JOURNAL OF BACTERIOLOGY, Feb. 1993, p. 1165–1172
0021-9193/93/041165-08$02.00/0

Vol. 175, No. 4

Identification of the Regulatory Sequence of Anaerobically Expressed Locus *aeg-46.5*

MUHYEON CHOE† AND WILLIAM S. REZNIKOFF*

Department of Biochemistry, University of Wisconsin, 420 Henry Mall, Madison, Wisconsin 53706

Received 1 September 1992/Accepted 14 December 1992

A newly identified anaerobically expressed locus, *aeg-46.5*, which is located at min 46.5 on *Escherichia coli* linkage map, was cloned and analyzed. The phenotype of this gene was studied by using a *lacZ* operon fusion. *aeg-46.5* is induced anaerobically in the presence of nitrate in wild-type and *narL* cells. It is repressed by the *narL* gene product, as it showed derepressed anaerobic expression in *narL* mutant cells. We postulate that *aeg-46.5* is subject to multiple regulatory systems, activation as a result of anaerobiosis, *narL*-independent nitrate-dependent activation, and *narL*-mediated repression. The regulatory region of *aeg-46.5* was identified. A 304-bp DNA sequence which includes the regulatory elements was obtained, and the 5′ end of *aeg-46.5* mRNA was identified. It was verified that the anaerobic regulation of *aeg-46.5* expression is controlled on the transcriptional level. Computer analysis predicted possible control sites for the NarL and FNR proteins. The proposed NarL site was found in a perfect-symmetry element. The *aeg-46.5* regulatory elements are adjacent to, but divergent from, those of the *eco* gene.

Escherichia coli produces energy by breaking down a variety of carbon sources. It transforms the chemical energy of these sources into ATP. The most efficient means of generating ATP is through the use of oxygen as a terminal electron acceptor. If the oxygen supply is limited, *E. coli* seeks alternatives to oxygen or generates ATP through fermentation. Nitrate, trimethylamine-*N*-oxide, dimethyl sulfoxide, or fumarate can be used as a terminal electron acceptor instead of oxygen. The switch from aerobic to anaerobic physiology involves the turning on or off of many genes.

Studies of the genetic control of this switch have identified a locus, *fnr*, which is a regulatory gene whose product is required for the expression of enzymes for anaerobic respiration (22). The *fnr* gene product, FNR, activates the anaerobic expression of nitrate reductase (23), dimethyl sulfoxide reductase (6), and fumarate reductase (11).

In addition to the basic aerobic/anaerobic regulatory switch exemplified by the FNR system, there exist other regulatory controls which establish a hierarchy of anaerobically expressed genes. For instance, the *narL* gene was found to encode a positive regulator for nitrate reductase (*narGHJI*) in response to the presence of nitrate (23). During anaerobic growth, nitrate induces the *nar* operon about 100-fold through the action of *narL*. NarL also acts as a repressor for fumarate reductase (*frdABCD*), dimethyl sulfoxide reductase (*dmsABC*), and pyruvate formate lyase (*pfl*) synthesis when cells are cultured in nitrate-containing medium under anaerobic conditions (9, 12, 18). In this way, the expression of anaerobic respiratory enzymes is controlled both to reflect the availability of electron acceptors and to favor the use of the most efficient alternative electron acceptor (24).

An anaerobically expressed locus at min 46.5 on the *E. coli* linkage map, *aeg-46.5*, was identified by an operon fusion technique using the hybrid phage λ*plac*Mu53 (3). It was discovered that *aeg-46.5* is repressed by NarL. It showed high anaerobic expression in a *narL* mutant background. However, the phenotype of *aeg-46.5* is unusual in comparison with the previously studied NarL-regulated systems. NarL repression was observed in the absence of added nitrate, and the derepressed level of expression in a *narL* mutant was much higher than that obtainable in wild-type cells. Second, this gene is induced by nitrate. From these observations, we postulate that *aeg-46.5* is an anaerobically expressed locus which is negatively regulated by *narL* regardless of whether nitrate is present and that there is a second *narL*-independent system which activates *aeg-46.5* expression in the presence of nitrate.

The expression level of the chromosomal *lacZ* fusion to *aeg-46.5* is 1,000 to 5,000 Miller units of β-galactosidase when expression is derepressed by the mutation in the regulatory gene *narL* and cells are grown anaerobically in nitrate-containing medium, in comparison with an aerobic expression level in medium without nitrate of about 20 U. This 50- to 250-fold induction to yield high levels of expression makes the *aeg-46.5*-controlling elements an attractive promoter system for a protein expression vector because high-level induction would need only addition of nitrate and removal of oxygen.

Because of its unique regulatory features and its possible usefulness, we have initiated an analysis of the *aeg-46.5*-regulatory elements. We have cloned the *aeg-46.5* DNA from specialized transducing phage DNA. The upstream sequence was obtained, and the regulatory regions and 5′ end of the mRNA were identified. By an analysis of the amount of mRNA, the anaerobic induction control was found to be on the transcriptional level. Computer analysis predicted possible control sites for NarL and FNR. The predicted NarL site was found to be a perfectly symmetrical sequence.

MATERIALS AND METHODS

Strains and plasmids. All *E. coli* K-12 strains, phage strains, and plasmids used in this study are listed in Table 1. The construction of RZ4500W, RZ4546.5W, and their deriv-

* Corresponding author.
† Present address: Laboratory of Molecular Biology, National Cancer Institute, Bethesda, MD 20892-0037.

