used for gene replacement mutagenesis. The *Sal*I fragment shown in this figure was cloned into M13mp19 and used in the *hemA* high-resolution transcription mapping experiments described in Materials and Methods. Restriction enzymes: B, *Bgl*II; E, *Eco*RI; Hp, *Hpa*I; H, *Hin*dIII; S, *Sal*I.

```
-220                       GGGGCATAGTGATGACAAGTCCTTGAGATACGTTGCAGTT
                           CCCCGTATCACTACTGTTCAGGAACTCTATGCAACGTCAA
                           ORFD +1
                                                                hemA2
                                                                  *
-180 ATAACCCTTAATGCTAGCGTTACCGTCCGCTATCGTCTATGTTCAAGTTGTCTTAATTGC
     TATTGGGAATTACGATCGCAATGGCAGGCGATAGCAGATACAAGTTCAACAGAATTAACG
      ^^^

-120 CAGAATCTAACGGCTTTCGGCAATTACTCCAAAAGGGGGCGCTCTCTTTTATTGATCTTA
     GTCTTAGATTGCCGAAAGCCGTTAATGAGGTTTTCCCCCGCGAGAGAAAATAACTAGAAT

                           hemA1
                             *
 -60 CGCATCCTGTATGATGCAAGCAGACTAACCCTATCAACGTTGGTATTATTTCCCGCAGAC
     GCGTAGGACATACTACGTTCGTCTGATTGGGATAGTTGCAACCATAATAAAGGGCGTCTG

     hemA
  +1 ATGACCCTTTTAGCACTCGGTATCAACCATAAAACGGCACCTGTATCGCTGCGAGAACGT
      M  T  L  L  A  L  G  I  N  H  K  T  A  P  V  S  L  R  E  R
  61 GTATCGTTTTCGCCGGATAAGCTCGATCAGGCGCTTGACAGCCTGCTTGCGCAGCCGATG
      V  S  F  S  P  D  K  L  D  Q  A  L  D  S  L  L  A  Q  P  M
 121 GTGCAGGGCGGCGTGGTGCTGTCGACGTGCAACCGCACGGAACTTTATCTTAGCGTTGAA
      V  Q  G  G  V  V  L  S  T  C  N  R  T  E  L  Y  L  S  V  E
 181 GAGCAGGACAACCTGCAAGAGGCGTTAATCCGCTGGCTTTGCGATTATCACAATCTTAAT
      E  Q  D  N  L  Q  E  A  L  I  R  W  L  C  D  Y  H  N  L  N
 241 GAAGAAGATCTGCGTAAAAGCCTCTACTGGCATCAGGATAACGACGCGGTTAGCCATTTA
      E  E  D  L  R  K  S  L  Y  W  H  Q  D  N  D  A  V  S  H  L
 301 ATGCGTGTTGCCAGCGGCCTGGATTCACTGGTTCTGGGGGAGCCGCAGATCCTCGGTCAG
      M  R  V  A  S  G  L  D  S  L  V  L  G  E  P  Q  I  L  G  Q
 361 GTTAAAAAAGCGTTTGCCGATTCGCAAAAAGGTCATATGAAGGCCAGCGAACTGGAACGC
      V  K  K  A  F  A  D  S  Q  K  G  H  M  K  A  S  E  L  E  R
 421 ATGTTCCAGAAATCTTTCTCTGTCGCGAAACGCGTTCGCACTGAAACAGATATCGGTGCC
      M  F  Q  K  S  F  S  V  A  K  R  V  R  T  E  T  D  I  G  A
 481 AGCGCTGTGTCTGTCGCTTTTGCGGCTTGTACGCTGGCGCGGCAGATCTTTGAATCGCTC
      S  A  V  S  V  A  F  A  A  C  T  L  A  R  Q  I  F  E  S  L
 541 TCTACGGTCACAGTGTTGCTGGTAGGCGCGGGCGAAACTATCGAGCTGGTGGCGCGTCAT
      S  T  V  T  V  L  L  V  G  A  G  E  T  I  E  L  V  A  R  H
 601 CTGCGCGAACACAAAGTACAGAAGATGATTATCGCCAACCGCACTCGCGAACGTGCCCAA
      L  R  E  H  K  V  Q  K  M  I  I  A  N  R  T  R  E  R  A  Q
 661 ATTCTGGCAGATGAAGTCGGCGCGGAAGTGATTGCCCTGAGTGATATCGACGAACGTCTG
      I  L  A  D  E  V  G  A  E  V  I  A  L  S  D  I  D  E  R  L
 721 CGCGAAGCCGATATCATCATCAGTTCCACCGCCAGCCCGTTACCGATTATCGGGAAAGGC
      R  E  A  D  I  I  I  S  S  T  A  S  P  L  P  I  I  G  K  G
 781 ATGGTGGAGCGCGCATTAAAAAGCCGTCGCAACCAACCAATGCTGTTGGTGGATATTGCC
      M  V  E  R  A  L  K  S  R  R  N  Q  P  M  L  L  V  D  I  A
 841 GTTCCGCGCGATGTTGAGCCGGAAGTTGGCAAACTGGCGAATGCTTATCTTTATAGCGTT
      V  P  R  D  V  E  P  E  V  G  K  L  A  N  A  Y  L  Y  S  V
 901 GATGATCTGCAAAGCATCATTTCGCACAACCTGGCGCAGCGTAAAGCCGCAGCGGTTGAG
      D  D  L  Q  S  I  I  S  H  N  L  A  Q  R  K  A  A  A  V  E
 961 GCGGAAACTATTGTCGCTCAGGAAACCAGCGAATTTATGGCGTGGCTGCGAGCACAAAGC
      A  E  T  I  V  A  Q  E  T  S  E  F  M  A  W  L  R  A  Q  S
1021 GCCAGCGAAACCATTCGCGAGTATCGCAGCCAGGCAGAGCAAGTTCGCGATGAGTTAACC
      A  S  E  T  I  R  E  Y  R  S  Q  A  E  Q  V  R  D  E  L  T
1081 GCCAAAGCGTTAGCGGCCCTTGAGCAGGGCGGCGACGCGCAAGCCATTATGCAGGATCTG
      A  K  A  L  A  A  L  E  Q  G  G  D  A  Q  A  I  M  Q  D  L
1141 GCATGGAAACTGACTAACCGCTTGATCCATGCGCCAACGAAATCACTTCAACAGGCCGCC
      A  W  K  L  T  N  R  L  I  H  A  P  T  K  S  L  Q  Q  A  A
1201 CGTGACGGGGATAACGAACGCCTGAATATTCTGCGCGACAGCCTCGGGCTGGAGTAGCAG
      R  D  G  D  N  E  R  L  N  I  L  R  D  S  L  G  L  E  -
1261 TACATCATTTTCTTTTTTTACAGGGTGCATTTACGCCTATGAAGCCTTCTATCGTTG
                                             +1 prfA
```

