JOURNAL OF BACTERIOLOGY, Apr. 1992, p. 2323–2331
0021-9193/92/072323-09$02.00/0
Copyright © 1992, American Society for Microbiology
Vol. 174, No. 7

Structural and Biochemical Characterization of the *Escherichia coli argE* Gene Product

THIERRY MEINNEL,* EMMANUELLE SCHMITT, YVES MECHULAM, AND SYLVAIN BLANQUET

Laboratoire de Biochimie, Unité de Recherche Associée no. 240 du Centre National de la Recherche Scientifique, Ecole Polytechnique, F-91128 Palaiseau Cedex, France

Received 19 September 1991/Accepted 27 January 1992

The DNA sequence of a 2,100-bp region containing the *argE* gene from *Escherichia coli* has been determined. The nucleotide sequence of the *ppc-argE* intergenic region was also solved and shown to contain six tandemly repeated REP sequences. Moreover, the *oxyR* gene has been mapped on the *E. coli* chromosome and shown to flank the *arg* operon. The codon responsible for the translation start of *argE* was determined by using site-directed mutants. This gene spans 1,400 bp and encodes a 42,350-Da polypeptide. The *argE3* allele and a widely used *argE* amber gene have also been cloned and sequenced. *N*-Acetylornithinase, the *argE* product, has been overproduced and purified to homogeneity. Its main biochemical and catalytic properties are described. Moreover, we demonstrate that the protein is composed of two identical subunits. Finally, the amino acid sequence of *N*-acetylornithinase is shown to display a high degree of identity with those of the succinyldiaminopimelate desuccinylase from *E. coli* and carboxypeptidase G2 from a *Pseudomonas* sp. It is proposed that this carboxypeptidase might be responsible for the acetylornithinase-related activity found in the *Pseudomonas* sp.

In microorganisms, the biosynthetic pathway from glutamate to arginine proceeds through eight enzymatic steps (reviewed in references 10, 11, and 15). Depending on the organism, distinct enzymatic strategies are used to ensure the removal of the acetyl group from N_2-acetylornithine (step 5 of the pathway). In *Escherichia coli* an *N*-acetylornithinase (NAO; EC 3.5.1.16) transforms this precursor into acetate and ornithine, whereas in *Streptococcus*, *Neurospora*, or *Chlamydomonas* spp. an ornithine acetyltransferase (EC 2.3.1.35) catalyzes the transfer of the acetyl group onto glutamate, yielding ornithine and acetylglutamate. In the latter case, acetylglutamate, the precursor of acetylornithine, is therefore regenerated upon the formation of ornithine (anaplerotic function). However, a few organisms contain the two activities. In *Saccharomyces* spp., a NAO activity was identified and proved not to be a side reaction of ornithine acetyltransferase but proposed to be due to a carboxypeptidase (12); in *Pseudomonas aeruginosa*, distinct polypeptides were found to be responsible for the two activities (16).

At the genetic level, NAO is encoded by the *argE* gene in *E. coli* (10) whereas ornithine acetyltransferase is encoded by the *ARG7* locus in yeasts (11). To our knowledge, the locus responsible for the NAO activity has not yet been identified in organisms expressing both *N*-acetylornithinase and ornithine acetyltransferase.

To characterize the NAO enzyme, it was of interest to determine its amino acid sequence. We report the nucleotide sequence of the *argE* gene from *E. coli* (part of the divergent *argECBH* operon), which had been previously mapped (8) and cloned (7). This allowed the construction of an NAO-overproducing strain and the purification of the enzyme to homogeneity. The NAO enzyme, which displays a rather broad specificity in the hydrolysis of numerous substrates of the type R_1-CO-NH-CH($(CH_2)_2$-R_2)-COOH (2, 20, 43), was further analyzed.

* Corresponding author.

MATERIALS AND METHODS

Bacterial strains and plasmids, general techniques. General genetic techniques were as described by Miller (28). The bacterial strains and plasmids used in this study are listed in Table 1.

Enzymes used were purchased from Boehringer-Mannheim (restriction enzymes, *Taq* DNA polymerase, DNA polymerase Klenow fragment), Pharmacia (restriction enzymes, T7 DNA polymerase, exonuclease III, T4 polynucleotide kinase), and Bethesda Research Laboratories (T4 DNA ligase). Buffers used were those indicated by the suppliers. Oligonucleotides were synthesized on a Gene Assembler I (Pharmacia) and purified by chromatography on a MonoQ anion exchanger (0.5 by 5 cm; Pharmacia). Chromosomal DNA was prepared as described previously (17) and purified by isopycnic CsCl gradient centrifugation (12 h in a Beckman Vti 65.2 rotor at 60,000 rpm). DNA fragments carrying the *argE* region were usually purified by size exclusion chromatography (36) and subcloned into M13 mp18, M13mp19, pBluescriptM13$^+$SK (Stratagene, La Jolla, Calif.) or pRS414 (37). The nucleotide sequence was determined by the dideoxy-chain termination method (34) with single-stranded or CsCl-purified double-stranded DNAs as the template. Automatic DNA sequencing was also performed by using the ALF system (Pharmacia).

Oligonucleotides were 5′ labeled by using 70 μCi of [γ-^{32}P] ATP (10 μCi/μl, 3,000 Ci/mmol; NEN) and 30 pmol of the desired oligonucleotide in a buffer consisting of 20 mM Tris-HCl (pH 8.0), 10 mM $MgCl_2$, and 20 mM dithiothreitol at 37°C for 30 min. Polymerase chain reaction amplification was used to clone *argE* with the following oligonucleotides: 5′-GCGGCTGAGTAGCAGGA and 5′-GATACTATCTA GACCAGAGGT. This latter oligonucleotide changes the *Bsp*HI restriction site upstream from ATG1 (see below) into a unique *Xba*I restriction site, while introducing an amber stop codon in place of an ATG codon.