TABLE 1. *E. coli* strains, phages, and plasmids used

Name	Genotype	Source or reference
Strains		
MG1655	λ^- F^-	C. A. Gross
RZ4300W	RZ4500 λD69	This study
RZ4303	RZ4300 λRZ4546.5-J	This study
RZ4304	RZ4300 λRZ4546.5-K	This study
RZ4305	RZ4300 λRZ4546.5-I	This study
RZ4306	RZ4300 λRZ4546.5-FW	This study
RZ4500W	λ^- F^- *lacZΔ145*	5
RZ4500F	RZ4500 *fnr-501 zcj*::Tn*10*	This study
RZ4500L	RZ4500 *narL215*::Tn*10*	This study
RZ4546.5W	RZ4500 *aeg-46.5*::λp*lac*Mu53	5
RZ4546.5F	RZ4500 *aeg-46.5*::λp*lac*Mu53 *fnr-501 zcj*::Tn*10*	5
RZ4546.5L	RZ4500 *aeg-46.5*::λp*lac*Mu53 *narL215*::Tn*10*	5
RZ8426	MG1655 *fnr-501 zcj*::Tn*10*	C. A. Gross
RK5278	*narL215*::Tn*10*	V. Stewart
CAG12098	MG1655 *zeg-722*::Tn*10*	C. A. Gross
CAG18493	MG1655 *zbh-29*::Tn*10*	C. A. Gross
DH5α	F^- λ^- *endA1 hsdR17*(r_K^- m_K^+) *supE44 thi-1 recA1 gyrA96 relA1* φ80d*lacZ*ΔM15	
Phages		
λp*lac*Mu53	Mu *cIts62 ner*$^+$ *A*$^+$ *'uvrD'* Mu *'s 'trpAB' 'lacZ*$^+$ *lacY*$^+$ *lacA'*	T. Silhavy
λpMu507	*cIts857 Sam7* Mu*A*$^+$*B*$^+$	T. Silhavy
λRZ4546.5-I	Regulatory elements of *aeg-46.5*	This study
λRZ4546.5-J	Regulatory elements of *aeg-46.5*	This study
λRZ4546.5-K	Regulatory elements of *aeg-46.5*	This study
λRZ4546.5-FW	Regulatory elements of *aeg-46.5*	This study
λD69	λ *bam*λ1°Δ(*srl*λ-*srl*λ2) *imm*^{21} *nin5 shn6*°	C. A. Gross
Plasmids		
pMLB524	pBR322 derivative *'lacZ* Ap^r	G. M. Weinstock
pGEM7Zf(+)	Phagemid	Promega Corp.
pRZ4428	pGEM7Zf(+), *Eco*RI-*Hin*dIII fragment of pRZ4433	This study
pRZ4429	pGEM7Zf(+), *Eco*RI-*Hin*dIII fragment of pRZ4434	This study
pRZ4430	pGEM7Zf(+), *Eco*RI-*Hin*dIII fragment of pRZ4435	This study
pRZ4431	pMLB524, *Eco*RI fragment of λp*lac*Mu53	This study
pRZ4433	pMLB524, *Eco*RI fragment of λRZ4546.5-J	This study
pRZ4434	pMLB524, *Eco*RI fragment of λRZ4546.5-K	This study
pRZ4435	pMLB524, *Eco*RI fragment of λRZ4546.5-FW	This study
pRZ4446	pRZ4434, *Sph*I linker in *Nru*I site	This study
pRZ4457	Deleted pRZ4446, 2.6-kb *Eco*RI-*Hin*dIII fragment	This study
pRZ4458	Deleted pRZ4446, 2.5-kb *Eco*RI-*Hin*dIII fragment	This study
pRZ4459	Deleted pRZ4446, 2.4-kb *Eco*RI-*Hin*dIII fragment	This study
pRZ4460	Deleted pRZ4446, 2.2-kb *Eco*RI-*Hin*dIII fragment	This study
pRZ4461	Deleted pRZ4446, 2.1-kb *Eco*RI-*Hin*dIII fragment	This study
pRZ4463	Deleted pRZ4446, 1.95-kb *Eco*RI-*Hin*dIII fragment	This study
Kohara library clones		F. Blattner

atives has been described previously (5). Strain RK5278 was provided by Valley Stewart (23). Phage λp*lac*Mu53 was provided by Thomas J. Silhavy (4). Strains CAG12098, CAG12141, and CAG18493 were provided by Carol A. Gross (21).

Plasmid pMLB524 was used to clone the *Eco*RI-generated *aeg-46.5*-containing DNA fragment. This plasmid has the C-terminal *lacZ* sequence downstream of the *Eco*RI site. Ligation with *lacZ* sequences upstream of the *Eco*RI site reconstructs a functional *lacZ* gene. Plasmid pMLB524 was obtained from George M. Weinstock (2).

Introduction of *fnr-501 zcj*::Tn*10* or *narL215*::Tn*10* was done by P1 transduction (20). The *fnr-501 zcj*::Tn*10* transduced strains were tested for the *fnr* phenotype by examining their anaerobic growth on M9-glycerol-nitrate agar. The transductants that did not grow anaerobically on this medium were used.

Media and chemicals. LB and M9 minimal media were described previously (17). The indicator dye 5-bromo-4-chloro-3-indolyl-β-D-galactoside (XG; Sigma) was dissolved in *N,N*-dimethylformamide and was added to final concentration of 40 μg/ml in agar medium. Antibiotics were added to final concentrations of 30 μg/ml for kanamycin, 15 μg/ml for tetracycline, and 50 μg/ml for ampicillin. Potassium nitrate was supplied at a final concentration of 10 mM when the nitrate effect was tested. For anaerobic growth, the agar plates were incubated in an anaerobic jar (BBL Microbiology Systems; Becton Dickinson and Co.).

UV induction and phage screening. The UV induction of λp*lac*Mu53 fusion phage in strain RZ4546.5W was done according to a previously described procedure (20). The lysates were plated on XG-containing agar in a 1/100 dilution series with indicator cells. Well-separated blue turbid plaques were picked and purified. The resulting plaques were resuspended in 1 ml of phage dilution buffer, and the resuspended phages were used to test the phenotype of the phage. Three microliters of phage resuspension was spotted on an indicator *E. coli* strain on LB-XG agar. The differences of phenotype after aerobic versus anaerobic incubation and between indicator strains were observed after 24 h.

β-Galactosidase assays. β-Galactosidase assays were performed as described previously by Miller (17), using chloroform and 0.1% sodium dodecyl sulfate for permeabilization of cells. Anaerobiosis was accomplished as described previously (5) with a 95% nitrogen–5% carbon dioxide gas mixture.

Radioactive probe hybridization to miniset clones. The radioactive probes were prepared by nick translating the plasmids that have *aeg-46.5* DNA. The probes were hybridized to a nylon paper (GeneScreen Plus nylon membrane; DuPont) that has fixed DNA of Kohara miniset clones (13). Basic hybridization procedures were used as described previously (14). The hybridization result was analyzed by using a radioactivity image processing machine (BetaScope 603 blot analyzer; Betagen Corp.).