FIG. 2. Nucleotide sequence of the *hemA* gene and the predicted amino acid sequence of the *hemA* polypeptide. The sequence of each DNA strand is shown for the region that separates *hemA* from the putative translational start site of the divergent ORFD. The proposed transcriptional start sites of the two major *hemA* RNA species are indicated by asterisks, and sequences which constitute possible promoter elements are underlined. The likely promoter sequences of the divergent ORFD transcript are underscored with dashed lines, and arrows indicate the nucleotides which correspond to the 5′ end of the ORFD transcript. The nucleotide sequence which corresponds to the reverse complement of the synthetic oligonucleotide used for high-resolution S1 mapping of the *hemA* transcript is bracketed. This sequence has been deposited in the EMBL/GenBank data base (accession no. M25323).

Chelm (1). A synthetic oligonucleotide, 5′-ACGATACACG TTCTCGC-3′ (Fig. 2), was labeled at its 5′ end and annealed to a single-stranded template prepared from a recombinant M13mp19 phage containing the 675-bp *Sal*I fragment shown in Fig. 1. The hybridized primer was extended, permitting the synthesis of an end-labeled DNA strand that was complementary to the *hemA* message. This labeled probe was purified and used in S1 nuclease protection experiments as described above. The same oligonucleotide primer and single-stranded DNA template were used in a dideoxy chain termination reaction to generate a sequence ladder.

Cosmid libraries. The methods used to prepare cosmid libraries were modified from those published (3). Genomic DNA from CR63 was partially digested with *Sau*3A1 and size-fractionated. Fragments greater than approximately 20 kbp were ligated to the *Bgl*II cloning sites in pV35, a cosmid vector that contains the origin of replication and the ampi-Cosmids were plated on SASX41B in the presence of exogenous ALA and then screened for complementation of the *hemA* mutation.

A second cosmid library was prepared from *E. coli* ORN125, which possesses a transposon Tn*10* insertion at 26.5 min that is approximately 82% cotransducible with the *hemA* locus (45). Purified DNA was partially digested with *Sau*3A and cloned into the unique *Bam*HI site in pWH4 (21). Recombinant transformants of HB101 were selected on LB plus kanamycin and then screened for tetracycline resistance and complementation of the *hemA* mutation.

Maxicell experiments. Plasmid-encoded polypeptides were identified by the maxicell technique of Sancar et al. (40) except that *E. coli* HB101 and DH5 were used as the plasmid hosts. Plasmid pMR81 consists of a 2.5-kbp *Bgl*I-*Pvu*I insert that includes ORFA plus 895 bp of the downstream *prfA* coding region and 336 bp of upstream sequence (Fig. 1) cloned as a blunt-ended fragment into the unique *Hin*cII site of pUC19. This plasmid was digested with *Bgl*II and religated to yield pMR81ΔBgl, a plasmid with a deletion of the 278-bp *Bgl*II fragment within the ORFA coding sequence.

Enzyme assays. ALA synthase activity was assayed as described by Guerinot and Chelm (17) for 2 h at 37°C in a 1-ml reaction mix containing 0.5 to 2.0 mg of protein. Cells were pelleted by centrifugation at 4,000 × *g* for 10 min at 4°C, washed one time with HEPES buffer (200 mM HEPES [*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid, pH 8.0], 33 mM $MgCl_2$, 1 mM 2-mercaptoethanol, 10 mM pyridoxal 5-phosphate) and suspended in 3 to 5 ml of HEPES buffer. Cells were disrupted at 4°C by two passages through a French pressure cell at 12,000 lb/in^2, and cell debris was removed by centrifugation at 17,000 × *g* for 20 min at 4°C. The supernatants were used for determination of protein concentration by the method of Lowry et al. (29) and for the measurement of enzyme activity. ALA synthase assays were also performed by the method of Burnham (12) as modified by Tai et al. (46).

