Structure and Regulation of the AMP Nucleosidase Gene (*amn*) from *Escherichia coli*†

Hazel B. Leung,‡ Kalla L. Kvalnes-Krick,§ Sheryl L. Meyer,∥ J. Kim deRiel,⊥ and Vern L. Schramm*,§

Department of Biochemistry and Fels Research Institute, Temple University School of Medicine, Philadelphia, Pennsylvania 19140, and Department of Biochemistry, Albert Einstein College of Medicine of Yeshiva University, 1300 Morris Park Avenue, Bronx, New York 10461

Received April 5, 1989; Revised Manuscript Received June 26, 1989

ABSTRACT: The gene for AMP nucleosidase from *Escherichia coli* (*amn*) has been sequenced and characterized. The gene codes for a transcript of 1.7 ± 0.2 kb, and the open reading frame corresponds to a protein of 483 amino acids (M_r = 53 848). Amino acid sequences from tryptic peptides of AMP nucleosidase, N-terminal amino acid analysis, and the amino acid composition confirm the gene assignment and the open reading frame of *amn*. Primer extension studies determined the 5′-end of the *amn* transcript. The 5′-regulatory region contains overlapping sequences with similarity to the consensus sequences for binding cAMP receptor protein and inorganic phosphate repressor protein. Addition of exogenous cAMP to *E. coli* deficient in adenylate cyclase resulted in a 3-fold increase in AMP nucleosidase activity. Growth of *E. coli* on limiting phosphate resulted in an 8-fold increase in the production of AMP nucleosidase. The *amn* gene was expressed in AMP nucleosidase deficient strains of *Azotobacter vinelandii* and *E. coli*. A pUC·*amn* construct is described that causes approximately 20% of the total protein in *E. coli* to be produced as AMP nucleosidase. Comparison of the amino acid sequence for AMP nucleosidase with that for yeast AMP deaminase (see the following paper) indicates a region in which six of eight amino acids are identical but no other overall homology. The amino acid sequence showed poor agreement with consensus sequences for adenylate binding sites even though the enzyme is known to have a catalytic site for AMP and regulatory sites for MgATP and phosphate.

AMP nucleosidase (EC 3.2.2.4) from *Escherichia coli* catalyzes the hydrolysis of the N-glycosidic bond of AMP to produce adenine and ribose 5-phosphate. The enzyme is allosterically activated by MgATP which decreases the $S_{0.5}$ for AMP several hundredfold but has no effect on V_{max} (Leung & Schramm, 1980). Inorganic phosphate is an allosteric inhibitor, reversing the effects of MgATP. The gene for AMP nucleosidase has been cloned into pBR322 and used to overproduce the protein (Leung & Schramm, 1984). The kinetic properties of the *E. coli* AMP nucleosidase have been defined (Leung & Schramm, 1980), heavy-atom kinetic isotope effects have been used to characterize the transition-state structure (Parkin & Schramm, 1984), and studies of the X-ray crystal structure are in progress (Giranda et al., 1986). The DNA sequence and deduced amino acid sequence of *amn* were determined to aid in the crystal structure analysis and to identify the catalytic and regulatory sites by comparison with known binding sites for adenylate nucleotides and phosphate (Moller & Amons, 1985; Walker et al., 1982). The regulation of expression of *amn* is also of interest since AMP nucleosidase is a minor protein (0.01%) of *E. coli*. The expression of the gene is shown here to be responsive to cAMP and phosphate levels, compounds for which regulatory DNA sequences have been described (deCrombrugghe et al., 1984; Makino et al., 1986).

Regulation of AMP concentrations is thought to be governed by AMP nucleosidase in prokaryotes and AMP deaminase in eukaryotes [e.g., Leung and Schramm (1980, 1984)]. Both enzymes are activated by MgATP and inhibited by inorganic phosphate. No prokaryotes are known to contain AMP deaminase, and no eukaryotes are known to contain AMP nucleosidase. The similarity of function, regulation, and phylogenetic distribution of the enzymes suggested a common evolutionary precursor. This hypothesis is tested in this paper and in the following paper which describes *AMD*, the AMP deaminase gene from yeast (Meyer et al., 1989). In the third paper of this series, it is demonstrated that yeast and mammalian AMP deaminases share regions of sequence similarity (Marquetant et al., 1989).