DNA manipulation. Basic recombinant DNA procedures have been described previously (14). Restriction enzymes and DNA-modifying enzymes were purchased from New England BioLabs or Promega Corp. *Taq* DNA polymerase, exonuclease III, and the *Sph*I linker were from Promega Corp. The *Sph*I linker has the sequence 5′-d(pGGCATG CC)-3′.

Plasmid pRZ4446 was constructed for the unidirectional deletion study by inserting the *Sph*I linker into the *Nru*I site of pRZ4434 as shown in Fig. 3. The plasmid DNA preparation, which was the substrate for the deletion reaction, was pretreated with T4 DNA ligase in order to repair any nicks. Exonuclease III was used to degrade the DNA after *Sph*I and *Sma*I double digestion. The degraded plasmids were made flush ended and ligated by treatment with S1 nuclease, Klenow fragment, and T4 DNA ligase. Resulting plasmids were electroporated into *E. coli* DH5α. This procedure for the construction of nested deletions is a modification of those described by Ausubel et al. (1) and Henikoff (8).

Sequence data were obtained from deleted plasmids. *Taq* DNA polymerase was used for the sequencing reaction.

Primer extension experiment. From the sequence of the upper strand, we synthesized three 40-mer primers that will prime the polymerization of the bottom-strand DNA by annealing to the mRNA template. The sequences corresponding to these primers, 160U, 270U, and 380U, are shown in Fig. 5. The primers were designed to hybridize to positions approximately 110 bases apart from each other. mRNA preparation and primer extension reactions were done according to previously described procedures (19). The reverse transcriptase was from Molecular Genetic Resources, Inc.

Computer analysis of the *aeg-46.5* sequence. Sequence data were analyzed by using the Genetics Computer Group sequence analysis software package, version 8.0 (Genetics Computer Group, Inc.). The program Fasta and GenBank data base (release 68.0) were used to search for DNA sequence similarities. Peptide sequence similarity searches were done by using the Fasta program and GenPept data base (release 64.3). The program Find was used to find sequences that match with −35 and −10 consensus sequences and with NarL, FNR, catabolite activator protein, and integration host factor control sequences.

RESULTS

Isolation of *aeg-46.5* specialized transducing phages. *aeg-46.5* is a newly identified anaerobically expressed locus which is controlled by *narL* and nitrate. To study the structure of this locus and its regulatory elements, specialized transducing phages that carry the *aeg-46.5* locus were isolated. These phages were used as sources of *aeg-46.5* DNA for subcloning experiments.

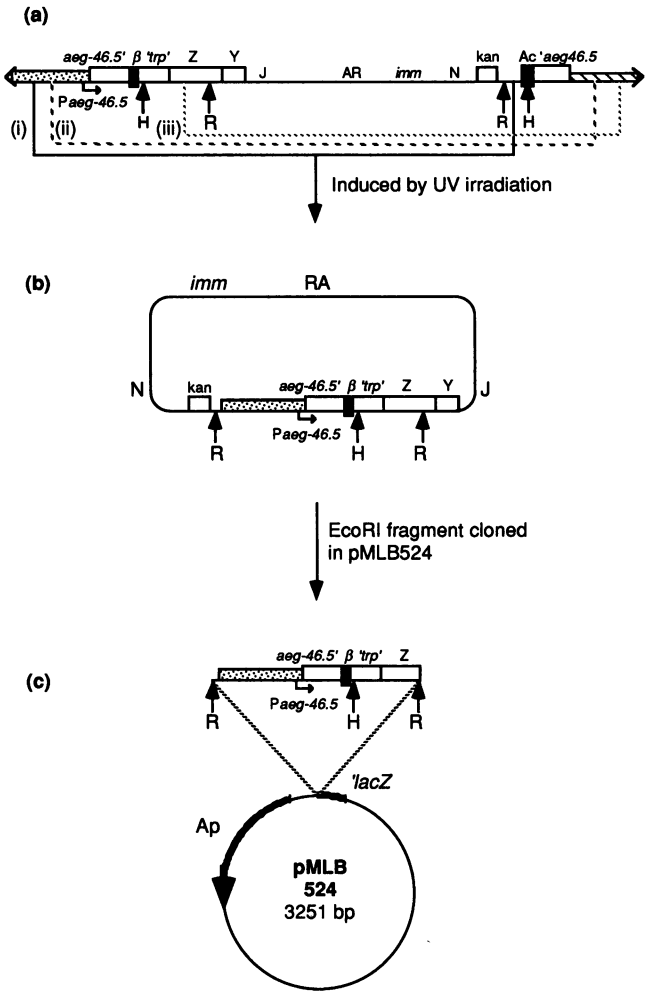

FIG. 1. Generation of specialized transducing phages and cloning of *aeg-46.5* DNA. (a) λ*plac*Mu53 fused to *aeg-46.5*. The β and Ac′ ends of phage Mu are drawn in black. The flanking chromosomal DNA on the left and right sides of the fusion are drawn as dotted and hatched arrows. The aberrant excision can include both sides of flanking chromosomal DNA (ii), or it could include chromosomal DNA from just one side (i) or the other (iii). The restriction enzyme site are indicated by arrows. R, *Eco*RI; H, *Hin*dIII. (b) The structure of specialized transducing phage for the case of excision (i). (c) Plasmid structure. The *Eco*RI fragment of a specialized transducing phage was cloned in *Eco*RI-cut pMLB524. This cloning reconstructs *lacZ*. The drawings are not to scale.

The strategy used to generate the specialized transducing phages is shown in Fig. 1. Strain RZ4546.5W has a λ*plac*Mu53 prophage inserted within the *aeg-46.5* locus. This prophage is a locked-in type; it can excise only by aberrant processes which are likely to incorporate flanking chromosomal DNA. To obtain the specialized transducing phages, the prophage was induced by UV irradiation. To screen the phages for the Lac phenotype, the induction suspensions were mixed with strain RZ4500W and plated on T agar containing the XG indicator. The phage plaques showed various phenotypes with respect to color, turbidity, and size. The variation in the phenotypes of the induced phages presumably reflects the fact that the aberrant exci-

sion removed various extents of the adjoining chromosomal regions for each excision event and in some cases lost phage sequences.