RESULTS

Isolation of the *hemA* gene. Two schemes were devised to isolate the *hemA* gene on a cosmid vector. The first approach was the complementation of the *hemA* mutation in SASX41B with a cosmid gene bank prepared from *E. coli* CR63. The second strategy was to clone a Tn*10* tetracycline resistance marker that was inserted adjacent to *hemA* on the *E. coli* chromosome in strain ORN125. We reasoned that some of the transformants which harbored Tn*10* on a recombinant cosmid would also contain *hemA*. The desired

cosmids were isolated as described in Materials and Methods; one cosmid obtained by the first screen (pMRC2) and two cosmids isolated from the second (pMRC5 and pMRC10) were characterized in detail. The apparent complementation of the *hemA* mutation by cosmid pMRC2 required that transformants be selected initially in the presence of exogenously added ALA. Transformants could then grow in the absence of ALA at rates comparable to the wild type, suggesting that recombination was required for complementation. This was confirmed by showing that the apparent complementation did not occur in strain EV10, a derivative of SASX41B made RecA$^-$ by P1 transduction from MC1024 (Table 1). In contrast, pMRC5 and pMRC10 complemented SASX41B with no requirement for prior selection in the presence of ALA or recombination function. Plasmid subclones of pMRC2 (pMR15) and pMRC10 (pMR57) were obtained (Fig. 1), and these exhibited the complementation characteristics of the parent cosmids described above. The presence of common restriction enzyme sites between the plasmid subclones suggested, and Southern hybridization experiments verified, that the DNA sequences isolated by the two different schemes were homologous (data not shown). The plasmid pMR57 contained the recombinant DNA insert of pMR15, along with additional flanking DNA that appeared to be necessary for *hemA* activity.

DNA sequence analysis. DNA sequencing of the cloned insert containing the putative *hemA* gene revealed two large open reading frames (ORFs), herein designated A and B. These two ORFs, along with their coding capacity and direction of transcription, are shown schematically in Fig. 1. Two other ORFs of undetermined lengths (ORFC and ORFD in Fig. 1) were also located in this region, but these have not been characterized further.

A computer search of GenBank sequences disclosed that ORFB corresponded to *prfA*, the gene that encodes peptide chain release factor 1 (RF1), which had been isolated and sequenced previously (14). Since both *prfA* and *hemA* map to about 27 min on the genetic linkage map (26, 39, 43), the cloned DNA sequences came from the *hemA* region of the chromosome. If ORFA corresponds to the *hemA* gene, it would explain why pMR15 does not, by itself, display *hemA* activity, since it lacks a portion of this ORF.

The coding sequence of the putative *hemA* gene, ORFA, and upstream and downstream flanking sequences are presented in Fig. 2. Assuming that the coding region begins at the first ATG, ORFA could theoretically code for a polypeptide consisting of 418 amino acids with a molecular weight of 46,312. The 41-bp intergenic region separating the ORFA stop codon and the first codon of *prfA* are also presented in the figure.

Maxicell experiments. Maxicell labeling experiments were performed in order to determine whether ORFA produced a polypeptide of the anticipated size. A polypeptide of approximately 46 kilodaltons was synthesized by the plasmid pMR81 (Fig. 3). No other ORFs were present on this plasmid which had a coding capacity sufficiently large to encode this polypeptide. Furthermore, an identical plasmid that was deleted for the 278-bp *Bgl*II fragment within ORFA (pMR81ΔBgl) failed to synthesize this polypeptide (Fig. 3). These experiments showed that ORFA can direct the synthesis of a polypeptide of a size in agreement with that predicted by the DNA coding sequence.

Construction of *hemA* mutant. The following experiments demonstrated that the 46-kilodalton polypeptide encoded by ORFA is required for ALA synthesis. The mutagenesis

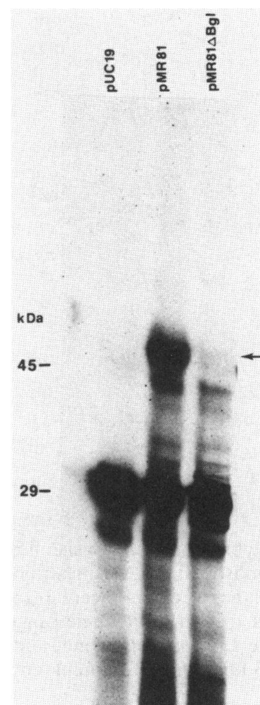

FIG. 3. Autoradiogram of plasmid polypeptides labeled in the maxicell procedure. The [^{35}S]methionine-labeled extracts were analyzed on a 10% sodium dodecyl sulfate-polyacrylamide gel; the plasmids harbored by the host *E. coli* DH5 are identified above the appropriate lanes and described in Materials and Methods. The electrophoretic mobility of the protein molecular mass markers egg albumin (45 kilodaltons) and carbonic anhydrase (29 kilodaltons) are indicated. The arrow identifies the position of the 46-kilodalton polypeptide (middle lane) that is coded for by the *hemA* gene. A polypeptide of approximately 40 kilodaltons that probably corresponds to the truncated RF1 protein is synthesized by both plasmids. The intact RF1 protein migrates as a 48-kilodalton polypeptide in this gel system (47; our unpublished results). In addition, a band that migrates below the 28,000-dalton β-lactamase protein is unique to pMR81ΔBgl and probably represents a product of the deleted *hemA* gene.

technique described in Materials and Methods was used to introduce the inactivated copy of ORFA on plasmid pMR61 into the genome, replacing the chromosomal copy of ORFA. Fourteen of the 14 kanamycin-resistant, ampicillin-sensitive transformants selected by this procedure were ALA auxotrophs, and one of these strains, designated EV61, was characterized further. In parallel experiments, we were unable to introduce into the genome a *hemA* interruption in which the *nptII* cassette was inserted in the opposite orientation (see Discussion).