DNA hybridization analysis was performed as follows. A 0.8% agarose gel run at 50 V for 12 h was denatured, neutralized (33), and dried under vacuum (30 min at room

TABLE 1. Bacterial strains and plasmid used in this study

Strain or plasmid	Genotype or markers	Source or reference
Strain		
K37	*galK rpsL*	27
UF121	*ara argE*(Am) Δ(*lac-proB*) *nalA rpoB thi* F'[*proA*$^+$ *B*$^+$ *lac*I-Z(Am)*181 kan*]	29
UF121R	Same as UF121 but *recA56 srl-300*::Tn*10*	38
PAL123	*ara* Δ(*lac-proB*) *nalA rpoB thi*	This work, derivative of UF121R
PAL3586	*thi-1 argE3 his-4* Δ*lacX74 galK2 mtl-1 xyl-5 tsx-29 supE44 rpsL*	25
PAL3587	Same as PAL3586 but *recA56 srl-300*::Tn*10*	25
Plasmids		
pMC7	pMB9 derivative carrying a 10-kbp *Eco*RI fragment containing the *argECBH* operon	8
pBSargE	pBluescript M13$^+$SK derivative carrying a *Stu*I-*Ecl*XI fragment from pMB9 between the *Eco*RV and *Ecl*XI sites, *argE*	This work
pBSargXE	pBluescript M13$^+$SK derivative carrying an *Eco*RV-*Xba*I fragment from the K37 chromosome between the *Hin*dII and *Xba*I sites, *argE*	This work
pBSargXEam	Same as pBSargXE, but cloned from the UF121 chromosome, *argE*(Am)	This work
pBSargXE3	Same as pBSargXE, but cloned from the PAL3586 chromsome, *argE3*	This work
pRSTT	pRS414 (37) derivative containing a *Pvu*II-*Sac*I fragment from pBluescriptM13$^+$KS between the *Sma*I and *Sac*I sites	This work
pRSargE	pRSTT derivative carrying a *Hpa*I-*Eco*RV fragment from pBSargE between the *Sma*I and *Hin*dII sites, *argE*	This work
pRSargEam1	Same as pRSargE but ATG1 codon changed to TAG	This work, Fig. 3
pRSargEop0	Same as pRSargE but ACA0 codon changed to TGA	This work, Fig. 3
pFam	pMC871 (5) derivative carrying a *Xho*I-*Bam*HI fragment from pBSTRNAfmetam (26) between the *Sal*I and *Bam*HI sites, expresses $tRNA_f^{Met}$ with CUA anticodon	This work

temperature followed by 45 min at 60°C). The dried gel was hybridized to about 50 × 10^6 dpm (30 pmol) of a 5' labeled oligonucleotide complementary to *oxyR* DNA (5'-AACTC GAATAACTAAAAGCCAA) at 37°C for 12 h in 50 ml of a buffer consisting of 0.5 M Na_2HPO_4 (pH 7.2), 2% sodium dodecyl sulfate (SDS), 1 mM EDTA, and 0.5% skimmed milk. The gel was then washed for 1 h in 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) (33)–0.1% SDS at room temperature and autoradiographed.

Site-directed mutagenesis. Oligonucleotide site-directed mutagenesis (35) was used to introduce mutations in the translation start region of the *argE* gene. For this purpose, the 789-bp *Hpa*I-*Hin*dIII fragment (bp 3084 to 3873 in Fig. 1) carrying *argE* plus the 5' extremity of *argC* and the promoter of the *arg* operon was cloned between the *Hin*dIII and *Hin*dII sites of the M13mp19 vector (M13HH1). The *Hin*dIII-*Eco*RI fragment of pRSargE (defined in Table 1) was then replaced by one of the mutagenized ones to yield either pRSargEop0 or pRSargEam1.

NAO assay. NAO was assayed by the method of Vogel and McLellan (44) in 250-μl standard assay mixtures containing 6 mM acetylornithine (Serva), 0.8 mM reduced gluthatione, 0.1 M K_2HPO_4 (pH 7.0), and 0.2 mM $CoCl_2$. Reactions were stopped by adding 0.75 ml of a solution prepared by mixing 2 volumes of ninhydrin (0.056 M in 2-methoxyethanol) and 1 volume of 0.4 M citric acid. After the mixture had been boiled for 10 min, 1.5 ml of 0.7 M NaOH was added. The resulting mixture was left for 20 min, and the A_{470} was measured. Under the standard assay conditions, molar extinction coefficients at 470 nm of 17,560 $M^{-1} \cdot cm^{-1}$ (ornithine) and 21.8 $M^{-1} \cdot cm^{-1}$ (acetylornithine) were measured and systematically used.

Protein concentrations were determined by the method of Bradford (4), using the Bio-Rad protein assay kit.

Purification of NAO. UF121R cells carrying the pBSargXE plasmid (Table 1) were used to inoculate a flask of 2× TY medium (1 liter) containing 50 μg ampicillin per ml. Cultures were grown overnight at 37°C, harvested by centrifugation, and resuspended in 85 ml of buffer A (0.1 M K_2HPO_4 [pH 7.0], 1 mM dithiothreitol). The sample was sonicated, and cell debris were removed by centrifugation. Streptomycin sulfate (3%, wt/vol) was added to the supernatant. After centrifugation, the supernatant was subjected to an $(NH_4)_2SO_4$ precipitation (70% saturation). The pellet was redissolved in 10 ml of buffer A and applied to a gel filtration column (Superose-6 [1.6 by 50 cm]; Pharmacia) equilibrated in the same buffer and eluted at 0.2 ml/min. Pooled active fractions (diluted to a final volume of 30 ml in buffer A) were applied to an anion exchanger (Q-Hiload [1.6 by 10 cm]; Pharmacia) equilibrated in buffer A. A 0 to 0.5 M KCl gradient (3 ml/min, 0.2 M/h) was used; the enzyme eluted at 0.24 M KCl. Fractions containing NAO activity were pooled, dialyzed, and stored at −30°C in buffer A containing 55% glycerol.

Protein samples (300 ng) were analyzed by SDS-polyacrylamide gel electrophoresis (PAGE) (12.5% homogeneous gels) or native PAGE (8 to 25% gradient gels) on the Phast-System (Pharmacia).

Nucleotide sequence accession number. The nucleotide sequence reported in this paper has been submitted to the GenBank/EMBL data bank as accession number X55417.

RESULTS

Genetic organization of the *argE* region. (i) Restriction analysis of the pMC7 plasmid. The pMC7 plasmid, obtained by inserting the 9.8 ± 0.3-kb *Eco*RI fragment of the λ50 bacteriophage into the pMB9 vector (8), was already known to carry the whole *argECBH* operon and suspected to carry a part of the *ppc* gene (7). It also contains a fragment (1.1 kb) of the λ genome.

