

## Repression of the *Escherichia coli modABCD* (Molybdate Transport) Operon by ModE†

AMY M. GRUNDEN, RAMESH M. RAY, JOSEPH K. ROSENTEL,  
FRANK G. HEALY,‡ AND K. T. SHANMUGAM\*

Department of Microbiology and Cell Science, University of Florida, Gainesville, Florida 32611

Received 23 August 1995/Accepted 27 November 1995

The *modABC* gene products constitute the molybdate-specific transport system in *Escherichia coli*. Another operon coding for two proteins which diverges from the *modABCD* operon has been identified. The first gene of this operon codes for a 262-amino-acid protein, designated ModE (28 kDa), and the second gene codes for a 490-amino-acid protein, ModF (54 kDa). The role of ModF has not yet been determined; however, mutations in *modE* derepressed *modABCD* transcription even in the presence of molybdate, suggesting that ModE is a repressor. ModE, in the presence of 1 mM molybdate, repressed the production of plasmid-encoded ModA and ModB' proteins in an *in vitro* transcription-translation system. DNA mobility shift experiments confirmed that ModE binds to an oligonucleotide derived from the operator region of the *modABCD* operon. Further experimentation indicated that ModE binding to target DNA minimally requires an 8-bp inverted-repeat sequence, TAAC · GTTA. A highly conserved amino acid sequence, TSARNQXXG (amino acids 125 to 133), was identified in ModE and homologs from *Azotobacter vinelandii*, *Haemophilus influenzae*, *Rhodobacter capsulatus*, and *Clostridium pasteurianum*. Mutants with mutations in either T or G of this amino acid sequence were isolated as "superrepressor" mutants. These mutant proteins repressed *modABCD* transcription even in the absence of molybdate, which implies that this stretch of amino acids is essential for the binding of molybdate by the ModE protein. These results show that molybdate transport in *E. coli* is regulated by ModE, which acts as a repressor when bound to molybdate.

The presence of molybdenum is required for the activity of several enzymes found in animals, plants, and bacteria, such as sulfite oxidase, xanthine dehydrogenase, nitrate reductase, formate dehydrogenase, and nitrogenase (46). In these organisms, the molybdenum is present in the molybdoenzymes in the form of a pterin-containing molybdenum cofactor (MoCo), with the exception of that found in nitrogenase, which has an iron-molybdenum cofactor (FeMoCo) (1, 16, 34).

In *Escherichia coli*, the successful production of molybdoenzymes relies upon the efficient uptake of molybdate via the molybdate-specific transporter encoded by the *modABCD* operon (formerly known as the *chID* locus) (18, 26, 35, 42). The *E. coli* molybdate transport machinery, in which *modA* encodes a molybdate-specific periplasmic binding protein, *modB* encodes an integral membrane channel-forming protein, and *modC* encodes an ATP-binding energizer, closely resembles the established ATP-binding cassette (ABC) transporter motif (15, 26). Homologous molybdate transport systems have also been described for *Azotobacter vinelandii*, *Rhodobacter capsulatus*, and *Haemophilus influenzae* (9, 23, 45).

Initial studies of the characterization of *mod* mutants revealed that strains harboring these mutations were incapable of producing nitrate reductase or formate dehydrogenase activity without molybdate supplementation (11, 14, 27, 36, 39, 41, 43). There are data to suggest that when *mod* mutants are

grown in molybdate-supplemented media, molybdate enters the cells by means of the sulfate transport system as well as other nonspecific anion transporters (22, 36). Some of the *mod* mutants were also analyzed for their molybdate uptake kinetics and were shown to transport molybdate at a much lower rate than *mod*<sup>+</sup> strains did (5, 14). Further investigations of *mod* mutants have demonstrated that the *modABCD* operon is regulated by the intracellular concentration of molybdate. Specifically, high levels of molybdate reduced the level of transcription of the transport genes (27, 35, 36). This reduction in the level of transcription of the *modABCD* operon is most likely mediated by a molybdate-activated repressor protein.

An analysis of the *modABCD* DNA sequence revealed a region in between the transcription and translation start sites of the *modA* gene which contains an 8-base inverted repeat (TAAC · GTTA) flanked by two CAT (CA) boxes (26, 35, 36). There are indications that either the CAT boxes or the inverted repeat or both are involved in the binding of the *mod* operon repressor (35). The inverted repeat found in this region is analogous to the inverted-repeat sequences implicated as the target binding sites for both the Met and Trp repressor (MetJ and TrpR) proteins (21, 32, 33).

In our laboratory, a derivative of strain SE2069 [ $\phi(modA-lacZ)$ ] (36) which derepressed *modA* transcription in the presence of molybdate was isolated. The complementation of this mutation resulted in the identification of two genes that are located upstream of the *modA* gene and are transcribed in the opposite orientation relative to *modABCD*. These genes have been designated *modE* and *modF*, as they appear to constitute a two-gene operon. This communication provides the genetic and physiological evidence for the *modEF* operon as well as biochemical evidence which verifies the presumptive function of *modE* as the repressor protein that regulates the *modABCD* operon.

\* Corresponding author. Mailing address: Department of Microbiology and Cell Science, Box 110700, University of Florida, Gainesville, FL 32611-0700. Phone: (904) 392-2490. Fax: (904) 392-5922. Electronic mail address: Shanmugam@ICBR.IFAS.UFL.EDU.

† Florida Agricultural Experiment Station Journal Series no. R-04745.

‡ Present address: Department of Plant Pathology, Cornell University, Ithaca, N.Y.

TABLE 1. Bacterial strains, phages, and plasmids used in this study

Strain, phage, or plasmid	Genotype	Source and/or reference
<b>Strains</b>		
BW545	$\Delta(lacU)169 rpsL$	Laboratory collection
MC4100	$araD139 \Delta(argF-lacU)205 rpsL150 relA1 flbB5301 deoC1 ptsF25$	CGSC 6152
RK4353	MC4100 <i>gyrA219 non-9</i>	V. Stewart
VJS720	<i>chlD247::Tn10 (modB247)</i>	V. Stewart
VJS1779	RK4353 <i>moe-251::Tn10d(Tc)</i>	V. Stewart
VJS1780	RK4353 <i>mob-252::Tn10d(Tc)</i>	V. Stewart
VJS1782	RK4353 <i>moa-254::Tn10d(Tc)</i>	V. Stewart
VJS1784	RK4353 <i>mog-256::Tn10d(Tc)</i>	V. Stewart
BL21 $\lambda$ DE3	<i>hsdS gal lacUV5-gene1 (T7)</i>	F. Studier (44)
SE1325	<i>cysC43 srl-300::Tn10 thr-1 leu-6 thi-1 lacY1 galK2 ara-14 xyl-5 mtl-1 proA2 his-4 argE3 rpsL31 tsx-33 supE44 modB138::Tn5</i>	Laboratory collection
SE1811	SE2069 <i>modE1</i>	This study
SE2069	BW545 $\phi(modA'-lacZ^+)102$	36
SE1910	BW545 ( $\Delta modE$ -Km) $_2$	This study
SE1938	BW545 $\lambda$ RM26	This study
SE1934	Rk4353 ( <i>modF</i> -Km) $_1$	This study
SE1940	BW545 ( <i>modF</i> -Km) $_1$ $\lambda$ RM26	P1 transduction (SE1938 $\times$ SE1934)
SE1942	BW545 ( $\Delta modE$ -Km) $_2$ $\lambda$ RM26	P1 transduction (SE1938 $\times$ SE1910)
SE1952	BW545 <i>modB138::Tn5</i> $\lambda$ RM26	P1 transduction (SE1938 $\times$ SE1325)
<b>Phages</b>		
$\lambda$ RZ5	$\lambda$ 'bla' <i>lacZ lacY</i> $^+$	Laboratory collection
$\lambda$ RM26	$\lambda$ $\phi(modE'-lacZ^+)lacY^+ bla^+$	This study
<b>Plasmids</b>		
pTrc99A	<i>P<sub>trc</sub> lacI<sup>f</sup> Amp<sup>r</sup></i>	Pharmacia
pT7-7	<i>pT7 lacZ' Amp<sup>r</sup></i>	Pharmacia
pUC4K	<i>lacZ' Kan<sup>r</sup> Amp<sup>r</sup></i>	Pharmacia
pZ1918	<i>'lacZ Amp<sup>r</sup></i>	H. Schweizer (40)
pFGH15	<i>modABCDEF<sup>+</sup> Amp<sup>r</sup></i>	26
pSE1004	<i>modABCDE<sup>+</sup> Amp<sup>r</sup></i>	22
pSE1009	<i>modABC<sup>+</sup> Amp<sup>r</sup></i>	22
pRM1	<i>modA<sup>+</sup>B' Amp<sup>r</sup></i>	This study
pAG1	<i>modE<sup>+</sup> Amp<sup>r</sup></i>	This study
pRM12	<i>modEF<sup>+</sup> Amp<sup>r</sup></i>	This study
pRM13	<i>modEF<sup>+</sup> Amp<sup>r</sup> Tet<sup>r</sup></i>	This study
pRM14	$\phi(modE'-lacZ^+) modF+ Ampr$	This study
pRM16	<i>modE<sup>+</sup> modF-Km Amp<sup>r</sup> Tet<sup>r</sup></i>	This study
pRM17	<i>'modE <math>\phi(modF-Km) 'lacZ rep(Ts) Cam<sup>r</sup></math></i>	This study
pRM25	<i>modEF<sup>+</sup> Amp<sup>r</sup></i>	This study
pRM26	$\phi(modE'-lacZ) modF+ Ampr$	This study
pRM22	<i>modE<sup>+</sup> Amp<sup>r</sup></i>	This study
pRM23	<i>modF<sup>+</sup> Amp<sup>r</sup></i>	This study
pRM9	<i>modE-Km modABCD<sup>+</sup> Amp<sup>r</sup></i>	This study
pMAK705	<i>'lacZ rep(Ts) Cam<sup>r</sup></i>	12
pRM10	$\phi(modE'-Km) 'lacZ rep(Ts) Camr$	This study