Experimental Procedures

Materials. Restriction endonucleases, DNA polymerase I, polynucleotide kinase, M-MuLV-reverse transcriptase, some sequencing oligonucleotide primers, and nucleotide mixes for ^{35}S-labeled DNA sequencing were purchased from New England Biolabs. Cloned Klenow fragment, pUC-phagemid (pTZ18), Sequenase (T7) DNA polymerase and Sequenase DNA sequencing kit were obtained from U.S. Biochemical Corp. Inhibit-ACE was purchased from 5 Prime – 3 Prime, Inc. [γ-^{32}P]ATP, [α-^{32}P]dATP, [α-^{32}P]dCTP, and [α-^{35}S]-dATPαS were acquired from Amersham Corp. Nitrocellulose and nytran (nylon membrane) were from Schleicher and Schuell. M13mp18, M13mp19, Mono Q HR5/5 (FPLC column), and NAP-10 desalting columns were purchased from Pharmacia LKB Biotechnologies. *E. coli* strain DH5α was from Bethesda Research Laboratories. All other chemicals were of reagent grade from various sources.

† Support was provided by Research Grant GM 21083 (V.L.S.) and Postdoctoral Fellowship GM 12373 (K.L.K.-K.) and CA 12227 (J.K. deR.) from the National Institutes of Health and SIG-6 from the American Cancer Society (J.K.deR.). Most of the experimental work was performed in the Deparment of Biochemistry, Temple University School of Medicine.

* To whom correspondence should be addressed.

‡ Department of Biochemistry, Temple University School of Medicine.

§ Department of Biochemistry, Albert Einstein College of Medicine.

∥ Present address: Cephalon Inc., 145 Brandywine Parkway, West Chester, PA 19380.

⊥ Fels Research Institute, Temple University School of Medicine.

0006-2960/89/0428-8726$01.50/0

Protein and Peptide Analysis. Peptides from purified AMP nucleosidase were isolated after trypsin digestion of the citraconylated protein (Atassi & Habeeb, 1972). Peptide separation was on a Vydak C18TP column (Bodman) using a 130-min acetonitrile gradient of 10–70% in 0.1% trifluoroacetic acid at pH 3.0. A further purification using 1-propanol from 0 to 60% in 0.1% trifluoroacetic acid, pH 3.0, was also used (Hermodson & Mahoney, 1983). The N-terminal and peptide sequences were determined by Dr. Mark Hermodson, Purdue University. Amino acid composition was determined by HPLC of derivatized amino acids using the Picotag system of Waters Associates. The determination was done by W. Bear of Albert Einstein College of Medicine.

Bacterial Plasmids, Strains, and Media. The preparation of pHL8 plasmid that contains the *E. coli* gene for AMP nucleosidase was described earlier (Leung & Schramm, 1984). It was produced in quantity for sequencing from *E. coli* in strain HL359 (Leung & Schramm, 1984). The shuttle vector for *E. coli* and *Azotobacter vinelandii* was pKT230, a derivative of RSF1010 (Bagdasarian et al., 1981), provided by Dr. Paul Bishop, University of North Carolina, in *E. coli* strain SK1590. A strain of SK1590 free of plasmid pKT230 was selected after rapid growth and serial transfers in antibiotic-free liquid medium. This strain was used as the host for recombinants of pKT230. *E. coli* strain HB101 was the gift of Dr. John Manfredi, University of North Carolina at Chapel Hill. *A. vinelandii* strains and their growth conditions have been reported previously (Leung & Schramm, 1981). The culture medium used for cultivation of *A. vinelandii* before and after transformation was described by David et al. (1981). *E. coli* strains were grown in LB medium (Miller, 1972). The synthetic phosphate-free medium used was buffered with 4-morpholinepropanesulfonic acid (MOPS) (Neidhardt et al., 1974), and thiamin (10 μg/mL) was also added.