To sequence the *argE* gene, we established a detailed

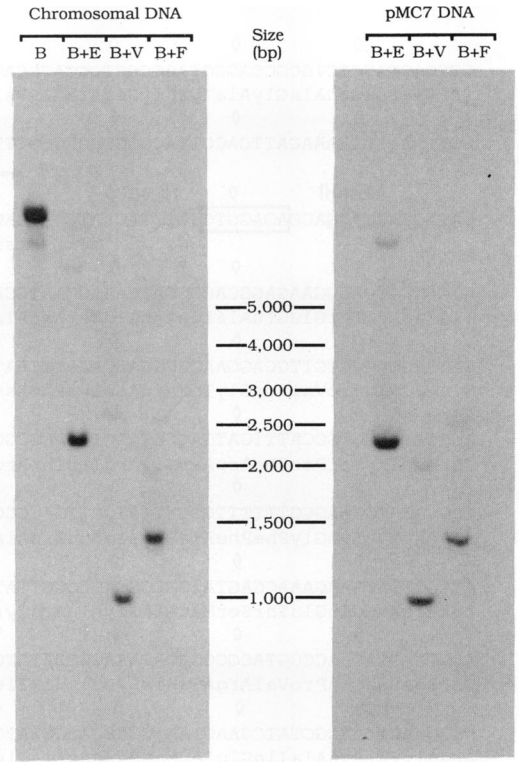

FIG. 2. *oxyR* is located on the pMC7 plasmid and flanks the *argH* gene. Samples (30 μg) of chromosomal DNA from strain K37 were restricted and analyzed by electrophoresis on an agarose gel (left). Samples (2 μg) of restricted pMC7 were applied to the same gel (right). Restriction enzymes: B, *Bam*HI; E, *Eco*RI; F, *Eco*RV; and V, *Pvu*II.

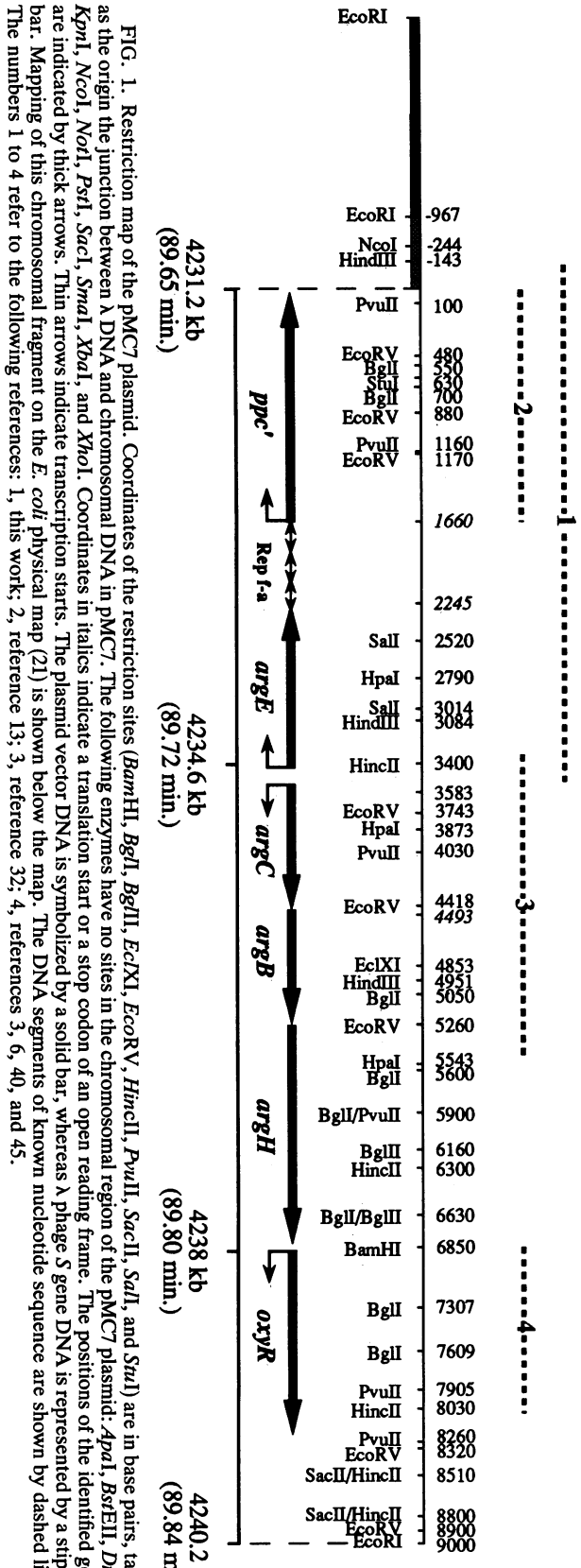

FIG. 1. Restriction map of the pMC7 plasmid. Coordinates of the restriction sites (*Bam*HI, *Bgl*I, *Bgl*II, *Ecl*XI, *Eco*RV, *Hin*cII, *Pvu*II, *Sac*II, *Sal*I, and *Stu*I) are in base pairs, taking as the origin the junction between λ DNA and chromosomal DNA in pMC7. The following enzymes have no sites in the chromosomal region of the pMC7 plasmid: *Apa*I, *Bst*EII, *Dra*II, *Kpn*I, *Nco*I, *Not*I, *Pst*I, *Sac*I, *Sma*I, *Xba*I, and *Xho*I. Coordinates in italics indicate a translation start or a stop codon of an open reading frame. The positions of the identified genes are indicated by thick arrows. Thin arrows indicate transcription starts. The plasmid vector DNA is symbolized by a solid bar, whereas λ phage *S* gene DNA is represented by a stippled bar. Mapping of this chromosomal fragment on the *E. coli* physical map (21) is shown below the map. The DNA segments of known nucleotide sequence are shown by dashed lines. The numbers 1 to 4 refer to the following references: 1, this work; 2, reference 13; 3, reference 32; 4, references 3, 6, 40, and 45.

restriction map of pMC7 (Fig. 1). The *Hin*cII pattern is different from the one already published (7). This could be due to a misinterpretation of the insertion point of the λ50 chromosomal fragment in the λ genome (see below). To subclone the *argE* gene, we inserted the *Stu*I (bp 630)-*Ecl*XI (bp 4853) fragment of pMC7 (Fig. 1) between the *Eco*RV and *Ecl*XI sites of pBluescriptM13$^+$SK. The pBSargE plasmid thus obtained was capable of relieving the arginine auxotrophy of the *argE recA* strain UF121R.

The restriction map in Fig. 1 was compared with the *E. coli* physical map (21). The *arg* operon could be located at 89.7 min and is likely to be carried fully by the λ 4G11 phage and partly by the λ 8H10 phage.

This mapping unambiguously established the presence on pMC7 of the 5′ part of *ppc* since the restriction pattern in the bp 0 to 1500 region (Fig. 1) matched that deduced from the *ppc* nucleotide sequence (13), with the exception of two additional *Bgl*I sites close to each other (positions 550 and 700). Accordingly, one of these *Bgl*I sites has been reported in the Kohara map (21).