## MATERIALS AND METHODS

**Bacterial strains, bacteriophages, and plasmids.** The origins and genotypes of the bacterial strains, phages, and plasmids used in this study are presented in Table 1. All of the strains are *E. coli* K-12 derivatives. The techniques used for the plasmid constructions have been described previously (25). Strain SE1811 was isolated as a spontaneous Lac<sup>+</sup> derivative of strain SE2069 on lactose-MacConkey-molybdate (1 mM) medium. By using phage P1, the mutation responsible for the derepressed ModA phenotype was shown to cotransduce with  $\phi(modA-lacZ)$  at 94%.

**Media and growth conditions.** Bacterial cultures were grown in Luria broth (LB) supplemented as needed with 0.3% glucose (LBG) and 1 mM sodium molybdate. Growth conditions for cultures used in  $\beta$ -galactosidase assays were described previously (36). Unless otherwise indicated, all cultures were incubated at 37°C. When required, ampicillin (100  $\mu$ g/ml), kanamycin (50  $\mu$ g/ml), chloramphenicol (15  $\mu$ g/ml), and tetracycline (15  $\mu$ g/ml) were added to the growth media.

**Sequencing of *modEF* DNA.** The *modEF* DNA was sequenced by the Sanger

dideoxy method (13, 38). The plasmids from which the *modEF* sequence was determined are presented in Fig. 1. Primers synthesized by the DNA Synthesis Core Facility of the Interdisciplinary Center for Biotechnology Research at the University of Florida based on the DNA sequence were used in the DNA sequence experiments. The DNA sequence was analyzed by using computer software programs GCG (2, 4, 7) and Genepro (Riverside Scientific, Seattle, Wash.). The DNA sequence is available from GenBank under accession no. U27192.

**Construction of *modE-lacZ* fusion plasmids.** For ease of later genetic manipulations, a plasmid pUC19-based  $\phi(modE-lacZ)$  construct was prepared by first removing the 3.2-kb *KpnI-HindIII* fragment from plasmid pFGH15 and placing it into the *KpnI-HindIII* site of pUC19 to yield plasmid pRM12. The 3.2-kb *SmaI* fragment containing the *lacZ* gene from plasmid pZ1918 (40) was ligated into the *AflIII* site of plasmid pRM12, which had been previously modified with the Klenow fragment of DNA polymerase I (Klenow), yielding plasmid pRM14. Since a lower-copy-number plasmid was desired for the purpose of monitoring *in vivo* levels of transcription of *modE-lacZ*, a 3.2-kb *EcoRI-HindIII* fragment from

TABLE 2. Purification of ModE protein

Fraction	Total protein (mg)	Yield (%)	Purification (fold)
Extract	238	100	1.0
Q Sepharose fast flow	57	23.9	4.2
HiTrap Heparin	16.5	6.9	14.4
Chromatofocusing	12	5.0	19.8

plasmid pRM12 was ligated into the *EcoRI-HindIII* sites of plasmid vector pBR322, resulting in plasmid pRM13. The *mod* insert in plasmid pRM13 was unstable. The removal of the *tet* gene as a *HindIII-AvaI* fragment from plasmid pRM13, yielding plasmid pRM25, resolved this problem. Next, the *lacZ* gene was placed into the *AflII* site of plasmid pRM25 by the same method as that used for the construction of plasmid pRM14, thereby creating the  $\phi(modE-lacZ)$  plasmid pRM26. The *bla* and *modE'-lacZ* from plasmid pRM26 were recombined with  $\lambda$ RZ5 to produce  $\lambda$ RM26.  $\lambda$ RM26 was transferred to the chromosome of strain BW545, probably to the *latt* site as previously described (36).

**Construction of ModE expression plasmids.** Two different *modE* fusion derivatives were constructed for the purpose of ModE overexpression and purification. The ModE expression plasmid pAG1 was constructed by ligating an 870-bp *EspI* fragment containing the promoterless *modE* gene from plasmid pFGH15 (Fig. 1) that had been modified by using the Klenow enzyme into the *SmaI* site of the *ptac*-based expression vector, plasmid pTrec99A (Pharmacia Biotech, Piscataway, N.J.). High-level expression of *modE* was not achieved with this construct; however, this plasmid was later used for mutagenesis experiments. A second *modE* expression plasmid, pRM22, which was used for the overexpression and purification of ModE was constructed by ligating an *EcoRI-HindIII* fragment from plasmid pRM12 into the *EcoRI-HindIII* site of plasmid pT7-7 (Pharmacia).

**Construction of a  $\Delta modE$ -Km and a *modF*-Km strain.** Strain SE1910 ( $\Delta modE$ -Km) was constructed by replacing the 570-bp *KpnI* fragment within the *modE* gene in plasmid pSE1004 (Fig. 1) with the *KpnI* fragment containing the *Km<sup>r</sup>* gene from plasmid pUC4K (Pharmacia), resulting in plasmid pRM9. The *PvuII-EcoRV* fragment from plasmid pRM9 was removed and ligated into the *BamHI* site of plasmid pMAK705 (12), which had been modified with Klenow enzyme, to yield plasmid pRM10. The *modE'-Km* region from plasmid pRM10 was recombined into the chromosome of strain BW545, creating strain SE1910. The *modF*-Km strain was constructed by ligating a 1.3-kb *HincII* fragment containing the *Km<sup>r</sup>* gene from plasmid pUC4K into the *BstBI* site in the *modF* gene in plasmid pRM13, which had been modified with Klenow enzyme, to yield plasmid pRM16. A 2.9-kb *EcoRV-ScaI* fragment from plasmid pRM16 was removed and ligated into the Klenow enzyme-modified *BamHI* site of plasmid pMAK705, resulting in plasmid pRM17. The *modF*-Km region of plasmid pRM17 was recombined into the chromosome of strain RK4353, producing strain SE1934.

**Construction of a *modAB'* plasmid.** A 1.7-kb *MscI* fragment containing the entire *modA* gene and the truncated *modB* gene (*modB'* lacks the C-terminal 10 amino acids) from plasmid pSE1009 (Fig. 1) was ligated into the *SmaI* site of vector plasmid pUC19, resulting in plasmid pRM1.

**Enzyme assays.**  $\beta$ -Galactosidase activity was assayed by measuring the rate of hydrolysis of *o*-nitrophenyl- $\beta$ -D-galactopyranoside (ONPG) as described by Miller (28). The activities are expressed as nanomoles of *o*-nitrophenol produced per minute per milligram of protein. The data presented are the averages for at least three independent experiments which varied by less than 15%. Formate hydrogenylase (FHL) activity was determined as described previously (36).

**Hydroxylamine mutagenesis of *modE*.** A 7.5- $\mu$ g amount of plasmid pAG1 DNA in 50  $\mu$ l of deionized H<sub>2</sub>O was mixed with 40  $\mu$ l of phosphate-EDTA buffer (0.5 M KPO<sub>4</sub>, pH 6.0; 5 mM EDTA) and 80  $\mu$ l of hydroxylamine-HCl solution, pH 6.0 (0.35 g of NH<sub>2</sub>OH-HCl, 0.56 ml of 4 M NaOH, 4.44 ml of distilled H<sub>2</sub>O) (6). This mixture was incubated at 37°C for 18 h. After the incubation period, the hydroxylamine was removed from the DNA by three successive 1:1 phenol extractions followed by a 1:1 chloroform extraction and alcohol precipitation of the DNA. Since the target strain, SE1811, has a lower level of transformation efficiency, the mutated DNA was first transformed into strain RK4353, which has a higher level of transformation efficiency. Total plasmid DNA isolated from strain RK4353 transformants was subsequently transformed into strain SE1811. Transformants were plated on lactose-MacConkey agar with and without 1 mM sodium molybdate. White colonies were picked from media lacking molybdate, and dark red colonies were selected from media containing molybdate. The levels of  $\beta$ -galactosidase activity produced by these mutants were determined to confirm the presumptive phenotypes. The entire *modE* gene sequences for the isolated *modE* mutants were determined.