Total genomic DNA was extracted from EV61 and JC7623 and used in Southern hybridization studies in order to verify that the expected gene replacement event had occurred. The DNA samples were digested with *Eco*RI, size-fractionated on 1% agarose gels, transferred to cellulose nitrate, and hybridized to two different DNA probes: the 1.3-kbp *nptII* gene and a 1.9-kbp *Hpa*I fragment which contains the major portion of ORFA along with 700 bp of upstream DNA sequences (Fig. 1). The *nptII* probe hybridized to a 5.2-kpb *Eco*RI fragment of DNA derived from the mutant but did not hybridize to DNA isolated from the parent strain (Fig. 4, lanes 1 and 2). The *Hpa*I probe hybridized to a 4.2-kbp

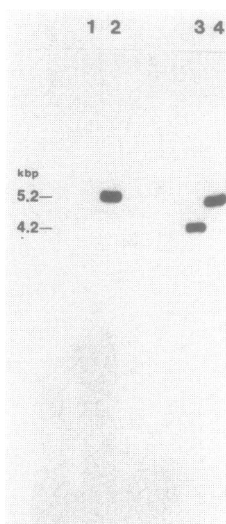

FIG. 4. Southern hybridization analysis of EV61 genomic DNA. *Eco*RI digests of chromosomal DNA from strain JC7623 (lanes 1 and 3) and strain EV61 (lanes 2 and 4) were transferred to cellulose nitrate and hybridized to two radiolabeled probes. Lanes 1 and 2 were hybridized to the 1.4-kbp *nptII* gene, and lanes 3 and 4 were hybridized to the 1.8-kbp *Hpa*I fragment containing the *E. coli hemA* gene.

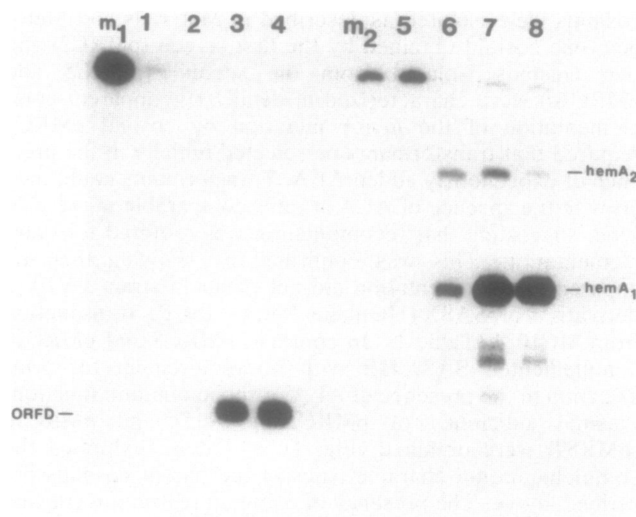

FIG. 5. S1 nuclease protection analysis. The probes used in S1 nuclease protection experiments were each of the 5′-end-labeled strands of the 447-bp *Bst*NI-*Sal*I fragment described in Materials and Methods. Each probe was hybridized to the following RNA types: aerobic M9-succinate (lanes 2 and 6); aerobic M9-glucose (lanes 3 and 7); and anaerobic M9-glucose (lanes 4 and 8). Control lanes: lanes m_1 and m_2 contain the probes alone; lanes 1 and 5 contain the labeled probes hybridized with salmon sperm DNA.

*Eco*RI fragment in DNA derived from strain JC7623, but, as expected, this fragment was shifted to 5.2 kbp in DNA derived from strain EV61 (Fig. 4, lanes 3 and 4). This analysis showed clearly that the wild-type chromosomal copy of ORFA was absent in EV61 and had been replaced by the inactive copy.

P1 transductions. If ORFA corresponds to the *hemA* gene, then the kanamycin resistance marker present in EV61 should map at approximately 26.7 min and should be cotransduced with *trp* at approximately 27.5 min (4). To test this, a P1 lysate was prepared from JK268 (*trpA trpE hemA*$^+$) and used to transduce strain EV61, selecting for HemA$^+$ (ability to form normal colonies on LB with no ALA supplement) and then testing for the Trp$^-$ phenotype. The *hemA* and *trp* markers were cotransduced at a frequency of 15.4% (10 of 65), and all of the HemA$^+$ transductants were sensitive to kanamycin. These genetic data corroborated the DNA sequence analysis that located ORFA in the 26.7-min region of the *E. coli* linkage map. For unknown reasons, we were unable to transduce the kanamycin resistance marker from EV61 into various recipient strains (see Discussion).

ALA synthase assays. Extracts from various strains of *E. coli*, including strains which harbored the putative *hemA* gene on high-copy-number plasmids, were tested for ALA synthase activity by the protocols described in Materials and Methods. We were unable to detect activity in any of these strains. However, ALA synthase activity could be readily demonstrated by these methods in an *E. coli* strain which contained the *Bradyrhizobium japonicum hemA* gene cloned into pUC9 (data not shown), indicating that the protocols were adequate for ALA synthase detection. Reports of ALA synthase activity in *E. coli* are inconsistent (44, 46), and recent experiments indicate that ALA is synthesized by the 5-carbon glutamate route in *E. coli* (see below). Additional experiments are required to resolve this question.