(ii) The *oxyR* gene is located immediately downstream from the *argH* gene on the pMC7 plasmid. Extracts of cells infected with a λ transducing phage carrying the *argH* gene showed the overexpression of a polypeptide of M_r 55,000 (24), indicating, therefore, that the *argH* gene is roughly 1,500 bases long. From this and from the DNA sequences of the *argECBH* promoter (9), the *argC* and *argB* genes, and the 5′ part of the *argH* gene (32), we concluded that the *argH*

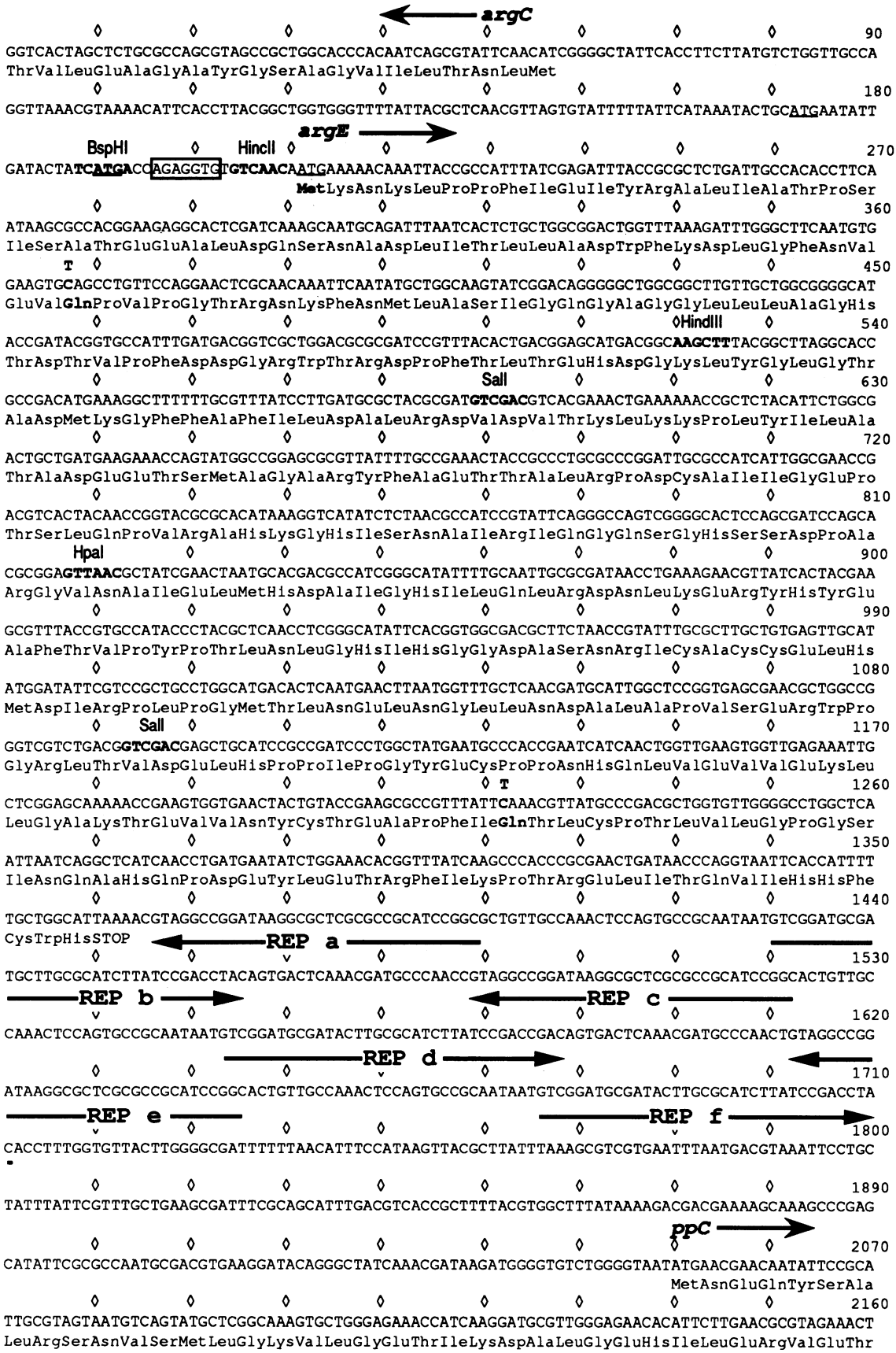
argC
argE
BspHI
HincII
HindIII
SalI
HpaI
SalI
REP a
REP b
REP c
REP d
REP e
REP f
ppC

FIG. 3. Nucleotide sequence of the *argE* gene. The DNA sequence and translated amino acid sequence are shown. The three possible start codons of *argE* are underlined. The third one is responsible for the translation start (see Results). The Shine-Dalgarno consensus sequence is boxed. Restriction sites are shown in boldface type. The REP sequences (numbered from a to f) are indicated with arrows. The single transitions found in each of the *argE*(Am) (codon 53) and *argE3* (codon 338) genes are shown in boldface type above the nucleotide sequence. The beginnings of the flanking genes (*argC* and *ppc*) are also indicated.

gene should extend nearly to the unique *Bam*HI site on the pMC7 plasmid (Fig. 1).

A recent map of the *E. coli* chromosome (22) suggested that the *oxyR* gene was located near the *arg* operon at 89.4 min; moreover, *oxyR* has recently been cloned and sequenced from the λ 4G11 phage (40). Since the deduced *oxyR* restriction map perfectly matches that of the pMC7 plasmid from *Bam*HI (bp 6850) to *Eco*RV (bp 8320), it has been proposed to map close to the *arg* operon (3, 45). To confirm this localization of the *oxyR* gene, we performed the following hybridization analysis. An oligonucleotide, whose sequence was derived from the *oxyR* DNA sequence (3, 6, 40, 45), was hybridized to either restricted pMC7 plasmid DNA or restricted chromosomal DNA. As shown in Fig. 2, the hybridization patterns confirmed that *oxyR* was indeed located downstream from the *argH* gene on the pMC7 plasmid (Fig. 1). The next gene already mapped close to *oxyR* is *trmA*. It lies 2 to 3 kb downstream from *oxyR* on the next *Eco*RI fragment, as shown by the restriction map of pTN035, a plasmid encoding both the *argH* and *trmA* genes (31). *trmA* is transcribed counterclockwise (23). This set of data allows the determination of the entire gene linkage from *ppc* (89.4 min) to *rpoB* (90.4 min), with the exception of the above 2- to 3-kb *Eco*RI region. Since the gene encoding pantothenate kinase (*coaA*) has been mapped between *argH* and *rpoB* (41), it is tempting to propose that *coaA* could occupy this yet unsolved region. If this is true, this gene could easily be cloned from λ 4G11 and/or E11C11 of the *E. coli* λ library of Kohara et al. (21) or from plasmid pTN035 (31).