**Purification of the ModE protein.** A 1-liter LB culture of strain BL21- $\lambda$ DE3(pRM22) was incubated at 37°C with vigorous shaking (250 rpm) until an optical density at 420 nm of 1.0 was reached. Isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) was then added to a final concentration of 0.5 mM, and the culture was incubated for an additional 2 h. The cells were harvested and resuspended

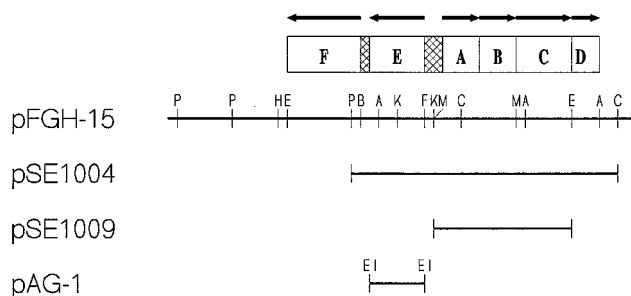


FIG. 1. Maps of plasmids used as sources of *mod* genes. The arrows show the directions of transcription of the indicated *mod* genes. A, *AseI*; B, *BstBI*; C, *ClaI*; E, *EcoRV*; EI, *EspI*; F, *AflII*; H, *HindIII*; K, *KpnI*; M, *MscI*; P, *PvuII*.

in 8 ml of extraction buffer (50 mM Tris-HCl, pH 7.5; 0.5 mM EDTA; 0.5% glycerol; 1 mM dithiothreitol). This cell suspension was passed through a French pressure cell (20,000 lb/in<sup>2</sup>) twice. The broken cell suspension was centrifuged at 100,000  $\times$  g for 60 min. The resultant supernatant was loaded on a 10-ml Q Sepharose fast-flow column (Pharmacia) equilibrated with 50 mM Tris buffer, pH 8. The column was washed with NaCl solution of increasing concentration in a stepwise manner. The ModE protein eluted with 0.2 M NaCl. The buffer was changed to 10 mM phosphate buffer, pH 7, by using an EconoPac-10 DG desalting column (Bio-Rad Laboratories, Hercules, Calif.) prior to loading of the protein sample on a 5-ml heparin column (HiTrap; Pharmacia) equilibrated with 10 mM phosphate buffer, pH 7. Proteins bound to the heparin column were eluted with NaCl solution of increasing concentration. ModE was eluted with 0.2 M NaCl. The ModE protein was concentrated by using a Centricon-20 spin cartridge (Polysciences, Warrington, Pa.) before the buffer was exchanged with 25 mM histidine-HCl buffer, pH 6.2, and the protein was then applied to a 10-ml polybuffer exchanger 94 column (Pharmacia) equilibrated with 25 mM histidine-HCl buffer, pH 6.2. The column was washed with polybuffer 74, pH 4.0, and ModE protein was eluted when the pH of the column reached approximately 4.5. The ModE protein buffer was exchanged with 10 mM phosphate buffer-10% glycerol, pH 7, and the final protein concentration of 0.42 mg/ml was obtained. All protein purification steps were conducted at 4°C. During the purification process, the ModE protein was monitored by the bicinchoninic acid protein determination assay as well as by electrophoresis in sodium dodecyl sulfate (SDS)-polyacrylamide gels, since the intrinsic properties of ModE did not lend the protein to conventional UV detection at 280 nm or detection by the Bradford assay. The identity of the purified protein as ModE was confirmed by N-terminal amino acid sequencing (Interdisciplinary Center for Biotechnology Research Protein Chemistry Core Laboratory at the University of Florida). The protein yield after each purification step is indicated in Table 2, and the purity of ModE after each purification procedure is shown in Fig. 2. About 5% of the protein in the crude extract was recovered as ModE protein. The protein obtained after the chromatofocusing step was judged to be pure by native polyacrylamide gel



FIG. 2. SDS-PAGE of proteins from different stages of ModE purification. Lane 1, molecular weight markers (phosphorylase b, 97,400; bovine serum albumin, 66,200; ovalbumin, 45,000; carbonic anhydrase, 31,000; trypsin inhibitor, 21,500; lysozyme, 14,400); lane 2, extract; lane 3, proteins after Q Sepharose fast-flow purification; lane 4, proteins after HiTrap Heparin purification; lane 5, proteins after chromatofocusing. A 25- $\mu$ g amount of protein was loaded into each of lanes 2 to 4. A 10- $\mu$ g amount of protein was loaded into lane 5.

← <i>modA</i>	
CGAGCCATTGGTAACCCCTTAATGTAACGTTATGTAGGGGACAATATAACGACTAAGTTCGGAAATTTCCAGCAGTTATTGCTAACCTTT	90
R A M	
GTGAGGTAGATAAGAAAAATATCGGCAGGAAAAGCAGGAAGTTGAGAAAAAGAAAATGCCCGACTAAGCGGGCATTACAGGGAATCAATG	180
→ ←	
ATTTGTCCGGCTGGTCTTTTTTACCAACACCAGAAAAGATGTTGAATACTTACCAAGACCGTAAATCAGACCCAGGATGATGGCCAT	270
CAGCAGGTACCATGATTACGGCGAATACCAGACTTTTTAATACTCTAACATGGTCAACTCCAGATATAGTCATGAGACTATTCTAAC	360
CGCTAAGCACAGAAAAGCACTCCCTTTTGTGCGGTGAGCTTTGCGTGGCGTTCTGTTTCCGTCACAATAAGACTTTTGCCAGGACATTG	450
→ <i>modE</i>	
TTATGCAGGCCGAAATCCCTTCTCACCCCTTAAGCTCCAACAAAAATATTTCGCCGACCCGCGCCGCTTTTCGCTACTAAAACACATTGCGC	540
M Q A E I L L T L K L Q Q K L F A D P R R I S L L K H I A L	
TTTCCGGTTCATTAGCCAGGAGCGAAAAGATGCCGGTATTAGCTATAAAAGCGCCTGGGATGCCATTAACGAGATGAATCAGTTAAGTG	630
S G S I S Q G A K D A G I S Y K S A W D A I N E M N Q L S E	
AGCATATCTGGTCGAGCGCGCAACAGGCGGTAAAGGTGGCGCGCGCAGTACTGACCCGCTATGGTCAGCGACTGATTACGCTCTATG	720
H I L V E R A T G G K G G G A V L T R Y G Q R L I Q L Y D	
ACTTACTGGCGCAATCCAGCAAAAAGCCTTTGATGTGTTAAGTACGATGACGCCCTGCGCTGAACAGCCTGCTGGCCGCGATCTCAC	810
L L A Q I Q K A F D V L S D D A L P L N S L L A A I S R	
GTTTTCTCACTGCAAAACAGCGCCCGTAAACAGTGGTTCGGTACCATCACCCCGCGATCATGATGACGTTCAACAGCATGTTGATGCT	900
F S L Q T S A R N Q W F G T I T A R D H D D V Q Q H V D V L	
TACTGGCTGACGAAAACACGCCCTGAAAGTCGCAATTACCCGACAAAGCGCGCGCTGCGGCTGGATGAAGGCAAGAAGTGTGA	990
L A D G K T R L K V A I T A Q Q S G A R L G L D E G K E V L I	
TATTGCTAAAAGCGCGTGGTGGTATTACTCAGGACGAGCGGTTCGCGCAAAACCGTGAACAACCAATTACCGGTATTATTAGTCATA	1080
L L K A P W V G I T Q D E A V A Q N A D N Q L P G I I S H I	
TTGAGCGCGCGCAGAGCAGTGCAGATTAATGGCGCTACCCGAGGGCAAACTGTGCGCCACAGTGCAGGTAATGAAGCGACTT	1170
E R G A E V L M A L P D G Q T L C A T V P V N E A T S	
CTCTTCAGCAAGGACAGAATGTCACGGCTACTTTAATGCCGACAGCGTATTATCGCCACGCTGTGCTAAGCGTGTGACAATTTGTTA	1260
L Q Q G Q N V T A Y F N A D S V I I A T L C *	
→ <i>modF</i>	
TGAAACACGTATCCCTGTCTAGTAATCGCTGCACAAAGTGGGATATAAAATGTCATCGTTGCAAATTTGCAAGGCACGTTTCGCTTAGC	1350
M S S L Q I L Q G T F R L S	
GACACAAAACCGTGCATTCGCCTCAGCTAACGTTAAACGGGGTGTAGTGGCGTTCGCGTTCGAATGGAAGCGGGAATCCGGCC	1440
D T K T L Q L P Q L T L N A G D S W A F V G S N G S G K S A	
CTGGCCCGCGCGCGGGGAACTCCGCTTTTGAAGGTGAACGGAAAGCCAGTTTCCACATCAGCTCCTCTCCTCCGAGCAA	1530
L A R A L A G E L P L L K G E R Q Q S Q F S H I T R L S F E Q	
TTGCAAAAACCTCGTCAGCGAATGGCAGCGAATAACCGGATGCTCGGCCAAGATGACACCGGACGCACTACCGCTGAG	1620
L Q K L V S D E W Q R N N T D M L G P G E D D T G R T T A E	
ATCATTCCAGGATGAAGTAAGGATGCACCGCTTGCATGCACTGGCGCAGCAGTTCGGTATTACCGCCCTCCTCGACCGAGCTTTAAA	1710
I I Q D E V K D A P R C M Q L A Q Q F G I T A L L D R R F K	
TACCTTTCCACTGGCGAGCGCAAAAACCTGCTGTGTGTCAGGCGTGTGTCGGAGCCTGACTGTTGATTCTTGATGAGCGTTCGAT	1800
Y L S T G E R K T L L C Q A L M S E P D L L I L D E P F D	
GGCCTGGATGTGCCTCAGCTCAGCAGCTGGCTGAGCGACTCGCCTCGTTACATCAGTCCGGTATTACTCTGGTACTGGTCAATCCG	1890
G L D V A S R Q Q L A E R L L A S L H Q S G I T L V L V L N R	
TTGATGAGATCCCGAGTTTGTCCAGTTTGTGCGGTGCTGGCGGATTCAGCTTAGCGGAAACTGGCGCTAAAGAGGAACCTCCAA	1980
F D E I P E F V Q F A G V L A D C T L A E T G A K E E L L Q	
CAAGCCTCGTCGCGCAACTGGCGCATAGTGAACAGCTTGAAGGTGTCGAACCTGCGGAGCCGGATGAACCTTCAGCAGCTCAGCCCTTA	2070
A C L V A H A S E Q L E G V Q L P E P D E P S A R H A L	
CCGCCCACGAAACCGCGCATTTGTGCTGAACAATGGCGTGGTTTCTTATAACGATCGCCCCATTCTTAATAACCTTAGCTGGCAGGTGAAT	2160
P A N E P R I V L N N G V V S Y N D R P I L N N L S W Q V N	
CCAGGCGAACACTGGCAAATTTGTCGGGCCAAATGGTCAGGAAAATCGACGTTATTAAAGCCTGGTTACTGGCGATCATCCGCAAGGTAC	2250
P G E H W Q I V G P N G A G K S T L L S L V T G D H P Q G Y	
AGCAACGATTTGACGCTTTTCGGACGAGCTCGCGGACGCGCGAACCATCTGGCCATCAAAAAGCATATCGGTTACGTCAGCAGTAGT	2340
S N D L T L F G R R R G S G E T I W D I K K H I G Y V S S S	
TTGACTCTGGATTACCGGGTCAGCACTACCGTGCCTAATGTGATCTTTCTGGCTATTTGATTGCGATTGGCATTATCAGGCGGTTTCG	2430
L H L D Y R V S T T V R N V I L S G Y F D S I G I Y Q A V S	
GATCGCCAGCAAAAACCTGTCAGCAGTGGCTGGATATTCTCGGATTGATAACGACGCGCTGACGCTCCGTTCCATAGCTTTTCTCTGG	2520
D R Q Q K L V L D I L G I D K R T A D A P F H S L S W	
GGACAGCAGCTGTCGGCGTATGTCGCGCACTGGTGAACACTCGGACGTTGCTTATTCTCGATGAACCACTACAGGGGGTTGATCCG	2610
G Q Q R L A L I V R A L V K H P T L L I L D E P L Q G L D P	
CTGAATCGCCAGCTTATCCGCCGTTTGTGATGCTGATTAGCGAAGGTGAACGCAATTTGTTGTTTTCGACCAGCTGAAGAT	2700
L N R Q L I R R F V D V L I S E G E T Q L L F V S H H A E D	
GCGCCTGCCTGATTACCCATCGTCTTGTGTTGTCGCGGACGGTGGACTCTATCGCTATGTGTCACAAAATATATTGAGTCGGTAGT	2790
A P A C I T H R L E F V P D G G L Y R Y V L T K I Y *	
GCTGACCTTGGCGGAGCGGCTTAGCACCCCTCTCCGCCAACGTTTCGACGATGACGAGCATGAAACCGCGCTTTTTTTCAGATAAAA	2880
GCGCAATCATTATAAACCTCTGTTTTATAATCACTTAATCGGCATAAAAAAGGCTAAATTCCTGTGTAACGATTCCACTAATTTA	2970
→ <i>gale</i>	
TTCCATGTCACACTTTTCGCATCTTGTATGCTATGTTATTTTCATACCATAAGCCATAATG	3033