Determination of the transcriptional start site of the *hemA* gene. We mapped the 5′ end of the *hemA* mRNA by S1 nuclease protection methods with the *Sal*I-*Bst*NI probe described in Materials and Methods. Each end-labeled strand of probe was hybridized to RNA prepared from cultures of *E. coli* CR63 grown under several different conditions, selected to represent various energy-generating modes of metabolism. The results of one set of experiments are presented in Fig. 5. Two RNA species of different sizes protected the probe corresponding to one of the DNA strands (lanes 6, 7, and 8), and one RNA species protected the probe corresponding to the opposite strand (lanes 3 and 4; a band in lane 2 is present but faint). These results suggested that within the DNA encompassed by the *Sal*I-*Bst*NI fragment, there were two divergent transcriptional units.

Electrophoresis of each of the protected products resulting from S1 nuclease experiments adjacent to a Maxam-Gilbert sequencing ladder generated from its corresponding full-length probe enabled identification of the 5′ end points of the presumptive transcriptional start sites (data not shown). Results of high-resolution S1 nuclease mapping of the *hemA* transcript are presented in Fig. 6. The nucleotides which correspond to the 5′ ends of the *hemA* message and the divergent ORFD transcript are indicated in Fig. 2. The 5′ ends of the *hemA* transcript were separated by 92 nucleotides and were evident in each of four different RNA preparations and in experiments with several different probes (Fig. 5 and 6, and data not shown). It is therefore unlikely that they are due to artifacts associated with the S1

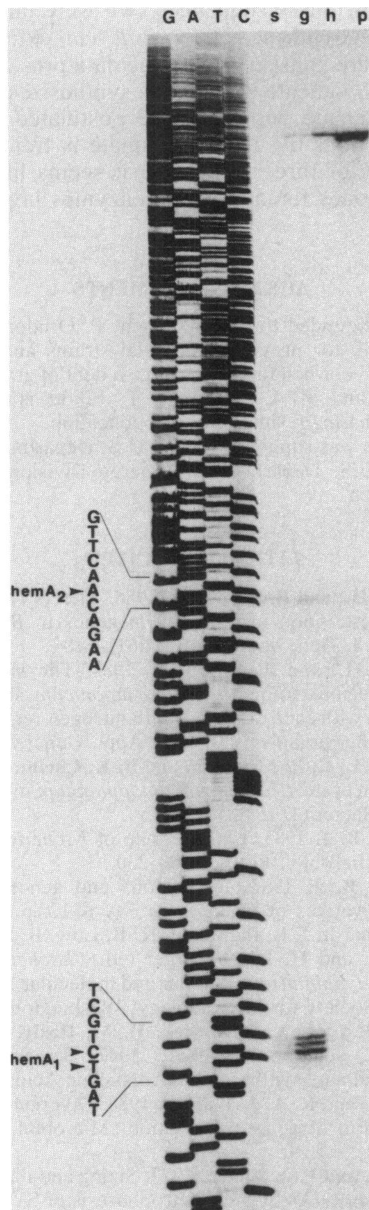

FIG. 6. High-resolution S1 mapping of the *hemA* mRNA transcript. A single-stranded DNA probe was generated by primer extension of the *hemA*-specific synthetic oligonucleotide (see Materials and Methods) and hybridized to 20 μg of RNA. The protected DNA fragments of S1 nuclease digestion were size-fractionated next to a DNA sequencing ladder which had been produced by primer extension of the same oligonucleotide. Lane s, Probe hybridized to RNA extracted from *E. coli* cultures grown aerobically on M9 medium plus succinate; lane g, probe hybridized to RNA extracted from *E. coli* cultures grown aerobically on M9 medium plus glucose; lane h, probe hybridized to an equal amount of heterologous denatured salmon sperm DNA; lane p, probe not subjected to S1 nuclease digestion. Nucleotides depicted on the left correspond to the sequence of the *hemA* coding strand. Nucleotides which correspond to the most likely 5′ ends of *hemA* mRNA transcripts are indicated by arrowheads.

nuclease digestion or RNA isolation procedure. We cannot determine from these data whether there are two functional transcriptional start sites for the *hemA* message or whether some RNA processing (cleavage) event results in the two different RNA species.

The apparent start sites of the distal *hemA* transcript and the ORFD transcript did not overlap, but were separated by approximately 45 nucleotides (Fig. 2). Whether these sequences play a role in the regulation of either transcript has not yet been addressed experimentally.

DISCUSSION

We report in this investigation the isolation and DNA sequence of the *hemA* gene from *E. coli*. The cloned gene was identified as *hemA* based on the following results.

(i) The gene complemented the *hemA* mutation of SASX41B. The apparent complementation of the mutant *hemA* allele by truncated copies of the cloned gene required RecA function, indicating that the cloned DNA corresponded to *hemA* and eliminating the possibility that the cloned gene suppressed the mutant phenotype of SASX41B by a mechanism involving a different locus.

(ii) The gene was localized to approximately 26.7 min on the *E. coli* linkage map, the site mapped previously for the *hemA* locus. Sequence analysis of the cloned DNA disclosed that the *hemA* gene lay directly upstream of and was transcribed in the same direction as *prfA*, which encodes the RF1 protein and maps to 26.7 min. Transduction mapping with P1 also localized the cloned gene to the 27-min region of the chromosome, and in addition, the restriction map of the DNA sequence surrounding the isolated *hemA* gene corresponds to the chromosomal restriction map at 26.7 min (25).