Screening for phages with the *aeg-46.5*-regulatory elements. We wanted to screen for phages that have the gene expression regulatory elements that are responsible for the high anaerobic expression of *aeg-46.5*. To identify these phages, we tested the anaerobic plaque phenotype of several phages and checked the anaerobic expression patterns of lysogens of candidate phages. The phage suspension obtained by UV irradiation was plated, and turbid blue plaques were collected. Phages from these plaques were tested for the anaerobic XG plaque phenotype on lawns composed of RZ4500W, RZ4500F, and RZ4500L, which are wild type, *fnr*, and *narL* strains, respectively. Four phages out of about 100 gave plaques with the expected pattern for *aeg-46.5* expression, which is deeper blue on an RZ4500L lawn than on an RZ4500W lawn and very light blue on an RZ4500F lawn. These phages were designated λRZ4546.5-I, -J, -K, and -FW.

These four phages were further tested for their anaerobic *lacZ* gene expression properties by creating lysogens and then assaying the level of encoded β-galactosidase. A strain which has the λ*imm*^{21} prophage λD69 was constructed. This strain, RZ4300W, was used to obtain lysogens of λRZ4546.5-I, -J, -K, and -FW through homologous recombination with the prophage. The lysogenic cells were selected on LB plates containing XG and kanamycin. Colonies that were blue and resistant to kanamycin were purified. The locations of the λRZ4546.5-I, -J, -K, and -FW prophages were checked by P1 transduction. Strains that have the λRZ4546.5 prophages linked to Tn*10* at 17.75 min were used for further studies. The selected strains were transduced to be *fnr* or *narL* with P1 lysates prepared on a strain which has *fnr-501 zcj*::Tn*10* or *narL215*::Tn*10*. The lysogenic strains and their *fnr* or *narL* derivatives were compared for anaerobic β-galactosidase expression on LB-XG-nitrate agar. All four phages showed the same pattern of anaerobic regulation as did the original fusion strain and its *fnr* and *narL* derivatives (RZ4546.5W, RZ4546.5F, and RZ4546.5L). The *narL* derivatives were dark blue, the wild types were less blue, and the *fnr* derivatives were the least blue (see Fig. 2 for a quantitative comparison).

Cloning of *aeg-46.5* DNA into plasmid vectors. It was necessary to clone the *aeg-46.5* DNA from the phage genomes into a plasmid vector in order to facilitate further analysis of this locus. The structure of the *aeg-46.5* specialized transducing phage is presented in Fig. 1b. The relevant *Eco*RI fragments of λRZ4546.5-J, -K, and -FW were purified and ligated with *Eco*RI-cut pMLB524 as shown in Fig. 1c. The recombinant plasmids were named pRZ4433, -4434, and -4435, respectively (see Fig. 3 for structures). The *Hin*dIII-*Eco*RI fragments which carry the *aeg-46.5* sequences were subcloned into *Hin*dIII- and *Eco*RI-digested pGEM7Zf(+) by screening for white ampicillin-resistant colonies. They were named pRZ4428, -4429, and -4430, respectively. The 517-bp DNA sequence at the fusion point between *aeg-46.5* and λp*lac*Mu53 was obtained (data not shown). Each plasmid revealed the same sequence corresponding to the *trp'AB'* (103 bp)-β end of Mu (105 bp)-*aeg-46.5* (309 bp) DNA. This finding confirmed that the DNA contained in each plasmid is from same original construct and that no rearrangement occurred during phage excision.

Identification of Kohara library clone with *aeg-46.5* DNA. The plasmids obtained by cloning the *Eco*RI fragment from the induced phages were used to locate the *aeg-46.5* region

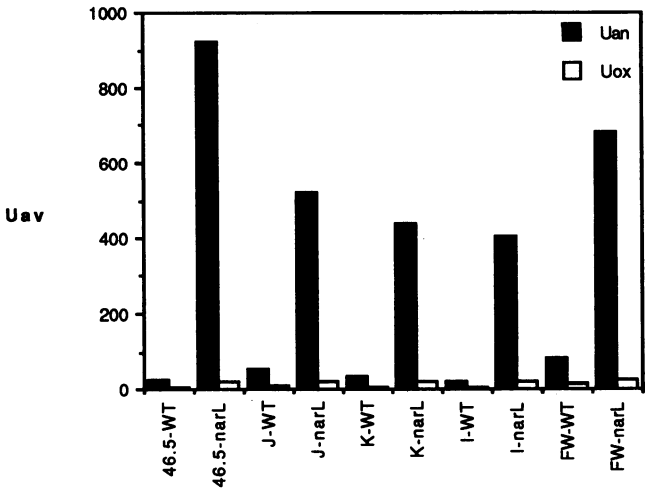

FIG. 2. β-Galactosidase expression in lysogenic strains. Anaerobic and aerobic β-galactosidase activities were measured in M9-glucose-nitrate medium. All of the lysogenic strains of four specialized transducing phages demonstrated the same pattern of anaerobic induction and derepression by the *narL* mutation in nitrate medium. WT, wild-type host (RZ4500W); narL, *narL* mutant host (RZ4500L); 46.5, original fusion strain; J, K, I, and FW, lysogens of λRZ4546.5-J, -K, -I, and -FW; Uan, anaerobic activity; Uox, aerobic activity; Uav, average of triplicate assays. The activity measurement error was less than ±30%.

in the Kohara library. We prepared radioactive probes of plasmids pRZ4433, -4434, and -4435 by nick translation. Plasmid pRZ4431, which has the *Eco*RI fragment of λp*lac*Mu53, was used as a negative control. Each *aeg-46.5* probe was hybridized to a nylon paper that had the immobilized DNA of the miniset phages. The *aeg-46.5* probe annealed uniquely to DNA from the λ19D1 clone of the Kohara library (clone 372 of its miniset), which is located at 46.5 ± 1 min on the *E. coli* linkage map (data not shown).