FIG. 3. DNA sequences of *modE* and *modF*. The amino acid sequences are listed below the DNA sequence. Asterisks represent the translation terminations. Presumptive -10 and -35 sequences are underlined, and the ribosomal binding sites are doubly underlined. The location of the *lacZ* insertion in the *modE* gene in plasmids pRM14 and pRM26 is shown by the down arrowhead. The *modA*-42mer (positions 3 to 44) and *modF*-42mer (positions 1362 to 1403) oligonucleotide sequences used in *in vitro* transcription-translation and DNA mobility shift experiments are highlighted by boldface type. The 8-bp inverted repeats present in the oligonucleotides are shown in inverse print. The ATP-GTP binding motif present in ModF is in boldface type. The stem-loop structure between the *modABCD* and *modEF* operons is indicated by the two opposing arrows.

electrophoresis (PAGE) and SDS-PAGE (Fig. 2) and N-terminal amino acid sequencing of the protein.

**Mobility shift experiments.** Mobility shift experiments (10) were performed as outlined below. Two 42-bp oligonucleotides containing the putative repressor binding region, one derived from the operator region of *modA* (*modA*-42mer) and the other derived from the 5' end of *modF* (*modF*-42mer DNA) (Fig. 3), were synthesized (Anagen, Palo Alto, Calif.). The *modA*-derived oligonucleotide contained a *Bst*EII site, while an *Eco*RI site was engineered into the *modF*-derived oligonucleotide. After digestion of the oligonucleotides with the appropriate endonucleases, Klenow enzyme was used to label the DNA with <sup>35</sup>S-

dATP. Removal of contaminating enzyme and dATP was achieved by phenol extraction and alcohol precipitation of the oligonucleotides. Binding reactions were initiated by mixing the appropriate amounts of labeled DNA and ModE with 10 mM Tris-HCl, pH 7.9–10 mM MgCl<sub>2</sub>–50 mM NaCl–1 mM dithiothreitol–10% glycerol. When appropriate, sodium molybdate was added to a final concentration of 1 mM. The reaction mixtures were incubated at 30°C for 30 min prior to being loaded onto a 5% polyacrylamide gel in Tris-borate-EDTA buffer (30). For certain experiments, 1 mM sodium molybdate was incorporated into the gel and the electrophoresis buffer. The gel was prerun at 100 V, at room temperature, for at least 90 min to stabilize the current prior to loading of the

TABLE 3. Complementation of mutation in strain SE1811 for molybdate-dependent repression of the *modABCD* operon

Strain	Relevant genotype	$\beta$ -Galactosidase activity <sup>a</sup>		FHL <sup>b</sup>
		Without molybdate	With molybdate	
SE2069	$\phi(modA-lacZ)$	1,300	200	-
SE2069(pSE1004)	$\phi(modA-lacZ)$ <i>modABCDE</i> <sup>+</sup>	300	300	+
SE2069(pSE1009)	$\phi(modA-lacZ)$ <i>modABC</i> <sup>+</sup>	300	300	+
SE1811	$\phi(modA-lacZ)$ <i>modE</i>	1,500	1,500	-
SE1811(pSE1004)	$\phi(modA-lacZ)$ <i>modE</i> <i>modABCDE</i> <sup>+</sup>	290	260	+
SE1811(pSE1009)	$\phi(modA-lacZ)$ <i>modE</i> <i>modABC</i> <sup>+</sup>	1,500	1,500	+
SE1811(pAG1)	$\phi(modA-lacZ)$ <i>modE</i> <i>modE</i> <sup>+</sup>	600	100	-

<sup>a</sup> Cells were grown in LBG medium. The molybdate concentration was 1 mM.  $\beta$ -Galactosidase activities are expressed as nanomoles minute<sup>-1</sup> milligram of cell protein<sup>-1</sup>.

<sup>b</sup> -, absent; +, present.

samples. The running buffer was continuously recirculated during the electrophoresis by using a Minipuls II pump (Gilson). After completion of the electrophoresis, the gel was transferred to Whatman filter paper and dried under a vacuum. Autoradiography was used to visualize the resulting oligonucleotide-protein migration patterns. The amount of radioactivity in each band was determined by using a phosphorimager (Molecular Dynamics, Sunnyvale, Calif.).

**In vitro transcription-translation experiments.** The in vitro transcription-translation experiments using S30 extract and L-[<sup>35</sup>S]methionine to monitor the expression of *modA* and *modB'* from plasmid pRM1 in the presence and absence of ModE and 1 mM molybdate were conducted as suggested by the supplier of the S30 extract (Promega, Madison, Wis.). Three micrograms of plasmid pRM1 DNA and, when appropriate, 40 pmol of ModE were included in the S30 reaction mixtures. Proteins in a sample of each S30 reaction mixture (50,000 cpm) were separated in an SDS-15% polyacrylamide gel, and the resultant reaction products were visualized by autoradiography. Quantification of the labelled proteins produced was accomplished by using a phosphorimager.

## RESULTS AND DISCUSSION

**Evidence for the existence of a regulator of the *modABCD* operon.** When the strain with the *modA-lacZ* fusion, strain SE2069, was assayed for  $\beta$ -galactosidase activity after growth in LBG with or without 1 mM sodium molybdate, a high level of  $\beta$ -galactosidase activity was obtained only with cells grown in medium lacking molybdate (Table 3). Furthermore, introducing plasmids pSE1004 and pSE1009 carrying different regions of the *modABCD* operon into strain SE2069 resulted in a low level of  $\beta$ -galactosidase activity production regardless of whether molybdate was present in the growth medium. Strain SE2069 with the plasmids produced FHL activity, which is in agreement with the Mod<sup>+</sup> phenotype. These results show that saturating the molybdate pool in the cytoplasm, as determined by the ability to produce FHL activity, leads to repression of the *modABCD* operon.