(iii) Inactivation of the chromosomal homolog of the cloned gene created an ALA auxotroph. Strain EV61 contained a defined insertion in the *hemA* gene that removed 278 nucleotides of the coding sequence. This strain was respiration deficient and lacked catalase activity, but formed small colonies on agar plates if provided with a fermentable carbon source. The mutant phenotype was completely reversible when the medium was supplemented with ALA. Thus, the *hemA* gene probably does not encode a polypeptide vital to the survival of *E. coli*, although the possibility that strain EV61 harbors a compensatory suppressor mutation or that some ALA is synthesized via an alternate pathway has not been ruled out by these experiments.

The inability to construct a gene replacement mutant in which the *nptII* cassette is transcribed in a direction opposite to that of *hemA* implies that this insertion may be polar on the downstream *prfA* gene. This effect may be less severe or absent when the *nptII* gene is transcribed in the same direction as *prfA*, perhaps because the *nptII* gene provides a new promoter for *prfA*. Attempts to create chromosomal deletions that extended into the *prfA* coding region were unsuccessful (data not shown), suggesting that RF1 is essential for the viability of *E. coli*. This would not be surprising in light of its role in polypeptide chain termination and also because *prfB*, which encodes peptide chain release factor 2 (RF2), is indispensable (23). These observations, in conjunction with the inability to transduce the *hemA* insertion from EV61 into different backgrounds, suggest that some insertions in *hemA* may be polar on *prfA*. It is possible that EV61 possesses a secondary mutation that suppresses the polar effect on *prfA* expression.

The computer analysis program of Pustell and Kafatos (38) was used to compare the DNA and amino acid sequences of the *E. coli hemA* gene and ALA synthase sequences which had been reported from several organisms. The *E. coli hemA* sequence exhibited no significant similarities with the ALA synthase sequences of *B. japonicum*, chicken, mouse, or human (data not shown). This observation was unexpected,

as substantial similarities exist between *B. japonicum* and chicken (30), mouse, and human (our unpublished results) and among the various ALA synthase sequences of eucaryotic origin (6). However, the *E. coli hemA* gene exhibits extensive similarities with the DNA and amino acid sequences of the cloned *Salmonella typhimurium hemA* gene (T. Elliott, personal communication). The predicted amino acid sequences share 94.3% identical residues over the entire length of the polypeptides, and the amino acids are 96.9% conserved (data not shown).

This lack of similarity to cloned ALA synthase sequences, coupled with our inability to detect ALA synthase activity in *E. coli* cell extracts, suggested that the primary route of ALA synthesis in *E. coli* might not be via ALA synthase. Interestingly, the *B. japonicum hemA* gene cloned into pUC9 complemented the *E. coli hemA* strain SASX41B. This result implies that ALA synthase can function in *E. coli* but does not prove that this is the usual means of ALA biosynthesis in this organism. A recent report (28) presents convincing evidence that the 5-carbon glutamate pathway operates in *E. coli*, which would be consistent with our results and which would suggest that the *hemA* gene may code for one of the enzymes involved in this pathway.

The results of S1 nuclease protection experiments suggest that transcription of the *hemA* message may initiate at two different sites or, alternatively, that there may be a processing event at the 5′ nontranslated end of the message. Two protected fragments of different sizes were apparent when experiments were performed with RNA samples isolated from cells grown under several different cultural conditions with respect to carbon substrate and anaerobicity. It will now be possible to examine whether transcription of *hemA* is regulated under various growth conditions and whether *hemA* expression plays a role in the regulation of heme biosynthesis.

Nucleotides upstream of each of the transcriptional start sites can be identified that display similarity to the canonical -10 and -35 sequences of *E. coli* σ^{70} consensus promoter elements (20) (Fig. 2). For *hemA1*, two different hexanucleotides are denoted in Fig. 2 that show similarity to the -35 promoter element. Although the sequence proximal to *hemA1* is better matched to the consensus, it is not spaced appropriately with respect to the -10 element and thus may not function with the designated -10 element as an RNA polymerase-binding site. Whether these or other sequences constitute functional promoters has not been addressed in this study. Regions of divergent transcription are common in *E. coli*, and in many cases the product of one transcript plays a regulatory role in the level of expression of transcripts that originate in the region (7). The possibility that ORFD could play a role in the level of *hemA* mRNA therefore warrants investigation.

A 41-bp region separates the stop codon of the *hemA* ORF from the start codon of the *prfA* gene. Interestingly, the stop codon of the *hemA* gene is recognized specifically by RF1 (14), but whether this observation is of any functional significance has not been determined. A putative promoter sequence for the *prfA* gene was identified by Craigen et al. (14) in what we have shown is the COOH-terminal region of the *hemA* coding sequence. There has been no demonstration that the sequence serves as a functional promoter, and it has not been determined directly whether the expression of the two genes is coordinated in any manner.

In summary, we have presented evidence that the *hemA* gene from *E. coli* encodes a polypeptide which functions in the synthesis of ALA but which shows no similarity to cloned ALA synthase sequences. We were unsuccessful in detecting ALA synthase activity in *E. coli* extracts. Both of these results are consistent with the data presented by Li et al. (28), which indicate that *E. coli* synthesizes ALA by the 5-carbon glutamate pathway. The postulated sequence of catalytic reactions by which glutamate is transformed into ALA consists of three steps, and it seems likely that the *hemA* gene codes for one of the enzymes involved in this process.