Nucleotide sequence of the *argE* gene region. The nucleotide sequence of the *Hin*dIII (bp −143)-*Hin*cII (bp 3400) fragment containing the whole *argE* gene was determined (Fig. 3). Two open reading frames oriented counterclockwise and starting at positions 1359 (396 codons) and 2070 (553 codons) were found. The sequence of the latter open reading frame, which does not end within the sequenced fragment, matches the 5′ part of the *ppc* gene (Fig. 1). The 396-codon open reading frame most probably encodes the NAO. This conclusion is established further below by the sequencing of *argE* mutants.

```
REP
 a        ••------C•GC•C••••••••GC••••A••GC••••••••G•••••
 c,e      ••••A•••CGG••C••••••••GC••••A••GC••••••••G•••••
 b        ••••••••T•A••T••••••••AT••••••••AT••••••••A•••••
 d        ••••••••T•A••T••••••••AT••••••••AT••••••••A••G••
 f        ••••••••T•A••T••••••••ATA•••••••AT••••••••A•••••

consensus TGCCGCAAYARTGYCGGATGCGRYGCTTGCGCRYCTTATCCGRCCTAC

canonical          GCCKGATGCGRCGYNNNNRCGYCTTATCMGGCCTAC
```

FIG. 4. Extended REP sequences in the *argE-ppC* intergenic region. The consensus sequence generated from the six REP units in the *argE-ppc* intergenic region is shown. For each REP unit, only the bases different from the consensus sequence are listed, while bases identical to the consensus sequence are indicated by dots. A dash signifies the deletion of a base from the consensus sequence. The canonical REP sequence (39) is also reported. K = G or T; M = C or A; R = G or A; Y = C or T.

Nucleotide sequencing of the intergenic *argE-ppc* region revealed the occurrence of a triplicated region. Each repeated unit was composed of a 57- to 61-bp sequence with dyad symmetry (Fig. 3). From these three homologous units, a consensus sequence could be defined (Fig. 4). It matched the well-known canonical REP (repetitive extragenic palindromic, also called PU [palindromic unit]) sequence, which is found in the intergenic regions of several operons (18, 39). The consensus sequence obtained here is 12 bp longer than the canonical REP sequence (Fig. 3).

Finally, examination of the nucleotide sequence from the *Hin*dIII site (bp −143 in Fig. 1) to the junction with *ppc* confirmed that the chromosomal insertion within λ50 occurred in the *S* gene region and not in the pR one, as suggested earlier (7). Linkage between *ppc* and the λ part of the *Eco*RI insert occurred roughly 100 bases downstream from the second *Pvu*II site (bp 100 in Fig. 1) of *ppc*.

Determination of the translation start of *argE*. The *argE* gene sequence revealed three possible initiator ATG codons (Fig. 3), of which the third (underlined in Fig. 3) is the best candidate because of the presence upstream of a well-located typical Shine-Dalgarno sequence (boxed in Fig. 3). To prove that this codon (referred to as ATG1) was indeed the initiation codon of *argE*, we carried out the following experiment. The codon preceding ATG1 was mutagenized into an opal (TGA) one. The NAO activity expressed from pRSargEop0, the plasmid carrying this opal mutation, was identical to that expressed from the control pRSargE plasmid (Table 2). This result strongly suggested that ATG1 was the initiator codon of *argE*. Furthermore, ATG1 was changed into an amber codon (TAG) to give plasmid pRS argEam1. As expected, the resulting plasmid was unable to relieve the arginine auxotrophy of the Su^- strain UF121R, whereas it conferred an Arg^+ phenotype on strain PAL3587 (*argE3 supE44*). In the latter case, the NAO activity in cell extracts did not exceed 1% of that obtained with pRSargE. This could be explained by a low efficiency of translation

TABLE 2. Determination of the translation start of *argE*[a]

Strain	pRSargE derivative	Phenotype[b]	NAO relative activity (%)[c]
UF121R	pRSargE	Arg^+	72
UF121R	pRSargEop0	Arg^+	100
UF121R	pRSargEam1	Arg^-	0
PAL3587	pRSargEam1	Arg^+	0.8
PAL3587(pFam)	pRSargEam1	Arg^+	63
PAL3587	pRSargE	Arg^+	83

[a] The indicated strains were transformed by the indicated plasmid.

[b] The Arg^+ phenotype corresponded to the ability to form colonies on minimal M9 plates without arginine after 48 h of incubation at 37°C.

[c] Cells were grown in LB medium (26) supplemented with 50 μg of ampicillin per ml. Samples (1 ml) were withdrawn during exponential growth (optical density at 650 nm of ca. 1). After centrifugation, the cells were resuspended in buffer A and sonicated. NAO activity was assayed and normalized to the A_{280} of the extract. The NAO activity is indicated as the percentage of the activity measured with UF121R carrying pRSargEop0 (0.13 U/A_{280} unit). One enzyme unit (U) is defined as the amount of enzyme capable of hydrolyzing 1 μmol of NAO per s in the standard assay.

initiation upstream from ATG1 (Table 2). It has already been reported that a plasmid expressing a $tRNA_f^{Met}$ derivative with an amber anticodon (CUA) allowed translation initiation of the *cat* gene carrying an amber codon as the translation initiator (42). Similarly, if ATG1 was indeed the initiator codon, a large increase of the NAO activity should occur upon introduction of a similar plasmid (pFam in Table 1) in the strain PAL3587(pRSargEam1). As expected, expression of this amber $tRNA_f^{Met}$ caused an increase (80-fold) in the NAO activity, which became identical to that obtained with pRSargE (Table 2). That this amber mutation at codon 1 was much better suppressed by an initiator suppressor tRNA than by an elongator one confirmed that ATG1 was responsible for the translation initiation of the *argE* gene.

Cloning and sequencing of the *argE3* (Oc) and *argE*(Am) mutant genes. Many *E. coli* strains of genetic interest carry a mutation mapped in the *argE* gene and conferring an Arg^- phenotype. Two alleles have been widely used. The *argE3* allele, found in strains derived from AB1157 (1), is known to carry an ochre mutation (19). Another *argE* mutation, carried by the XA100 derivatives (29), can be suppressed by amber suppressors. These two mutations could be mapped by comparing the capacity of a set of DNA fragments covering the *argE* open reading frame to cure the *argE*(Am) or *argE3* mutations of the UF121 and PAL3586 strains. Four M13 derivatives carrying either the *Hpa*I (bp 3873)-*Hin*dIII (bp 3084) (M13HH1), the *Hin*dIII (bp 3084)-*Hpa*I (bp 2790) (M13HH2), the *Hpa*I (bp 2790)-*Stu*I (bp 630) (M13HS), or the *Sal*I (bp 2520)-*Eco*RI (bp 1170) (M13SE) fragments were used to infect the recombination-proficient *argE* strains UF121 and PAL3586. Infected cells were scored for arginine prototrophy on M9 minimal medium. For the UF121 strain, only the cells infected with M13HH1 yielded Arg^+ cells at high frequency (10^{-4}) with respect to the reversion frequency ($<10^{-7}$). The *argE*(Am) mutation could therefore be located in the region containing the first 100 codons of *argE*. With strain PAL3586, a high efficiency of curing occurred by using the M13HS and M13SE phages, thus indicating that the *argE3* mutation lies within the distal 100 codons of *argE*.