Strain SE1811 produced high levels of  $\beta$ -galactosidase activity even in the presence of molybdate (Table 3). When strain SE1811 was transformed with plasmid pSE1009, the resulting strain, SE1811(pSE1009), produced FHL activity, suggesting the restoration of the Mod<sup>+</sup> phenotype. However, the *modA-lacZ* fusion in strain SE1811 was still derepressed, indicating a lack of a second component, besides molybdate, required for the molybdate-dependent repression of *modA-lacZ*. Strain SE1811(pSE1004) produced low levels of  $\beta$ -galactosidase activity. In light of these findings, it was proposed that a region of DNA present in pSE1004 but not in pSE1009 is responsible for

complementation of the second mutation in strain SE1811 for the molybdate-dependent repression of *modA-lacZ*. To determine if this is indeed the case, a plasmid containing this region of DNA, plasmid pAG1 (Fig. 1), was constructed and transformed into strain SE1811. The resulting  $\beta$ -galactosidase activities produced by strain SE1811(pAG1) indicated that this plasmid restored the molybdate-dependent repression of *modA-lacZ*. Plasmid pAG1, as expected, did not complement the *modA* mutation for FHL production. The lower level of  $\beta$ -galactosidase activity obtained with the cultures grown in LBG medium could be a consequence of the higher copy number of the repressor gene.

***modE* gene.** Sequencing the DNA that was responsible for the restoration of the molybdate-dependent repression of *modA-lacZ* in strain SE1811 revealed two open reading frames starting 444 bp upstream of the *modA* translation start site (Fig. 3). These two open reading frames, denoted *modE* and *modF*, are transcribed in the opposite direction in comparison with *modA* and encode a 262-amino-acid protein (28,200 Da) and a 490-amino-acid protein (53,900 Da), respectively. The expected mass of ModE was in agreement with the value obtained after SDS-PAGE (29,000 Da) (Fig. 2). Matrix-assisted laser desorption ionization time-of-flight mass spectrometry also gave an apparent molecular mass of 28,271 Da for ModE. The translation start site for *modE* was confirmed by the N-terminal amino acid sequence of purified ModE. The first 12 amino acids, MQAEILLTLKIQ, matched the amino acids predicted on the basis of the DNA sequence of *modE*.

A search for proteins that bear similarity to the *E. coli* ModE protein resulted in the detection of four homologs (Fig. 4). These proteins were identified as ModA from *A. vinelandii* (35% identity, 45% similarity) (23) (renamed ModE [29]), ModE from *H. influenzae* Rd (48% identity, 58% similarity) (9), MopA from *R. capsulatus* (34% identity, 46% similarity) (45), and MopII from *Clostridium pasteurianum* (29% identity, 46% similarity) (17). A possible helix-turn-helix motif which is a known prokaryotic DNA-binding protein structure motif (31) has been identified for the *E. coli* ModE, as indicated in Fig. 4.

***modF* gene.** The predicted size of ModF (54 kDa) was also verified by SDS-PAGE of the protein produced after overexpression of ModF by using a *ptac* system (plasmid pRM23). Although no protein similar to ModF was found during database searches, analysis of the *modF* gene revealed the presence of an ATP-GTP binding site motif (GSNGSGKS) (Fig. 3).

To date, the function of ModF has not been elucidated, as deletions in the *modF* gene do not result in a detectable phenotype for FHL production or *modA* gene synthesis or activity under the conditions used in this study. Dorrell et al. (8) have described and presented the sequence of a gene involved in photoreactivation (*phrA*). This DNA sequence resides within the *modF* gene described in this study. On the basis of the DNA sequence, these investigators proposed that PhrA is a 38-kDa protein involved in photorepair. The facts that a 54-kDa and not a 38-kDa product was found when the *modF* DNA was expressed in a T7-driven expression system and that expression of a plasmid-borne  $\phi(modF-lacZ)$  responds to the presence of molybdate (high levels of *modF-lacZ* expression are attained under low-molybdate conditions; data not shown) suggest that this segment of DNA codes for the *modF* gene.

**Regulation of *modA* by ModE.** The ModE protein was used in an in vitro transcription-translation experiment with the objective of determining whether expression of the *modABCD* operon is repressed in the presence of ModE and molybdate. To this end, plasmid pRM1, containing *modAB'* was incubated with *E. coli* S30 extract with and without 1 mM molybdate and 40 pmol of ModE.

AvModE	MTATRFLARMSLDTDVG--TALSDTRIRLLEAIEREGSINRAAKVVPLSYKAAWDAIDTMNN	60
EcModE	MQAEILLTLKLQOK--LFADPRRISLLKHIALSGSISOGAKDAGISYKSAWDAINEMNQ	57
HiModE	MKNTEILLTIKLQQA--LFIDPKVRLLKEIQCCGSINQAAKNAKVSYSKSAWDHLEAMNK	58
RcMopA	MNEQPLIAALSILQRAGAPRVGGDRIRLLEAIARHGTIAGAAREVGLSYKTAWDAVGTLNN	60
AvModE	LAPEPLVVRVAGGRQGGGTQLTDYGRIRIVAMYRALEIEYQSALDRLSERLNEVTGGDIQA	120
EcModE	LSEHILVERATGGKGGGAVLTRYGQRLIQLYDLLAQIQKAFDVLSDDDALPLNSLLAA	117
HiModE	ISPRPLLERNTGGKNGGTALTYYAERLLQLYDLLERTQEHAFHILQDE-SVPLDSLTA	117
RcMopA	LFEQPLVEAAPGGRTGGNARVTEAGQALIAAGFLLEGALTKALGVLEGGVSAPEKALNTL	120
AvModE	FQRLMHSMKTSARNQFAGIVTGLRVGGVDYEVRI RL-DAENEIAAVITKASAENLELA	179
EcModE	ISRF---SLQTSARNQWFGTITARDHDDVQQHVDVLLADGKTRLKVAITAQSGARLGLD	173
HiModE	TARF---SLQSARNQFFGRVAQQRIDSRVVDVNVQGLPTPLQVSIITTKSSARLKLI	173
RcMopA	WSL---TMRTSNRNTLRCTVTRVTLGAVNAEVELALTDGHS-LTAVITERSATEMGLA	174
CpMopII	MSISARNQLKKGKVVGLKKGVVTAEVVLEIAGGN-KITSIISLDSVEELGVK	50
AvModE	IGKEVFALVKSSVMLTTEPSLKL-TARNQLWGEVIDIHEGPNNEVTALPGRSVTCV	238
EcModE	EGKEVLILLKAPWVGITQDEAVAQ-NADNQLPGIISHIERGAEQCEVLMALPDGQTL CAT	232
HiModE	TEKEVMLMFKAPWVKISEQPLA--NQPNQFPVNIKSLN---EEEAILQFAESNIEFCAT	219
RcMopA	PGVEVFALIKASFVMLAAGGDPGRISACNRLTGIVAARTDGPVNTEIILDGNCCKSITAV	234
CpMopII	EGAELTAVVKSTDVMILA	68
AvModE	VTADSKALGLAPGVAACAFFKSSSVILAVYG	270
EcModE	VPVN--EATSLQGGQNVITAYFNADSVIIATLC	262
HiModE	V-H--QPNQWQIEQQVWIHIDQEQIILATLG	255
RcMopA	ITHTSADALGLAPGVPATALFKASHVILAMP	265

FIG. 4. Amino acid sequences of homologs of the *E. coli* ModE protein. AvModE, *A. vinelandii* ModE (23); EcModE, *E. coli* ModE; HiModE, *H. influenzae* Rd ModE (9); RcMopA, *R. capsulatus* MopA (45); CpMopII, *C. pasteurianum* MopII (17). Double dots represent identity, and single dots indicate conservative substitutions. Inverse print marks the conserved SARNQ sequence, and the bold italics indicate the amino acids that have been mutated in the *E. coli* ModE mutant proteins. All identity and similarity designations are in relation to *E. coli* ModE. A putative helix-turn-helix DNA binding region in EcModE is underlined.

On the basis of the resultant autoradiogram (Fig. 5) and phosphorimager data, it can be concluded that the presence of ModE reduced the production of the 26-kDa ModA protein and the 24-kDa ModB' protein by 48 and 65%, respectively (lanes 1 and 2). With the addition of molybdate, this repression of *modAB'* by ModE became even more pronounced (greater than 90% decrease in the production of both ModA and ModB') (Fig. 5, lanes 3 and 4). The decrease in the levels of ModA and ModB' production even in the absence of added molybdate could be due to contaminating molybdate present in the commercial S30 extract. It should be noted that although equal amounts of radioactivity were loaded in all lanes, the sample containing both ModE and molybdate (Fig. 5, lane 4) had less labelled protein than the other lanes. This disparity in amounts of labelled protein is due to the presence in this sample of low-molecular-weight (partially degraded?) protein products which migrated faster than the dye front during electrophoresis. The reduction of *modABCD* expression in the in

vitro transcription-translation reaction mixtures containing ModE shows that ModE is a repressor of the *modABCD* operon, and this repression is enhanced by molybdate. Furthermore, addition of *modA*-42mer DNA (Fig. 3) to the reaction mixtures containing molybdate and ModE reversed the repression of *modAB'* by ModE. A 50-fold excess of the oligonucleotide (compared with the amount of plasmid pRM1) restored the production of ModA to 66% of the levels for the reaction without ModE (data not shown). Increasing the oligonucleotide concentration to a 100-fold excess completely restored *modAB'* production, indicating that the *modA* operator region oligonucleotide DNA is capable of titrating the ModE protein and relieving repression of transcription through the plasmid-borne *modA* gene. These results also show that the ModE protein functions as a classical repressor by interacting with the DNA between the transcription and translation start sites of the *modABCD* operon.