ACKNOWLEDGMENTS

Thanks are extended to B. Bachmann, P. Orndorff, L. Snyder, and C. P. Wolk for providing bacterial strains and plasmid and cosmid vectors. For helpful discussions, a debt of gratitude is owed to L. Snyder and M. L. Guerinot. T. Elliott is acknowledged gratefully for sharing results prior to publication.

This research was supported by the U.S. Department of Agriculture and the U.S. Department of Energy Division of Biological Energy Research.

LITERATURE CITED

1. **Adams, T. H., and B. K. Chelm.** 1988. Effects of oxygen levels on the transcription of *nif* and *gln* genes in *Bradyrhizobium japonicum*. J. Gen. Microbiol. **134**:611–618.
2. **Adams, T. H., and B. K. Chelm.** 1984. The *nifH* and *nifDK* promoter regions from *Rhizobium japonicum* share structural homologies with each other and with nitrogen-regulated promoters from other organisms. J. Mol. Appl. Genet. **2**:392–405.
3. **Adams, T. H., C. R. McClung, and B. K. Chelm.** 1984. Physical organization of the *Bradyrhizobium japonicum* nitrogenase gene region. J. Bacteriol. **159**:857–862.
4. **Bachmann, B. J.** 1983. Linkage map of *Escherichia coli* K-12, edition 7. Microbiol. Rev. **47**:180–230.
5. **Bachmann, B. J.** 1987. Derivations and genotypes of some mutant derivatives of *Escherichia coli* K-12, p. 1190–1224. *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.
6. **Bawden, M. J., I. A. Borthwick, H. M. Healy, C. P. Morris, B. K. May, and W. H. Elliott.** 1987. Sequence of human 5-aminolevulinate synthase cDNA. Nucleic Acids Res. **15**:8563.
7. **Beck, C. F., and R. A. J. Warren.** 1988. Divergent promoters, a common form of gene organization. Microbiol. Rev. **52**:318–326.
8. **Berk, A. J., and P. A. Sharp.** 1977. Sizing and mapping of early adenovirus mRNAs by gel electrophoresis of S1 endonuclease-digested hybrids. Cell **12**:721–732.
9. **Biel, S. W., M. S. Wright, and A. J. Biel.** 1988. Cloning of the *Rhodobacter capsulatus hemA* gene. J. Bacteriol. **170**:4382–4384.
10. **Borthwick, I. A., G. Srivastava, A. R. Day, B. A. Pirola, M. A. Snoswell, B. K. May, and W. H. Elliott.** 1985. Complete nucleotide sequence of hepatic 5-aminolevulinate synthase precursor. Eur. J. Biochem. **150**:481–484.
11. **Boyer, H. W., and D. Roulland-Dussoix.** 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. **41**:459–472.
12. **Burnham, B. F.** 1970. δ-Aminolevulinic acid synthase. Methods Enzymol. **17**:195–204.
13. **Casadaban, M. J., and S. N. Cohen.** 1980. Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. J. Mol. Biol. **138**:179–207.
14. **Craigen, W. J., R. G. Cook, W. P. Tate, and C. T. Caskey.** 1985. Bacterial peptide chain release factors: conserved primary structure and possible frameshift regulation of release factor 2. Proc. Natl. Acad. Sci. USA **82**:3616–3620.
15. **Davis, R. W., D. Botstein, and J. R. Roth.** 1980. Advanced bacterial genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