To precisely localize the mutations, we cloned the mutated *argE*(Am) and *argE3* genes from the UF121 and PAL3586 chromosomes, respectively, by using the polymerase chain reaction technique. As a control, the wild-type allele was also cloned from the Arg^+ strain K37 by the same technique. The amplified genes were cloned in pBSM13$^+$SK under the control of the *lac* inducible promoter, to yield pBSargXE (fragment from K37), pBSargXEam (fragment from UF121), and pBSargXE3 (fragment from PAL3586). The binding sites for the *argR* repressor (15) are not present on the cloned fragments. It was verified that pBSargXE relieved the Arg auxotrophy of the UF121R strain (Su^-) whereas pBSargXEam and pBSargXE3 did not. In each of the three experiments, two independently obtained plasmids were fully sequenced. The sequence of the *argE* gene from K37 was identical to that obtained from the pMC7 insert. The *argE* sequence from pBSargXEam carried a single mutation with codon Gln-53 (CAG) changed into a TAG amber codon (Fig. 3). The *argE3* allele also showed a point transition with the Gln-338 CAA codon changed into a TAA ochre codon (Fig. 3). The latter mutation creates a *Dra*I restriction site.

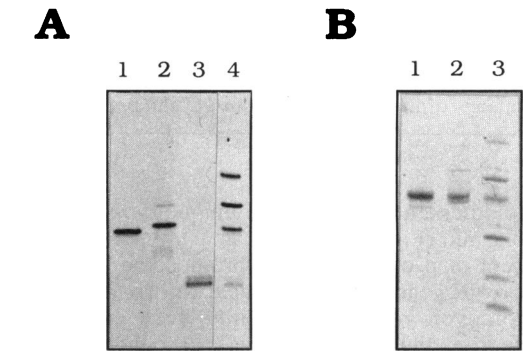

FIG. 5. NAO is a homodimer. (A) Samples containing 300 ng of either bovine serum albumin (67 kDa; lane 1), purified NAO (lane 2), or ovalbumin (43 kDa; lane 3) were analyzed by PAGE on an 8 to 25% gradient gel. Lane 4 was loaded with a mixture of molecular mass markers (152 kDa, *E. coli* methionyl-tRNA synthetase; 108 kDa, *E. coli* valyl-tRNA synthetase; 67 kDa, bovine serum albumin; 43 kDa, ovalbumin). (B) Samples containing 300 ng of ovalbumin (43 kDa; lane 1) or purified NAO (lane 2) were analyzed by SDS-PAGE. Lane 3 was loaded with the low-molecular-mass markers from Pharmacia (94, 67, 43, 30, 20, and 14.4 kDa).

Purification and biochemical characterization of the *E. coli* NAO. The predicted molecular weight of the *argE* product (M_r 42,350) is smaller than the one previously measured by either molecular sieving (M_r 62,000 [personal communication in reference 15]) or analysis by denaturing SDS-PAGE (M_r 52,000 [24]). To investigate this discrepancy, we purified NAO to homogeneity from the UF121R(pBSargXE) strain, which overexpresses the NAO activity 600-fold compared with the PAL123 strain grown under identical conditions. The purification procedure is described in Materials and Methods section and is summarized in Table 3. SDS-PAGE analysis of pure NAO (Fig. 5B) showed that the enzyme behaves as an M_r 42,000 peptide, in good agreement with the

TABLE 3. Purification of NAO

Purification step	Amt of protein (mg)[a]	Total activity (U)[b]	Sp act (U/mg of protein)	Yield (%)	Relative purification
Crude extract	520	785	1.5	100	
Ammonium sulfate	385	645	1.7	82	1.1
Superose-6	108	570	5.3	73	3.5
Q-Hiload	4.5	392	87.1	52	58

[a] Assayed as described previously (4).
[b] One enzyme unit (U) is defined as the amount of enzyme capable of hydrolyzing 1 μmol of NAO per s in the standard assay.

TABLE 4. Biochemical and catalytic properties of NAO

Quaternary structure	Biochemical property			Catalytic property	
	M_r/promoter	pI	ϵ_M at 280 nm (M^{-1} cm^{-1})	k_{cat} (s^{-1})	K_m (mM)
α_2	42,350[a]	6.14[a]	101,300[b]	9,400 ± 600	1.56 ± 0.32

[a] Computed from the amino acid sequence.
[b] The molar absorption coefficient at 280 mm was deduced from the measurements of the A_{280} of protein samples and of the protein concentration in the same samples by the Bradford technique (4) with bovine serum albumin as a standard.
[c] NAO michaelian parameters were derived from iterative nonlinear least squares fits of the Michaelis-Menten equation to the experimental values. Confidence limits on the fitted values were determined by 30 Monte Carlo stimulations, using the experimental standard deviations on individual measurements. The K_m value is that of acetylornithine as deduced from initial reaction rates in the standard assay containing 1 to 40 mM acetylornithine.

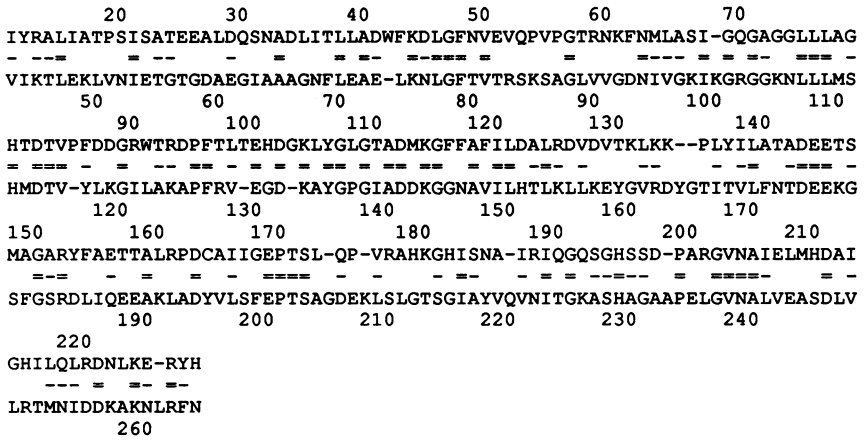

FIG. 6. Alignment of the amino acid sequences of *E. coli* NAO and *Pseudomonas* CPG2. Aligned are positions 11 to 228 from NAO (top) and 43 to 264 from CPG2 (bottom). Symbols: =, strictly conserved, −, conservative replacement.

value calculated from the DNA sequence (M_r 42,350). PAGE analysis under nondenaturing conditions (Fig. 5A) indicated an apparent molecular weight of 77,000, suggesting that the protein was a homodimer. The latter conclusion was confirmed by the elution profile of NAO on a Superose-6 molecular sieving column (data not shown). The above properties, as well as the michaelian parameters of the reaction catalyzed by the purified NAO, are summarized in Table 4. Note the very high k_{cat} value of the reaction catalyzed by purified NAO (Table 4).