**Binding of ModE to target DNA.** Since the in vitro transcrip-

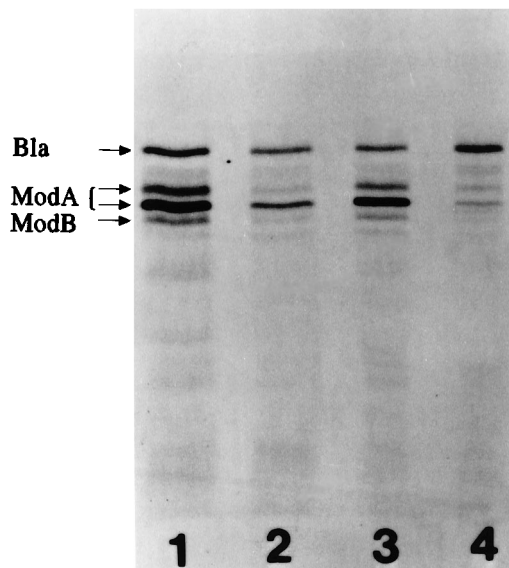


FIG. 5. SDS-PAGE analysis of proteins produced in an in vitro transcription-translation experiment in which plasmid pRM1 served as the template DNA. Molybdate and ModE concentrations were 1 mM and 40 pmol per reaction mixture, respectively. Lane 1, without ModE; lane 2, with ModE; lane 3, without ModE, with molybdate; lane 4, with ModE, with molybdate. Bla,  $\beta$ -lactamase. The upper ModA protein band represents the precursor protein bearing the signal peptide, while the lower ModA protein band represents the protein with the signal peptide removed. See text for details.

tion-translation experiments have shown that ModE does act as a repressor of the *modABCD* operon, it became important to determine directly whether ModE also manifests DNA binding capabilities. The DNA binding capabilities of ModE were demonstrated by using DNA mobility shift experiments (Fig. 6). Two different oligonucleotides were used in these DNA mobility shift experiments. One of the oligonucleotides was the *modA*-42mer DNA. This 42-bp DNA carries the inverted repeat and CAT sequences (CA TAAC · GTTA CATT AA [Fig. 3]) (also see references 26 and 35). For the purpose of distinguishing whether the inverted repeat or the flanking CAT sequences are required for binding ModE, a second 42-bp-long oligonucleotide derived from the *modF* gene containing only the TAAC · GTTA inverted-repeat sequence (Fig. 3) (*modF*-42mer DNA) was also included in the mobility shift experiments.

As the data presented in Fig. 6 indicate, the addition of ModE to both the *modA*-42mer and the *modF*-42mer in the presence of 1 mM molybdate caused a distinct shift in the migration of the oligonucleotides (Fig. 6, lanes 2 and 4), leading to a unique band containing the DNA-protein complex. However, only a small percentage (6%) of *modF*-42mer was complexed with ModE compared with the percentage for *modA*-42mer (50%). The difference in the amounts of binding between the two oligonucleotides and ModE must correlate with the differences in the target binding sequences found within the oligonucleotides.

When the mobility shift experiment was performed in the absence of molybdate (Fig. 6, lanes 5 to 8), a lesser amount of *modA*-42mer was retarded by binding with ModE and there was no detectable *modF*-42mer-ModE complex formed. Curiously, in lane 6 of the gel shown in Fig. 6 a second shifted band (upper band) not present in the lane containing molybdate (lane 2) can be observed; this band represented 19% of the complexed oligonucleotide (38% total for both shifted bands

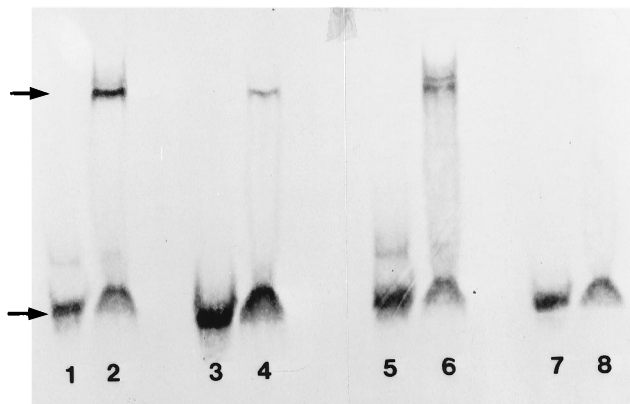


FIG. 6. Binding of ModE to appropriate target *mod* DNA. Lanes 1, 2, 5, and 6 contain *modA*-42mer (Fig. 3), and lanes 3, 4, 7, and 8 contain *modF*-42mer (Fig. 3). Lanes 2, 4, 6, and 8 contain ModE (80 pmol). Lanes 1 to 4 contain molybdate (1 mM), and lanes 5 to 8 contain no added molybdate. All samples contained 100,000 cpm of DNA. The top arrow indicates the ModE-bound DNA, and the bottom arrow marks the unbound DNA.

combined). The nature of this second band is not known. However, it is possible that this band represents a molybdate-independent complex possibly involving the CAT sequences or a dimer of ModE bound to the DNA. Furthermore, the smear below the DNA-protein complex observed in lane 6 of Fig. 6 suggests that in the absence of added molybdate, ModE and *modA*-42mer are forming only loose associations which are disrupted during electrophoresis. Thus, the mobility shift experiments demonstrated binding of ModE in the presence of molybdate to the 8-bp inverted repeat (TAAC · GTTA), with enhanced binding demonstrated for the DNA containing the flanking CA and CATTAA sequences as well.

The requirement of the CATAA sequence for effective repression was also reported by Rech et al. (35). In those experiments, the putative *modABCD* operon repressor binding region (CTACATAACGTTACATTAA) was mutated at specific sites (underlined) and the resultant expression of *modA-lacZ* was determined. Replacement of the C and last two A's of the first CATAA sequence with G, C, and T, respectively, reduced overall expression by twofold compared with wild-type expression but completely eliminated molybdate-induced repression. Singly mutating the C to a G and the second A from the CATAA sequence to a T resulted in partial repression of  $\phi(modA-lacZ)$ . Mutating bases 5' to the first CATAA sequence and bases constituting the second CATTAA sequence did not alter the regulation of *modA*. From these experiments, it can be surmised that of the two CAT boxes, only the first CATAA sequence is necessary for binding the repressor. However, the data do not resolve whether the ModE binding domain contains only the CATAA sequence or the overlapping 8-bp inverted repeat (underlined) (CATAAC · GTTA) or both.

**Implications of the differential binding of ModE to the two binding sites.** The fact that ModE is capable of binding to DNA containing the TAAC · GTTA inverted repeat which is not flanked by the CAT boxes (*modF*-42mer DNA), albeit at a much reduced level compared with the binding of the CAT box-flanked inverted repeat (*modA*-42mer DNA), recommended the analysis of *modABCD* DNA for the presence of other similar inverted repeats which might serve as potential sites of additional regulation. This search identified two candidate sequences, one positioned 17 bp upstream of the *modB*

Plasmid	Mutation	$\beta$ -Galactosidase Activity	
		-Mo	+Mo
pAG1	1 <span style="float: right;">262</span>	600	100
pAG1(A76V)	A76V	1,300	600
pAG1(T125I)	T125I	125	110
pAG1(G133D)	G133D	40	40
pAG1(Q216*)	Q216*	<10	<10

FIG. 7. Analysis of ModE mutants.  $\beta$ -Galactosidase activity produced by strain SE1811 carrying the indicated mutated pAG1 plasmid derivatives was determined and expressed in nanomoles minute<sup>-1</sup> milligram of cell protein<sup>-1</sup>. Cells were grown in LBG medium with (+) or without (-) sodium molybdate (Mo) (1 mM). \* represents the translation termination.

translation start site (TAAAC · GTTA) and another inverted-repeat sequence (AAAC · GTTT) 199 bp within the *modD* gene (26). The presence of the inverted repeat 5' to the *modB* translation start site is of particular interest since ModB is an integral membrane protein and as such must be carefully regulated. Results from the in vitro transcription-translation experiment (Fig. 5), in which a greater reduction in the production of ModB' than of ModA occurred, provides some evidence that there is additional modulation of *modB* expression, seemingly mediated by ModE. Although the functions of ModD and ModF, and therefore the need for their further regulation by ModE, are unclear,  $\phi(modF-lacZ)$  expression levels were higher by approximately fourfold when the cells were grown in low-molybdate-glucose-minimal medium in comparison with levels for cells grown in molybdate-supplemented medium, indicating molybdate-induced repression of *modF* expression (data not shown). Thus, these data suggest that ModE may exert its regulatory control in one of two ways: (i) transcriptional control by binding to the *modA* operator target sequence or (ii) additional modulation of expression of other *mod* genes by binding to the 8-bp inverted repeats.