Location of the regulatory region by unidirectional deletion. To locate the *aeg-46.5*-regulatory elements, we deleted DNA from the upstream boundary of the cloned fragment in the plasmid and tried to obtain the smallest fragment (largest deletion) that retains the *aeg-46.5*-regulatory elements. The lengths of the remaining chromosomal DNA fragments in 80 colony isolates were measured by double digesting the resulting plasmids with *Eco*RI and *Hin*dIII and analyzing the resulting products by agarose gel electrophoresis. The deletion end points in these plasmids span the 2.6-kb (same as in pRZ4434 [Fig. 3]) chromosomal sequence of pRZ4446 (data not shown). We selected 10 clones whose deletion end points are 300 to 500 bp apart from each other and transformed them into RZ4500W and RZ4500L.

The β-galactosidase assay results of these strains after anaerobic growth in M9-glucose-nitrate medium showed that only those plasmids which retained at least 2.4 kb of chromosomal DNA had high-level anaerobic β-galactosidase synthesis (data not shown). The next-largest plasmid in the series, pRZ4461, which retained 1.9 kb of chromosomal DNA, showed a low level of β-galactosidase synthesis and therefore had lost the regulatory element that is responsible for high anaerobic expression of *aeg-46.5* in the presence of the *narL* mutation. Four plasmids, pRZ4457, -4458, -4459, and -4460, which have 2.4, 2.3, 2.2, and 2.0 kb of chromosomal DNA, and pRZ4461 were tested to narrowly define the DNA that is responsible for anaerobic regulation. The re-

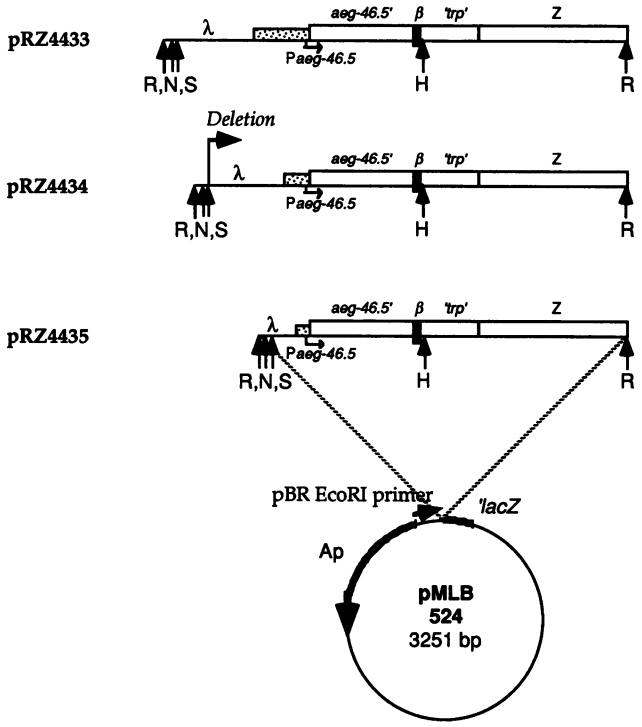

FIG. 3. Structures of *Eco*RI fragments cloned in plasmid pMLB524. Plasmids pRZ4433, -34, and -35 were constructed by cloning the relevant *Eco*RI fragments. The size and restriction map of each cloned fragment was obtained by agarose gel electrophoretic analysis of relevant restriction enzyme digests (data not shown). pRZ4433 has a 5.0-kb *Hin*dIII-*Eco*RI fragment which contains *aeg-46.5* DNA. pRZ4434 and -35 have 4.6- and 3.2-kb corresponding fragments, respectively. These *Hin*dIII-*Eco*RI fragments are drawn to scale to show the estimated sizes of λ and chromosomal DNA incorporated in each plasmid. pRZ4433, -34, and -35 had about 3.1, 2.6, and 2.4 kb of chromosomal DNA, respectively. The pBR322 *Eco*RI clockwise primer, which was used for DNA sequencing, is drawn as an arrow on pMLB524. Restriction enzyme sites are drawn with vertical arrows. R, *Eco*RI; N, *Nru*I; S, *Sma*I; H, *Hin*dIII. The *Nru*I site was used to insert the *Sph*I linker. Double digestion of the resulting plasmid with *Sph*I and *Sma*I leaves a 3′ protruding end toward the vector sequence and a blunt end toward *aeg-46.5* DNA. The orientation of deletion digestion by exonuclease III is drawn on pRZ4434 as an angled arrow marked *Deletion*.

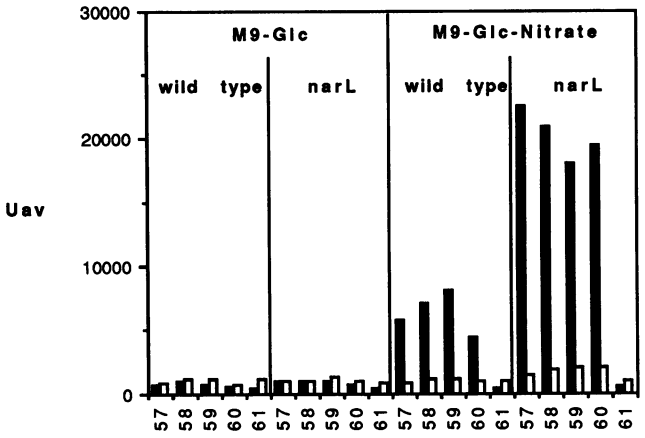

FIG. 4. β-Galactosidase assays of deletion plasmids. Levels of expression of β-galactosidase by pRZ4457, -58, -59, -60, and -61 (as indicated at the bottom) were measured in wild-type and *narL* mutant host cells in M9-glucose and M9-glucose-nitrate media. pRZ4461 lost the regulatory element that is responsible for *narL* derepression in nitrate medium. Nitrate induction of *aeg-46.5* in the wild-type strain is shown amplified on these multicopy plasmids. ■, anaerobic activity; □, aerobic activity; Uav, average of triplicate assays. The error of activity measurements was less than ±30%.

sults (Fig. 4) indicate that pRZ4460 is the plasmid which has the smallest extent of *aeg-46.5* DNA with high-level-regulated *lacZ* expression and thus all of the *cis*-acting regulatory elements. The difference in size between pRZ4460 and pRZ4461 is just 105 bp. The deletion of this 105 bp of DNA in pRZ4461 must have removed part of the regulatory elements.