NAO is homologous to CPG2 and SDPD. An extensive search for homologies (using the NBRF protein library, release 27) enabled us to first align the *E. coli* NAO sequence with that of a protein originating from *Pseudomonas* sp. strain RS-16 (30) and previously reported to be a carboxypeptidase G2 (CPG2). In the comparison, 218 amino acids, corresponding to about two-thirds of the total length of either polypeptide, could be aligned (amino acids 11 to 228 in the NAO sequence and amino acids 43 to 264 in the CPG2 sequence). According to the alignment shown in Fig. 6, the N termini are shifted by 32 residues, in agreement with the length of the signal peptide of CPG2. Of the 218 aligned residues, 64 identities (29.4% identity) and 47 conservative replacements (50.9% similarity) were found. In addition, hydrophobic cluster analysis (14) of the sequences of the two proteins suggested similar secondary structures within the conserved regions (data not shown).

Finally, succinyldiaminopimelate desuccinylase (SDPD; EC 3.5.1.18) also shared significant homology with NAO between amino acids 79 and 117 (60.5% identity). The DNA sequence of the *E. coli dapE* gene, encoding SDPD, is not yet published, but is available from the EMBL data bank with accession number X57403 (submitted by P. Stragier, Institut de Biologie Physico-Chimique, Paris, France).

DISCUSSION

This study establishes the nucleotide sequence of the *argE* gene from *E. coli*. In vivo mapping of the translation start site shows that this gene encodes an M_r 42,350 polypeptide. The expressed NAO enzyme behaves as a homodimer under its native form in vitro. Moreover, NAO is partly identical to CPG2, a protein originating from a *Pseudomonas* sp., and SDPD from *E. coli*.

Interestingly, the enzymes CPG2 and SDPD are known to hydrolyze peptidic bonds of substrates possessing the same consensus structure as the one recognized and cleaved by NAO, i.e., R_1-CO-NH-CH$(CH_2)_2$-R_2)-COOH, and, in the case of CPG2, to also have a broad substrate specificity (30). Moreover, the organism (*Pseudomonas* sp.) from which CPG2 has been extracted contains an NAO-related activity. This activity behaves as a Zn^{2+}-dependent carboxypeptidase (12). Since CPG2 is also described as a zinc-dependent enzyme, with an M_r of 2 × 43,932, it is tempting to propose that CPG2 itself corresponds to the NAO activity previously identified in *Pseudomonas* extracts.

This conclusion may be of particular interest for the understanding of the structure-function relationships of NAO and CPG2. The two enzymatic activities are characterized by the ability to hydrolyze an amino acid with a relatively broad specificity. The conserved residues shown in the comparison in Fig. 6, particularly an 18-amino-acid peptide (positions 106 to 123 in the NAO sequence), also strongly conserved in SDPD, may therefore have a crucial role in the hydrolytic activity.

Finally, we wondered whether, reciprocally, *E. coli* NAO could express such a carboxypeptidase activity. However, when using 6 mM *N*-α-benzoylornithine as the substrate, we could not show ornithine synthesis in a standard assay containing up to 1 μM homogeneous NAO. This result may reflect a hydrolytic specificity of the *E. coli* NAO narrower than that of the CPG2 enzyme.

ACKNOWLEDGMENTS

We thank N. Glansdorff for the generous gift of plasmid pMC7 and helpful comments on our manuscript. F. Dardel and P. Plateau are acknowledged for critical reading of the manuscript.

This work was supported by grants from the Ministère de la Recherche et de l'Enseignement supérieur (87-CO392) and the Fondation pour la Recherche Médicale.

REFERENCES

1. **Bachmann, B. J.** 1987. Derivations and genotypes of some mutant derivatives of *Escherichia coli*, p. 1190–1219. *In* F. C.

Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.