**Analysis of ModE mutants.** Hydroxylamine, which causes G-to-A transitions, was used to mutate the DNA coding for ModE (plasmid pAG1) in order to identify the critical amino acids responsible for the interaction with DNA and/or molyb-

date. After mutagenesis, two different mutant phenotypes were identified: partial loss of repression of the *modABCD* operon in the presence of molybdate and complete repression of the *modABCD* operon in the absence of molybdate. Strain SE1811 containing these mutant plasmids were assayed for  $\beta$ -galactosidase activity, and the results are presented in Fig. 7.

The partially derepressed ModE mutant (A76V) contains a replacement of alanine at amino acid position 76 with valine. Three mutants isolated as "superrepressors" repressed transcription of *modA-lacZ* even in the absence of molybdate. Two of these mutants harbor the following mutations: threonine at position 125 changed to isoleucine (T125I) and glycine at position 133 replaced by aspartic acid (G133D). In the third mutation (Q216\*), the C-terminal 47 amino acids starting from the amino acid at position 216 were deleted.

In the partially derepressed (A76V) ModE mutant, the mutation lies immediately after a stretch of three glycines which are identical in three of the four homologs, the fourth having two of the three glycines (Fig. 4). Two of the four proteins have alanine at this position, while the other two have threonine. Two of the superrepressor mutants, T125I and G133D, have mutations situated near a 5-amino-acid-long sequence, SARNQ (amino acids 126 to 130), which is conserved among the five homologs, with the exception of *R. capsulatus* MopA, which has an asparagine in place of alanine and a threonine in place of the glutamine. In fact, the T125I mutation precedes the SARNQ sequence and the threonine is itself conserved among three of the homologs, with serine substituting for threonine in the *H. influenzae* protein. The G133D mutation lies 2 amino acids further on the C-terminal side of the SARNQ sequence. The glycine also is conserved in four of the proteins. The glutamine which is altered in the third superrepressor mutant (Q216\*) is unique to *E. coli* ModE at this position; however, there are several regions of conserved amino acids in the missing C-terminal end of ModE whose loss contributes to the phenotype of the mutant.

**Molybdate binding domain.** Given that the SARNQ region is conserved among four of the five homologs and partially conserved in the fifth and that two of the superrepressor mu-

TABLE 4. Regulation of *modE* in different *mod* strains

Strain	Relevant genotype	$\beta$ -Galactosidase activity <sup>a</sup>	
		Without molybdate	With molybdate
SE1938	$\phi(modE-lacZ)$	324	297
SE1952	$\phi(modE-lacZ) modB::Tn5$	316	249
SE1942	$\phi(modE-lacZ) \Delta(modE-Km)$	351	324
SE1940	$\phi(modE-lacZ) modF-Km$	378	351

<sup>a</sup> Cells were grown in LBG medium with or without sodium molybdate at a final concentration of 1 mM. Enzyme activities are expressed in nanomoles minute<sup>-1</sup> milligram of cell protein<sup>-1</sup>.



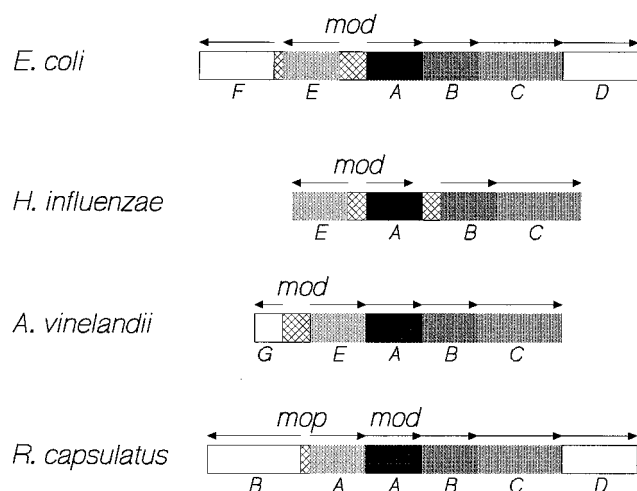


FIG. 8. Organization of genes coding for molybdate transport system in various organisms. Similar shading represents similar genes or functions. Unshaded genes are not similar. Arrows indicate the direction of transcription of the various *mod* genes.

tations are situated near the SARNQ region, it is reasonable to conclude that these amino acids are important for the proper functioning of ModE. Moreover, the phenotype associated with these mutations suggests that these amino acids are involved with activating the aporepressor by means of binding molybdate. Also, binding of molybdate by ModE would most likely require positively charged amino acids like those present in the SARNQ sequence.

Regions of sequence similar to SARNQ have also been found in other proteins that are known to interact with molybdate. For instance, the NarX protein, which functions to activate the regulator protein, NarL, in response to the presence of nitrate and molybdenum, contains the sequence SGRNE (amino acid positions 204 to 208) (19). Interestingly, specific mutations in the SGRNE region result in the molybdate-independent activation of NarL (19). Another nitrate sensor protein, NarQ, has the sequence SLRMQ (amino acid positions 48 to 52) (3), which is similar to the general motif. Furthermore, the rat estrogen receptor-hsp90 complex which is stabilized by molybdate (37) contains two sequences which bear similarity to SARNQ, one defined as PATNQ which begins at amino acid position 227 and a second sequence, LDRNQ, starting at amino acid position 415 (24). The human estrogen receptor also has a similar sequence, LARGQ, located at amino acid positions 190 to 194 (20). Taken together, these data show that the molybdate binding domain (motif) contains the amino acid sequence SARNQ and in ModE this domain is extended to include additional amino acids [(T/S)SARNQXXG)].

**Regulation of *modE*.** The regulation of *modE* was investigated by introducing  $\phi(modE-lacZ)$  borne on  $\lambda$ RM26 into various backgrounds and assaying for  $\beta$ -galactosidase activity. The results presented in Table 4 indicate that there is little difference in the levels of  $\beta$ -galactosidase activity obtained from a wild-type strain, SE1938; a *modB* mutant strain, SE1952; and a *modE* mutant strain, SE1942. These results indicate that *modE* is transcribed constitutively at low levels, which is consistent with its proposed role. A marginal increase in *modE-lacZ*-dependent  $\beta$ -galactosidase activity was observed for the fusion present in the *modF* mutant, strain SE1940. This suggests that there may be some form of interaction between

ModE and ModF, which is not surprising given that they are transcribed from the same operon and that ModE functions as a regulatory protein.

A plasmid-based  $\phi(modE-lacZ)$ , pRM14, was used to determine the effect of mutations in the global regulators, *fnr*, *arcA*, and *narL* gene products. Similarly, mutations in genes required for molybdate incorporation into molybdopterin, *moa*, *mob*, *moe*, and *mog*, were analyzed for their effect on *modE* expression (data not shown). No significant change in *modE* expression was detected in any of these backgrounds.

**Organization of the molybdate-specific transporter genes in several genera of bacteria.** Since organisms other than *E. coli* have been shown to possess molybdate-specific transporters (9, 23, 45), a comparison of the organizations of the genes coding for the molybdate transport machinery could shed light on the regulation of these genes as well. In *E. coli* the genes coding for the molybdate-specific transport proteins (*modABC*) are transcribed in the same direction from a single promoter. The repressor of the *modABCD* operon encoded by *modE* is transcribed in the opposite orientation relative to that of the other *mod* genes, with its translation start site situated 444 bp from the *modA* translation start site. An 8-bp stem and 6-bp loop structure which may serve to further spatially separate the two operons has been identified and is highlighted in Fig. 3.

An analysis of the organization of the molybdate transport genes from *H. influenzae* reveals an organization similar to that found in *E. coli* (Fig. 8), where the *modABC* genes are transcribed in the opposite orientation relative to that of *modE*. However, there does not appear to be a stem-loop structure between the two genes. It is also of interest that there is a 9-bp sequence, TAAC · GTTTA, which is located 17 bp (compared with 16 bp in *E. coli*) upstream of the translation start site of the *modA* gene in *H. influenzae* and that this inverted repeat is not flanked by the CA and CATTAA sequences (9). This sequence is similar to the inverted repeat which serves as the repressor binding site in *E. coli* except that there is an extra thymine. In the nitrogen-fixing organisms *A. vinelandii* and *R. capsulatus*, *modE* or, in *R. capsulatus*, *mopA* is transcribed in the same direction as *modA*. A search for sites in these two organisms' *mod* DNAs which are similar to the putative ModE repressor binding site resulted in the detection of only a GATAA sequence upstream of the *modA* gene in *A. vinelandii* (23). No similar sequences were found for *R. capsulatus* (45). The differences in the organization and presumably the regulation of the *mod* genes in *E. coli* compared with *A. vinelandii* and *R. capsulatus* are understandable given that the nitrogen fixation process undertaken by the last two organisms is a molybdate-requiring process.

#### ACKNOWLEDGMENTS

We thank H. Schweizer and V. Stewart for providing the strains used in this study and J. F. Preston for performing the matrix-assisted laser desorption ionization time-of-flight mass spectrophotometry experiment.

This work was supported by a Public Health Service grant from the National Institutes of Health (GM48667).