16. **Elhai, J., and C. P. Wolk.** 1988. A versatile class of positive-selection vectors based on the nonviability of palindrome-containing plasmids that allows cloning into long polylinkers. Gene **68**:119–138.
17. **Guerinot, M. L., and B. K. Chelm.** 1986. Bacterial δ-aminolevulinic acid synthase activity is not essential for leghemoglobin formation in the soybean/*Bradyrhizobium japonicum* symbiosis. Proc. Natl. Acad. Sci. USA **83**:1837–1841.
18. **Hadar, R., A. Slonim, and J. Kuhn.** 1976. Role of D-tryptophan oxidase in D-tryptophan utilization by *Escherichia coli*. J. Bacteriol. **125**:1096–1104.
19. **Hanahan, D.** 1985. Techniques for transformation of *E. coli*, p. 109–135. *In* D. M. Glover (ed.), DNA cloning, vol. I. IRL Press, Oxford.
20. **Hawley, D. K., and W. R. McClure.** 1983. Compilation and analysis of *Escherichia coli* promoter DNA sequences. Nucleic Acids Res. **11**:2237–2255.
21. **Herrero, A., J. Elhai, B. Hohn, and C. P. Wolk.** 1984. Infrequent cleavage of cloned *Anabaena variabilis* DNA by restriction endonuclease from *A. variabilis*. J. Bacteriol. **160**:781–784.
22. **Kannangara, C. G., S. P. Gough, P. Bruyant, J. K. Hoober, A. Kahn, and D. von Wettstein.** 1988. $tRNA^{Glu}$ as a cofactor in δ-aminolevulinate biosynthesis: steps that regulate chlorophyll synthesis. Trends Biochem. Sci. **13**:139–143.
23. **Kawakami, K., I. Toshifumi, and Y. Nakamura.** 1988. Conditionally lethal and recessive UGA-suppressor mutations in the *prfB* gene encoding peptide chain release factor 2 of *Escherichia coli*. J. Bacteriol. **179**:5378–5381.
24. **Keng, T., E. Alani, and L. Guarante.** 1986. The nine amino-terminal residues of 5-aminolevulinate synthase direct β-galactosidase into the mitochondrial matrix. Mol. Cell. Biol. **6**:355–364.
25. **Kohara, Y., K. Akiyama, and K. Isono.** 1987. The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. Cell **50**:495–508.
26. **Lee, C. C., Y. Kohara, K. Akiyama, C. L. Smith, W. J. Craigen, and C. T. Caskey.** 1988. Rapid and precise mapping of the *Escherichia coli* release factor genes by two physical approaches. J. Bacteriol. **170**:4537–4541.
27. **Leong, S. A., G. S. Ditta, and D. R. Helinski.** 1982. Heme biosynthesis in *Rhizobium*. Identification of a cloned gene coding for δ-aminolevulinic acid synthetase from *Rhizobium meliloti*. J. Biol. Chem. **257**:8724–8730.
28. **Li, J., O. Brathwaite, S. D. Cosloy, and C. S. Russell.** 1989. 5-Aminolevulinic acid synthesis in *Escherichia coli*. J. Bacteriol. **171**:2547–2552.
29. **Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall.** 1951. Protein measurement with the Folin phenol reagent. J. Biol. Chem. **193**:265–275.
30. **Maguire, D. J., A. R. Day, I. A. Borthwick, G. Srivastava, P. L. Wigley, B. K. May, and W. J. Elliott.** 1986. Nucleotide sequence of the chicken 5-aminolevulinic acid synthase gene. Nucleic Acids Res. **14**:1379–1391.
31. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
32. **McClung, C. R., J. E. Somerville, M. L. Guerinot, and B. K. Chelm.** 1987. Structure of the *Bradyrhizobium japonicum* gene *hemA* encoding 5-aminolevulinic acid synthase. Gene **54**:133–139.
33. **Messing, J.** 1983. New M13 vectors for cloning. Methods Enzymol. **101**:20–78.
34. **Messing, J., R. Crea, and P. A. Seeburg.** 1981. A system for shotgun DNA sequencing. Nucleic Acids Res. **9**:309–321.
35. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
36. **Poncz, M., D. Solowiejczyk, M. Ballantine, E. Schwartz, and S. Surrey.** 1982. "Nonrandom" DNA sequence analysis in bacteriophage M13 by the dideoxy chain-termination method. Proc. Natl. Acad. Sci. USA **79**:4298–4302.
37. **Powell, K. A., R. Cox, M. McConville, and H. P. Charles.** 1973. Mutations affecting porphyrin biosynthesis in *Escherichia coli*. Enzyme **16**:65–73.
38. **Pustell, J., and F. C. Kafatos.** 1984. A convenient and adaptable package of computer programs for DNA and protein sequence management, analysis and homology determination. Nucleic Acids Res. **12**:643–655.
39. **Ryden, M., J. Murphy, R. Martin, L. Isaksson, and J. Gallant.** 1986. Mapping and complementation studies of the gene for release factor 1. J. Bacteriol. **168**:1066–1069.
40. **Sancar, A., A. M. Hack, and W. D. Rupp.** 1979. Simple method for identification of plasmid-coded proteins. J. Bacteriol. **137**:692–693.
41. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
42. **Sasarman, A., M. Surdeanu, G. Szegli, T. Horodniceanu, V. Greceanu, and A. Dumitrescu.** 1968. Hemin-deficient mutants of *Escherichia coli* K-12. J. Bacteriol. **96**:570–572.
43. **Sasarman, A., M. Surdeanu, and T. Horodniceanu.** 1968. Locus determining the synthesis of δ-aminolevulinic acid in *Escherichia coli* K-12. J. Bacteriol. **96**:1882–1884.
44. **Schoenhaut, D. S., and P. J. Curtis.** 1986. Nucleotide sequence of mouse 5-aminolevulinic acid synthase cDNA and expression of its gene in hepatic and erythroid tissues. Gene **48**:55–63.
45. **Spears, P. A., D. Schauer, and P. E. Orndorff.** 1986. Metastable regulation of type 1 piliation in *Escherichia coli* and isolation and characterization of a phenotypically stable mutant. J. Bacteriol. **168**:179–185.
46. **Tai, T., M. D. Moore, and S. Kaplan.** 1988. Cloning and characterization of the 5-aminolevulinate synthase gene(s) from *Rhodobacter sphaeroides*. Gene **70**:139–151.
47. **Weiss, R. B., J. P. Murphy, and J. A. Gallant.** 1984. Genetic screen for cloned release factor genes. J. Bacteriol. **158**:362–364.
48. **Winans, S. C., S. J. Elledge, J. H. Krueger, and G. C. Walker.** 1985. Site-directed insertion and deletion mutagenesis with cloned fragments in *Escherichia coli*. J. Bacteriol. **161**:1219–1221.
49. **Yamamoto, M., N. S. Yew, M. Federspiel, J. B. Dodgson, N. Hayashi, and J. D. Engel.** 1985. Isolation of recombinant cDNAs encoding chicken erythroid 5-aminolevulinate synthase. Proc. Natl. Acad. Sci. USA **82**:3702–3706.
50. **Yanisch-Perron, C., J. Vieira, and J. Messing.** 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.