2. **Baumberg, S.** 1970. Acetylhistidine as substrate for acetylornithinase: a new system for the selection of arginine regulation mutants in *Escherichia coli*. Mol. Gen. Genet. **106:**162–173.
3. **Bölker, M., and R. Kahlman.** 1989. The *Escherichia coli* regulatory protein OxyR discriminates between methylated and unmethylated states of the phage Mu *mom* promoter. EMBO J. **8:**2403–2410.
4. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72:**248–254.
5. **Casadaban, M. J., J. Chou, and S. N. Cohen.** 1980. In vitro gene fusions that join an enzymatically active β-galactosidase segment to amino-terminal fragments of exogenous proteins: *Escherichia coli* plasmid vectors for the detection and cloning of translational initiation signals. J. Bacteriol. **143:**971–980.
6. **Christman, M. F., G. Storz, and B. N. Ames.** 1989. *OxyR*, a positive regulator of hydrogen peroxide-inducible genes in *Escherichia coli* and *Salmonella typhimurium*, is homologous to a family of bacterial regulatory proteins. Proc. Natl. Acad. Sci. USA **86:**3484–3488.
7. **Crabeel, M., D. Charlier, R. Cunin, and N. Glansdorff.** 1979. Cloning and endonuclease restriction analysis of *argF* and of the control region of the argECBH bipolar operon in *Escherichia coli*. Gene **5:**207–231.
8. **Crabeel, M., D. Charlier, N. Glansdorff, S. Palchaudhuri, and W. K. Maas.** 1977. Studies on the bipolar *argECBH* operon of *E. coli*: characterization of restriction endonuclease fragments obtained from λ*dargECBH* transducing phages and a colE1 *argECBH* plasmid. Mol. Gen. Genet. **151:**161–168.
9. **Cunin, R., A. Boyen, J. Piette, M. Crabeel, and N. Glansdorff.** 1982. Superposition of genetic sites in the regulatory region of the bipolar *argECBH* operon of *Escherichia coli*. Ann. Microbiol. (Paris) **133A:**235–241.
10. **Cunin, R., N. Glansdorff, A. Piérard, and V. Stalon.** 1986. Biosynthesis and metabolism of arginine in bacteria. Microbiol. Rev. **50:**314–352.
11. **Davis, R. H.** 1986. Compartmental and regulatory mechanisms in the arginine pathways of *Neurospora crassa* and *Saccharomyces cerevisiae*. Microbiol. Rev. **50:**280–313.
12. **Degryse, E.** 1974. Evidence that yeast acetylornithinase is a carboxypeptidase. FEBS Lett. **43:**285–288.
13. **Fujita, N., T. Miwa, S. Ishijima, K. Izui, and H. Katuki.** 1984. The primary structure of phosphoenolpyruvate carboxylase of *Escherichia coli*. Nucleotide sequence of the *ppc* gene and deduced aminoacid sequence. J. Biochem. **95:**909–916.
14. **Gaboriaud, C., V. Bissery, T. Benchetrit, and J. P. Mornon.** 1987. Hydrophobic cluster analysis: an efficient way to compare and analyze amino acid sequences. FEBS Lett. **224:**149–155.
15. **Glansdorff, N.** 1987. Biosynthesis of arginine and polyamines, p. 321–344. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
16. **Haas, D., and B. W. Holloway.** 1977. The genetic organization of arginine biosynthesis in *Pseudomonas aeruginosa*. Mol. Gen. Genet. **154:**7–22.
17. **Harris-Warrick, R. M., Y. Elkana, S. D. Ehrlich, and L. Lederberg.** 1976. Electrophoretic separation of *Bacillus subtilis* genes. Proc. Natl. Acad. Sci. USA **72:**2207–2211.
18. **Higgins, C. F., G. F. L. Ames, W. M. Barnes, J. M. Clement, and M. Hofnung.** 1982. A novel intercistronic regulatory element of prokaryotic operons. Nature (London) **298:**760–762.
19. **Kato, T., and S. Shinoura.** 1977. Isolation and characterization of mutants of *Escherichia coli* deficient in induction of mutations by ultraviolet light. Mol. Gen. Genet. **156:**121–131.
20. **Kelker, N. E., and W. K. Maas.** 1974. Selection for genetically repressible (Arg^+) strains of *Escherichia coli* K12 from genetically derepressed ($ArgR^-$) mutants using acetylnorvaline. Mol. Gen. Genet. **132:**131–136.
21. **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.
22. **Kröger, M., R. Wahl, and P. Rice.** 1990. Compilation of DNA sequences of *Escherichia coli* (update 1990). Nucleic Acids Res. **18:**2549–2552.
23. **Lindström, P. H. R., D. Stüber, and G. R. Björk.** 1985. Genetic organization and transcription from the gene (*trmA*) responsible for synthesis of tRNA (uracil-5)-methyltransferase by *Escherichia coli*. J. Bacteriol. **164:**1117–1123.
24. **Linn, T., M. Gomant, and J. Scaife.** 1979. Lambda transducing bacteriophage carrying deletions of the *argECBH-rpoBC* region of the *Escherichia coli* chromosome. J. Bacteriol. **140:**479–489.
25. **Mechulam, Y., S. Blanquet, and G. Fayat.** 1987. Dual-level control of the *E. coli pheST-himA* operon expression: $tRNA^{Phe}$-dependent attenuation and transcriptional operator repressor control. J. Mol. Biol. **197:**453–470.
26. **Meinnel, T., Y. Mechulam, D. Le Corre, M. Panvert, S. Blanquet, and G. Fayat.** 1991. Selection of suppressor methionyl-tRNA synthetase: mapping the anticodon binding site. Proc. Natl. Acad. Sci. USA **88:**291–295.
27. **Miller, H. I., and D. I. Friedman.** 1980. An *E. coli* gene product required for λ site-specific recombination. Cell **20:**711–719.
28. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
29. **Miller, J. H., and A. M. Albertini.** 1983. Effects of surrounding sequence on the suppression of nonsense codons. J. Mol. Biol. **164:**59–71.
30. **Minton, N. P., T. Atkinson, C. R. Bruton, and R. F. Sherwood.** 1984. The complete sequence of the *Pseudomonas* gene coding for carboxypeptidase G2. Gene **31:**31–38.
31. **Ny, T., and G. R. Björk.** 1980. Cloning and restriction mapping of the *trmA* gene coding for transfer ribonucleic acid (5-methyluridine)methyltransferase in *Escherichia coli* K-12. J. Bacteriol. **142:**371–379.
32. **Parsot, C., A. Boyen, G. N. Cohen, and N. Glansdorff.** 1988. Nucleotide sequence of *Escherichia coli argB* and *argC* genes: comparison of N-acetylglutamatekinase and N-acetylglutamate-γ-semialdehyde dehydrogenase with homologous and analogous enzymes. Gene **68:**275–283.
33. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
34. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:**5463–5467.
35. **Sayers, J. R., W. Schmidt, and F. Eckstein.** 1988. 5′→3′ exonucleases in phosphorothiate-based oligonucleotide-directed mutagenesis. Nucleic Acids Res. **16:**791–802.
36. **Schmitter, J. M., Y. Mechulam, G. Fayat, and M. Anselme.** 1986. Rapid purification of DNA fragments by high-performance size-exclusion chromatography. J. Chromatog. **378:**462–466.
37. **Simons, R. W., F. Houman, and N. Kleckner.** 1987. Improved single and multicopy *lac*-based cloning vectors for protein and operon fusions. Gene **53:**85–96.
38. **Springer, M., M. Graffe, J. Dondon, and M. Grünberg-Manago.** 1989. tRNA-like structures and gene regulation at the translational level: a case of molecular mimicry in *Escherichia coli*. EMBO J. **8:**2417–2424.
39. **Stern, M. J., G. Ferro-Luzzi Ames, N. H. Smith, E. C. Robinson, and C. F. Higgins.** 1984. Repetitive extragenic palindromic sequences: a major component of the bacterial genome. Cell **37:**1015–1026.
40. **Tao, K., K. Makino, S. Yonei, A. Nakata, and H. Shinagawa.** 1989. Molecular cloning and nucleotide sequencing of *oxyR*, the positive regulatory gene of a regulon for an adaptive response to oxidative stress in *Escherichia coli*: homologies between *oxyR* protein and a family of bacterial activator proteins. Mol. Gen. Genet. **218:**371–376.
41. **Vallari, D. S., and C. O. Rock.** 1987. Isolation and characteri-

zation of temperature-sensitive pantothenate kinase (*coa*) mutants of *Escherichia coli*. J. Bacteriol. **169:**5795–5800.
42. **Varshney, U., and U. L. RajBhandary.** 1990. Initiation of protein synthesis from a termination codon. Proc. Natl. Acad. Sci. USA **87:**1586–1590.
43. **Vogel, H. J., and D. M. Bonner.** 1956. Acetylornithinase of *Escherichia coli*: partial purification and some properties. J. Biol. Chem. **218:**97–106.
44. **Vogel, H. J., and W. L. McLellan.** 1971. Acetylornithinase (*Escherichia coli*). Methods Enzymol. **17A:**265–269.
45. **Warne, S. R., J. M. Varley, G. J. Boulnois, and M. G. Norton.** 1990. Identification and characterization of a gene that controls colony morphology and auto-aggregation in *Escherichia coli* K-12. J. Gen. Microbiol. **136:**455–462.