The assays described above were done both in the presence and in the absence of nitrate. The results for wild-type cells show that the dramatic increase of expression caused by the nitrate induction is also found when *aeg-46.5* is encoded by a multicopy plasmid as opposed to single-copy chromosomal DNA (5). The assay results obtained from the cultures grown in medium without nitrate showed almost no anaerobic induction regardless of the *narL* genetic background of the strains.

Sequence of regulatory elements and upstream DNA for the *aeg-46.5* locus. Sequencing of the *aeg-46.5* DNA in pRZ4460, -4461, and -4463 was done by using a pBR322 *Eco*RI clockwise primer (Fig. 3), which reads the upper strand of *aeg-46.5*. The sequence data were confirmed by reading the bottom strand utilizing primers whose hybridization sites were spaced to cover most of the *aeg-46.5* sequence as shown in Fig. 5. These primers were designed from the upper-strand sequence information. They were also used in the primer extension experiments described below. We used *Taq* DNA polymerase and, as a template, denatured double-stranded plasmid DNA. As shown in Fig. 5, 304 bp of sequence data were obtained from these three plasmids. The open reading frames are indicated under the DNA sequence.

Identification of mRNA 5′ ends. Three 40-mer primers were designed to use in a primer extension experiment in order to define the *aeg-46.5* mRNA 5′ ends. The sequences for these primers start 85 (160U), 195 (270U), and 306 (380U) bases downstream of the pRZ4461 deletion end point. Anaerobic cultures of *narL* cells containing pRZ4459, -4460, and -4461 were grown in M9-glucose-nitrate; an aerobic culture of cells containing pRZ4460 was also grown. About 6 μg of total RNA extracted from anaerobic cells and 47 μg extracted from aerobic cells were used for the primer extension reactions.

The results for primer 160U are displayed in Fig. 6. Plasmids pRZ4459 and -4460 program the same mRNA species anaerobically. This mRNA has one major 5′ end (at position 128 in Fig. 5) and a few minor 5′ ends. Plasmid pRZ4461, which did not show anaerobic β-galactosidase induction, did not produce these mRNAs. The aerobic culture of the pRZ4460-harboring cell gave major mRNA products which are three or four bases longer than the major anaerobic product (with 5′ ends at positions 124 and 125). The primer extension products from primers 270U and 380U confirm these results (data not shown). Thus, there are no other major *aeg-46.5*-related mRNAs.

The amounts of cDNA corresponding to the major species of anaerobic and aerobic mRNAs were quantitated by using a radioactivity image processing machine. The counts for the major anaerobic mRNA species of pRZ4459 and -4460 in column a were 103 and 102 cpm. The major aerobic mRNA

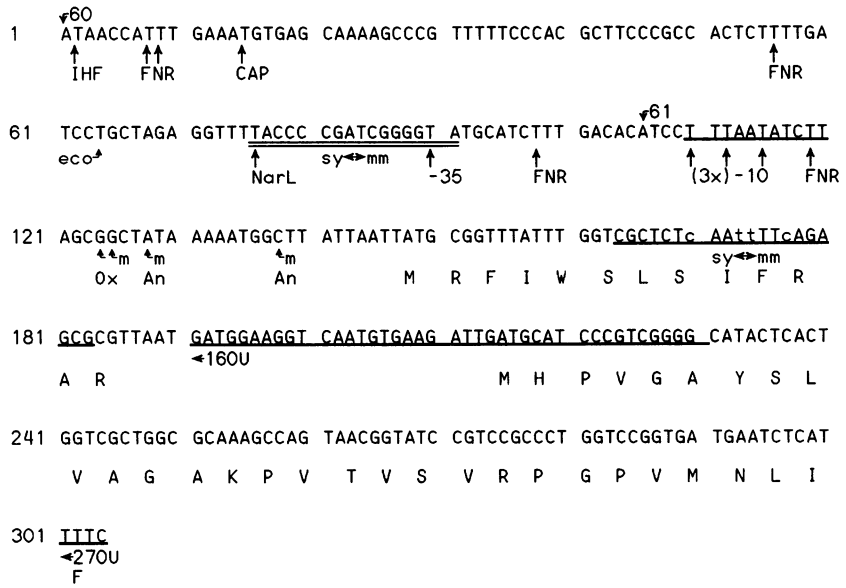

FIG. 5. Upstream DNA sequence of *aeg-46.5*. A 304-bp upstream DNA sequence was determined. Locations of the deletion end points for pRZ4460 (leaves regulatory elements intact) and pRZ4461 (prevents high-level regulated expression of *aeg-46.5*) are indicated by 60 and 61. The amino acid sequences of possible open reading frames are drawn under the DNA sequence. The primers used in 5′-end determinations are complementary to the sequences underlined, and their 3′ ends are indicated with ← and primer name. The 5′ ends of mRNAs are indicated with m and labeled An (anaerobic) or Ox (aerobic). The upstream boundaries of putative NarL, FNR, catabolite activator protein (CAP), and integration host factor (IHF) sites, and −35 and −10 regions, are indicated with ↑. Symmetry elements are underlined and labeled sy↔mm. The start point of similarity with the *E. coli* ecotin gene is indicated (eco). The downstream sequence data are not presented because of ambiguities. Primer 380U is 111 bp downstream from primer 270U.

of pRZ4460 in column a gave 63 cpm. From these data and the amount of total RNA extract used for the primer extension reactions, we calculated the ratio of anaerobic to aerobic mRNA to be about 13:1. This ratio corresponds very well with the anaerobic induction ratio of plasmid-encoded *aeg-46.5* in a *narL* mutant as shown in Fig. 4.