#### REFERENCES

- Allen, R. M., R. Chatterjee, M. S. Madden, P. W. Ludden, and V. K. Shah. 1994. Biosynthesis of the iron-molybdenum cofactor of nitrogenase. *Crit. Rev. Biotechnol.* **14**:225-249.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tools. *J. Mol. Biol.* **215**:403-410.
- Chiang, R. C., R. Cavicchioli, and R. P. Gunsalus. 1992. Identification and characterization of NarQ, a second nitrate sensor for nitrate-dependent gene regulation in *Escherichia coli*. *Mol. Microbiol.* **6**:1913-1923.
- Chou, P. Y., and G. D. Fasman. 1978. Prediction of the secondary structure of proteins from their amino acid sequence. *Adv. Enzymol.* **47**:45-147.

5. Corcuera, G. L., M. Bastidas, and M. Dubourdieu. 1993. Molybdenum uptake in *Escherichia coli* K12. *J. Gen. Microbiol.* **139**:1869–1875.
6. Davis, R. W., D. Botstein, and J. R. Roth. 1980. A manual for genetic engineering. Advanced bacterial genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
7. Devereux, J., P. Haeberli, and O. Smithies. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12**:387–395.
8. Dorrell, N., A. H. Ahmed, and S. H. Moss. 1993. Photoreactivation in a *phrB* mutant of *Escherichia coli* K-12: evidence for the role of a second protein in photorepair. *Photochem. Photobiol.* **58**:831–835.
9. Fleischmann, R. D., M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J. F. Tomb, B. A. Dougherty, J. M. Merrick, K. McKenney, G. Sutton, W. FitzHugh, C. A. Fields, J. D. Gocayne, J. D. Scott, R. Shirley, L. I. Liu, A. Glodeck, J. M. Kelley, J. F. Weidman, C. A. Phillips, T. Spriggs, E. Hedblom, M. D. Cotton, T. R. Utterback, M. C. Hanna, D. T. Nguyen, D. M. Saudek, R. C. Brandon, L. D. Fine, J. L. Fritchman, J. L. Fuhrman, N. S. M. Geoghagen, C. L. Gnehm, L. A. McDonald, K. V. Small, C. M. Fraser, H. O. Smith, and J. C. Venter. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**:496–512.
10. Fried, M., and D. M. Crothers. 1981. Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. *Nucleic Acids Res.* **9**:6505–6525.
11. Glaser, J. H., and J. A. DeMoss. 1971. Phenotypic restoration by molybdate of nitrate reductase activity in *chID* mutants of *Escherichia coli*. *J. Bacteriol.* **108**:854–860.
12. 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.
13. Hattori, M., and Y. Sakaki. 1986. Dideoxy sequencing method using denatured plasmid templates. *Anal. Biochem.* **152**:232–238.
14. Hemschemeier, S., M. Grund, B. Keuntje, and R. Eichenlaub. 1991. Isolation of *Escherichia coli* mutants defective in uptake of molybdate. *J. Bacteriol.* **173**:6499–6506.
15. Higgins, C. F. 1992. ABC transporters: from microorganisms to man. *Annu. Rev. Cell Biol.* **8**:67–113.
16. Hinton, S. M., and D. Dean. 1990. Biogenesis of molybdenum cofactors. *Crit. Rev. Microbiol.* **17**:169–188.
17. Hinton, S. M., C. Slaughter, W. Eisner, and T. Fisher. 1987. The molybdenum-pterin binding protein is encoded by a multigene family in *Clostridium pasteurianum*. *Gene* **54**:211–219.
18. Johann, S., and S. M. Hinton. 1987. Cloning and nucleotide sequence of the *chID* locus. *J. Bacteriol.* **169**:1911–1916.
19. Kalman, L. V., and R. P. Gunsalus. 1990. Nitrate- and molybdenum-independent signal transduction mutations in *narX* that alter regulation of anaerobic respiratory genes in *Escherichia coli*. *J. Bacteriol.* **172**:7049–7056.
20. Khavari, P. A., C. L. Peterson, J. W. Tamkun, D. B. Mendel, and G. R. Crabtree. 1993. BRG1 contains a conserved domain of the SW12/SNF2 family necessary for normal mitotic growth and transcription. *Nature (London)* **366**:170–174.
21. Lawson, C. L., and J. Carey. 1993. Tandem binding in crystals of a *trp* repressor/operator half-site complex. *Nature (London)* **366**:178–182.
22. Lee, J. H., J. C. Wendt, and K. T. Shanmugam. 1990. Identification and characterization of a new gene, *molR*, essential for utilization of molybdate by *Escherichia coli*. *J. Bacteriol.* **172**:2079–2087.
23. Luque, F., L. A. Mitchenall, M. Chapman, R. Christine, and R. N. Pau. 1993. Characterization of genes involved in molybdenum transport in *Azotobacter vinelandii*. *Mol. Microbiol.* **7**:447–459.
24. Maggi, A. M. A. 1991. GenBank accession number X61098.
25. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
26. Maupin-Furlow, J. A., J. K. Rosentel, J. H. Lee, U. Deppenmeier, R. P. Gunsalus, and K. T. Shanmugam. 1995. Genetic analysis of the *modABCD* (molybdate transport) operon of *Escherichia coli*. *J. Bacteriol.* **177**:4851–4856.
27. Miller, J. B., D. J. Scott, and N. K. Amy. 1987. Molybdenum-sensitive transcriptional regulation of the *chID* locus of *Escherichia coli*. *J. Bacteriol.* **169**:1853–1960.
28. Miller, J. H. 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
29. Mouncey, N. J., L. A. Mitchenall, and R. N. Pau. 1995. Mutational analysis of genes of the *mod* locus involved in molybdenum transport, homeostasis, and processing in *Azotobacter vinelandii*. *J. Bacteriol.* **177**:5294–5302.
30. Ogden, R. C., and D. A. Adams. 1987. Electrophoresis in agarose and acrylamide gels. *Methods Enzymol.* **152**:61–87.
31. Pabo, C. O., and R. T. Sauer. 1992. Transcription factors: structural families and principles of DNA recognition. *Annu. Rev. Biochem.* **61**:1053–1095.
32. Phillips, S. E. V., I. Manfield, I. Parsons, B. E. Davidson, J. B. Rafferty, W. S. Somers, D. Margarita, G. N. Cohen, I. Saint-Girons, and P. G. Stockley. 1989. Cooperative tandem binding of *met* repressor of *Escherichia coli*. *Nature (London)* **341**:711–715.
33. Phillips, S. E. V., and P. G. Stockley. 1994. Similarity of *met* and *trp* repressors. *Nature (London)* **368**:106.
34. Rajagopalan, K. V., and J. L. Johnson. 1992. The pterin molybdenum cofactors. *J. Biol. Chem.* **267**:10199–10202.
35. Rech, S., U. Deppenmeier, and R. C. Gunsalus. 1995. Regulation of the molybdate transport operon, *modABCD*, of *Escherichia coli* in response to molybdate availability. *J. Bacteriol.* **177**:1023–1029.
36. Rosentel, J. K., F. Healy, J. A. Maupin-Furlow, J. H. Lee, and K. T. Shanmugam. 1995. Molybdate and regulation of *mod* (molybdate transport), *fdhF*, and *hyc* (formate hydrogenlyase) operons in *Escherichia coli*. *J. Bacteriol.* **177**:4857–4864.
37. Rossini, G. P. 1994. The quaternary structures of untransformed steroid hormone receptors: an open issue. *J. Theor. Biol.* **166**:339–353.
38. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
39. Schlenso, V., A. Birkmann, and A. Böck. 1989. Mutations in *trans* which affect the anaerobic expression of formate dehydrogenase (*fdhF*) structural gene. *Arch. Microbiol.* **152**:83–89.
40. Schweizer, H. P. 1993. Two plasmids, X1918 and Z1918, for easy recovery of the *xylE* and *lacZ* reporter genes. *Gene* **134**:89–91.
41. Scott, D., and N. K. Amy. 1989. Molybdenum accumulation in *chID* mutants of *Escherichia coli*. *J. Bacteriol.* **171**:1284–1287.
42. Shanmugam, K. T., V. Stewart, R. P. Gunsalus, D. H. Boxer, J. A. Cole, M. Chippaux, J. A. DeMoss, G. Giordano, E. C. C. Lin, and K. V. Rajagopalan. 1992. Proposed nomenclature for the genes involved in molybdenum metabolism in *Escherichia coli* and *Salmonella typhimurium*. *Mol. Microbiol.* **6**:3452–3454.
43. Sperl, G. T., and J. A. DeMoss. 1975. *chID* gene function in molybdate activation of nitrate reductase. *J. Bacteriol.* **122**:1230–1238.
44. Studier, W. F., and B. A. Moffatt. 1986. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**:113–130.
45. Wang, G., S. Angermuller, and W. Klipp. 1993. Characterization of *Rhodobacter capsulatus* genes encoding a molybdenum transport system and putative molybdenum-pterin-binding proteins. *J. Bacteriol.* **175**:3031–3042.
46. Wootton, J. C., R. E. Nicolson, J. M. Cock, D. E. Walters, J. F. Burke, W. A. Doyle, and R. C. Cray. 1991. Enzymes depending on the pterin molybdenum cofactor: sequence families, spectroscopic properties of molybdenum and possible cofactor-binding domains. *Biochim. Biophys. Acta* **1057**:157–185.