Computer analysis of the *aeg-46.5* sequence. DNA sequence similarity and amino acid sequence similarity searches were executed by using the data base search program Fasta. One sequence was found to have 95.4% identity with the 65 bp at the extreme upstream end of the *aeg-46.5* sequence. This sequence is from upstream of the promoter for the *E. coli eco* gene (7, 15). The orientations of these two sequences would be divergent from the overlapping region. The start point of similarity is indicated in Fig. 5 as eco. No other sequence with significant similarity to *aeg-46.5* sequence was found. The amino acid sequences of the peptides which are from the open reading frames starting at position 148 (14 amino acids), 215 (31 amino acids), or 338 (64 amino acids) were used to search for similar peptides. No significant matches were found.

aeg-46.5 is regulated by NarL and possibly by FNR; thus, the presence of sequences that may be responsible for NarL or FNR control were sought in the upstream sequences by using the search program Find, allowing some mismatches. The hexamer sequence pattern TACTCC was found to be responsible for NarL control of the *narGHJI* operon through a saturation mutagenesis experiment (6a). The sequence TACCCC was found at position 76 and was part of a perfectly symmetrical sequence, TACCCCGA · TCGGG GTA. Two bases of this symmetrical sequence overlap with the putative −35 region at position 90. No other sequence with an equal or better match was found. For the FNR binding site search, the pattern sequence TTTGAT was used, revealing five possible FNR control sites (Fig. 5). Figure 5 also displays the locations of possible recognition sites for catabolite activator protein-cyclic AMP and integration host factor. It is not known whether these factors play a role in *aeg-46.5* expression. The symmetrical sequence CGC TCTcAAt · tTTcAGAGCG was found at position 164. This sequence is downstream of the 5′ end of the mRNA.

The consensus sequence of −35 and −10 regions was used to scan for possible σ^{70} promoter sequences. A perfect −35 sequence was found at position 99, but this −35 sequence is too close to the transcription initiation site. There were three −10 consensus sequences starting at position 110 with one mismatch, at position 112 with two mismatches, and at position 115 with two mismatches. The −35 region that could be compatible with these −10 sequences is at position 90, and this −35 region has 14-, 16-, and 19-base spacing with each −10 sequence.

DISCUSSION

The anaerobically expressed locus *aeg-46.5*, which was identified by an operon fusion technique using random phage insertion, was partially cloned and analyzed on the DNA level. The *aeg-46.5* DNA was first obtained as part of the genomes of illegitimately excised phages. The phages were screened for whether they carried the *aeg-46.5*-regulatory element by using indicator strains that showed different phenotypes of *aeg-46.5* because of mutation of the regulatory protein in each strain. It was demonstrated that the phage plaques could be screened for the anaerobic regulatory phenotype of their fusion constructs.

The DNA from the *aeg-46.5* locus has been cloned from the specialized transducing phage into a plasmid vector. These plasmids were used to identify which clone of the

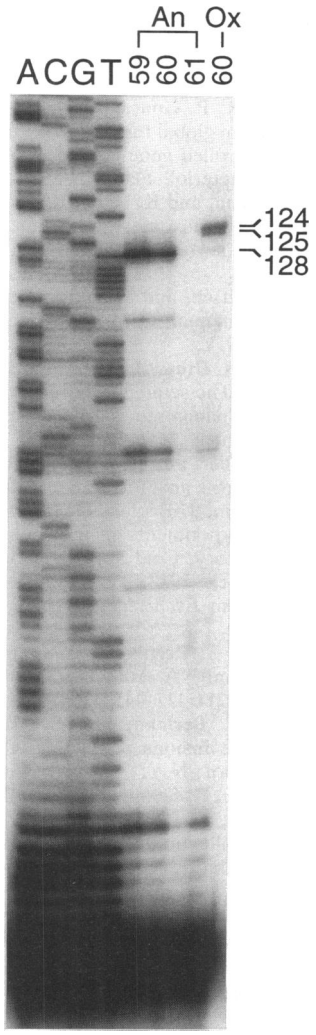

FIG. 6. Identification of the 5′ end of mRNA. The result of a primer extension reaction from primer 160U was analyzed by comparison with an adjoining sequencing ladder. The sequencing reactions were done by using the same primer as that used in the primer extension reaction. The major 5′ end of anaerobic mRNA is at position 128, and the 5′ ends of aerobic mRNA are at positions 124 and 125.

Kohara library has the *aeg-46.5* DNA, and *aeg-46.5* was located on the *E. coli* physical map. Analysis of the whole gene that can be obtained from the identified clone should help to elucidate the identity of this gene.

The pattern of *aeg-46.5* expression when encoded by a multicopy plasmid is similar in many respects to that found for the chromosome fusion. The maximal level of expression (in a *narL* background when grown anaerobically in the presence of nitrate) encoded by a multicopy plasmid is 20-fold higher than that encoded by the chromosome fusion. This enhanced expression approximates the plasmid's copy number. This finding indicates that no factor required for the derepressed expression of *aeg-46.5* is lacking in the multicopy situation. The plasmid-encoded *aeg-46.5* also responds to the presence or absence of oxygen, nitrate, and *narL*, although the patterns of the nitrate and *narL* responses differ somewhat from the single-copy situation. The most obvious difference is that the derepression effect of the *narL* mutation was not manifested in the absence of exogenous nitrate. The second difference is that the nitrate induction in the wild-type cells is more obvious. Third, the anaerobic/aerobic induction ratios are less dramatic. We do not have explanations for these results.

The observation that *aeg-46.5* responds to the presence of nitrate in the absence of a functional *narL* product and also responds to the presence of the *narL* product leads us to postulate that *aeg-46.5* is controlled by two independent nitrate-related regulatory systems. It is known that the genes for some fermentative enzymes are repressed by nitrate independently of *narL*, although the components of this regulatory system are unknown. Perhaps *aeg-46.5* is under the control of this system.

The deletion study identified the upper boundary of the regulatory region required for anaerobic control. The identification of the major 5′ end of the mRNA made it possible to locate the regulatory region to about 150 bp of obtained sequence. Within this region are several possible regulatory sites. One possible *narL* control sequence with a 1-bp mismatch was discovered. This sequence is within a perfect-symmetry element. There is no other case reported that implies that NarL might act on a symmetrical sequence. Further studies on this possible NarL site may lead to new insight into the nature of NarL regulation. The symmetry of this site could lead to an explanation for the differences in NarL regulatory pattern between previously studied NarL-regulated genes and *aeg-46.5*. The symmetric sequence overlaps with two bases of the −35 region. A simple hypothetical explanation for NarL repression of *aeg-46.5* could be that NarL binding on this symmetric sequence blocks the access of RNA polymerase to the −35 region of the *aeg-46.5* promoter.

aeg-46.5 is known to respond to the presence of FNR for anaerobic expression. Thus, we have looked for the presence of a possible FNR binding site. The FNR binding site search gave several possible FNR control sites. All of these sites have the highly conserved second T base, but the conserved four-base spacing between symmetric parts is found only in the one sequence that starts at position 9. The FNR control site is expected to be about 40 to 60 bases away from the transcription initiation site. The putative FNR control site at position 56 which has one half-perfect consensus sequence and two bases matching the other half site is in this expected location. It has its center 64.5 bases away from the transcription initiation site.

The upstream sequence associated with the *E. coli eco* gene (7, 15) overlaps by 65 bp with the sequence reported here. However, we found three differences between our sequence data and those previously reported. Though there are several anaerobically regulated genes which are related to protease activity (10, 16), it is not known whether the ecotin gene is regulated by anaerobiosis.

We found three possible −10 regions. It is possible that anaerobic and aerobic mRNAs are programmed from different −10 regions which give different 5′ ends to the mRNAs. Maybe the anaerobic and aerobic switches involve a choice of −10 region. The mRNA measurements were found to be compatible with the anaerobic induction ratio, which indicates that the anaerobic induction control is mainly on the transcriptional level. How the transcriptional controls occur remains to be determined.

ACKNOWLEDGMENTS

The nylon membrane which contains the Kohara library clones and the hybridization process were kindly provided by Shuang-En Chuang and Fred R. Blattner.

This work was supported by Public Health Service grant GM19670.

REFERENCES

1. **Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. A. Smith, J. G. Seidman, and K. Struhl (ed.).** 1987. Current protocols in molecular biology. John Wiley & Sons, Inc., New York.
2. **Berman, M. L., D. E. Jackson, A. Fowler, I. Zabin, L. Christensen, N. P. Fiil, and M. N. Hall.** 1984. Gene fusion techniques: cloning vectors for manipulating lacZ gene fusions. Gene Anal. Tech. **1:**43.
3. **Bremer, E., T. J. Silhavy, and G. M. Weinstock.** 1985. Transposable lambda *plac*Mu bacteriophages for creating *lacZ* operon fusions and kanamycin resistance insertions in *Escherichia coli*. J. Bacteriol. **162:**1092–1099.
4. **Bremer, E., T. J. Silhavy, and G. M. Weinstock.** 1988. Transposition of lambda *plac*Mu is mediated by the A protein altered at its carboxy-terminal end. Gene **71:**177–186.
5. **Choe, M. H., and W. S. Reznikoff.** 1991. Anaerobically expressed *Escherichia coli* genes identified by operon fusion techniques. J. Bacteriol. **173:**6139–6146.
6. **Cotter, P. A., and R. P. Gunsalus.** 1989. Oxygen, nitrate, and molybdenum regulation of *dmsABC* gene expression in *Escherichia coli*. J. Bacteriol. **171:**3817–3823.

6a. **DeMoss, J. A.** Personal communication.

7. **Erpel, T., P. Wang, C. S. Craik, R. J. Fletterick, and M. E. McGrath.** 1992. Physical map location of the new *Escherichia coli* gene *eco*, encoding the serine protease inhibitor ecotin. J. Bacteriol. **174:**1704.
8. **Henikoff, S.** 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene **28:**351–359.
9. **Iuchi, S., and E. C. Lin.** 1987. The *narL* gene product activates the nitrate reductase operon and represses the fumarate reductase and trimethylamine N-oxide reductase operons in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **84:**3901–3905.
10. **Jamieson, D. J., and C. F. Higgins.** 1986. Two genetically distinct pathways for transcriptional regulation of anaerobic gene expression in *Salmonella typhimurium*. J. Bacteriol. **168:** 389–397.
11. **Jones, H. M., and R. P. Gunsalus.** 1987. Regulation of *Escherichia coli* fumarate reductase (*frdABCD*) operon expression by respiratory electron acceptors and the *fnr* gene product. J. Bacteriol. **169:**3340–3349.
12. **Kalman, L. V., and R. P. Gunsalus.** 1989. Identification of a second gene involved in global regulation of fumarate reductase and other nitrate-controlled genes for anaerobic respiration in *Escherichia coli*. J. Bacteriol. **171:**3810–3816.
13. **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.
14. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
15. **McGrath, M. E., W. M. Hines, J. A. Sakanari, R. J. Fletterick, and C. S. Craik.** 1991. The sequence and reactive site of ecotin: a general inhibitor of pancreatic serine proteases from *Escherichia coli*. J. Biol. Chem. **266:**6620–6625.
16. **Miller, C. G., J. L. Miller, and D. A. Bagga.** 1991. Cloning and nucleotide sequence of the anaerobically regulated *pepT* gene of *Salmonella typhimurium*. J. Bacteriol. **173:**3554–3558.
17. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
18. **Sawers, G., and A. Bock.** 1988. Anaerobic regulation of pyruvate formate-lyase from *Escherichia coli* K-12. J. Bacteriol. **170:**5330–5336.
19. **Schulz, V. P., and W. S. Reznikoff.** 1990. In vitro secondary structure analysis of mRNA from lacZ translation initiation mutants. J. Mol. Biol. **211:**427–445.
20. **Silhavy, T. J., M. L. Berman, and L. W. Enquist.** 1984. Experiments with gene fusions. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
21. **Singer, M., T. A. Baker, G. Schnitzler, S. M. Deischel, M. Goel, W. Dove, K. J. Jaacks, A. D. Grossman, J. W. Erickson, and C. A. Gross.** 1989. A collection of strains containing genetically linked alternating antibiotic resistance elements for genetic mapping of *Escherichia coli*. Microbiol Rev. **53:**1–24.
22. **Spiro, S., and J. R. Guest.** 1990. FNR and its role in oxygen-regulated gene expression in *Escherichia coli*. FEMS Microbiol. Rev. **6:**399–428.
23. **Stewart, V.** 1982. Requirement of Fnr and NarL functions for nitrate reductase expression in *Escherichia coli* K-12. J. Bacteriol. **151:**1320–1325.
24. **Stewart, V.** 1988. Nitrate respiration in relation to facultative metabolism in enterobacteria. Microbiol. Rev. **52:**190–232.