Oligonucleotides. Oligonucleotides for radioactive probes, primer extension, and sequencing of *amn* were prepared as (17–21)-mers with the 5′-trityl group retained by using an Applied Biosystems Model 380A and B DNA synthesizer. Some oligodeoxynucleotides were purified by HPLC on a C_{18} μBondapak column (Waters) using a 40-min gradient of 5–45% acetonitrile in 0.1 M triethylammonium acetate, pH 7.0. The peak fraction at 254 nm from the column was lyophilized, treated with 200 μL of 80% acetic acid at room temperature for 20 min to remove the trityl group, lyophilized, resuspended in 100 μL of H_2O, and lyophilized to dryness. Other oligonucleotides were treated with 200 μL of 80% acetic acid for 20 min at room temperature to remove the trityl group and then desalted on a Sephadex G-25 column (1.3 × 2.7 cm) (NAP-10), equilibrated with 10 mM NaOH, pH 12, containing 0.5 M NaCl. The oligonucleotides were purified by FPLC on a Mono Q HR 5/5 column using a 30-min gradient of 0.5–1.0 M NaCl in 10 mM NaOH, pH 12. The peak fraction at 254 nm was desalted on a Sephadex G-25 column (1.3 × 2.7 cm) equilibrated with 10 mM Tris buffer, pH 8.0, containing 1 mM EDTA.

DNA Sequencing. Plasmid pHL8, a derivative of pBR322 containing *amn* and flanking regions, was purified by cesium chloride–ethidium bromide density gradient equilibrium centrifugation (Maniatis et al., 1982). Restriction fragments end-labeled with ^{32}P either were redigested with a second restriction enzyme and separated on 4% polyacrylamide gels or separate strands were isolated from strand-separating gels of 5% polyacrylamide (Maniatis et al., 1982). These fragments were then sequenced according to the method of Maxam and Gilbert (1977).

A novel technique of nearest-neighbor fragment analysis (developed by J.K.deR.) was used to map the location of the *Hin*fI and *Hpa*II fragments in pHL8. Isolated restriction fragments were labeled at the 3′ ends with [α-^{32}P]dATP or [α-^{32}P]dCTP and the Klenow fragment of DNA polymerase, denatured and hybridized to denatured pHL8 plasmid, extended with DNA polymerase, and recut with *Hin*fI or *Hpa*II to transfer label to the 5′-ends of adjacent fragments. This method proved useful in producing the restriction map for sequencing a portion of the gene by using Maxam and Gilbert (1977) sequence analysis. The remainder and confirming DNA sequence was obtained by using the dideoxy method of Sanger et al. (1977), as adapted for ^{35}S (Biggin et al., 1983) using either Klenow fragment of DNA polymerase or Sequenase (T7 DNA polymerase). Several areas of high G-C content were resolved by reverse transcriptase sequencing (Chen & Seeburg, 1985) or with Klenow fragment of DNA polymerase with incubation at 47° C. Oligonucleotide primers were designed to generate sequence at several locations along the 2.7-kb[1] *Hin*fI fragment of pHL8 that had been subcloned in both orientations into M13.

Construction of pKT230·amn and pTZ18·amn. An *E. coli–A. vinelandii* shuttle vector containing the regulatory and coding region of the AMP nucleosidase gene was constructed in pKT230, a broad range, high copy number vector derived from RSF1010 (Bagdasarian et al., 1981). A 2.7-kb *Hin*fI fragment from pHL8 (Leung & Schramm, 1984) that contained *amn* was cloned into the *Bam*HI site of pBR322 to form pBR·*amn*. A *Hin*dIII, *Sal*I digest of pBR·*amn* was ligated into a *Hin*dIII, *Sal*I digest of pUC18 to form pUC·*amn* in a 6-kb plasmid construct selected by ampicillin. The region of pUC·*amn* containing *amn* was excised with *Hin*dIII and *Bam*HI and ligated into pKT230 that had been digested with *Hin*dIII and *Bam*HI to remove a 0.6-kb fragment. This construct (pKT·*amn*, Figure 1) is a 14.5-kb plasmid containing resistance for streptomycin and the entire coding and regulatory region for AMP nucleosidase. The 2.7-kb *Hin*fI fragment of pHL8 in the M13·*amn* construct was restricted with *Kpn*I and *Hin*dIII and then ligated in both orientations into pTZ18 (a pUC18 phagemid). The pTZ18·*amn* was transformed (Leung & Schramm, 1984) into *E. coli* strains DH5α (a *rec*A⁻ strain) and HL359 (a *amn* mutant strain) (Leung & Schramm, 1984).

Expression of AMP Nucleosidase Genes. The effects of phosphate and cAMP on expression of *amn* and *lac*Z (β-galactosidase) were determined in *E. coli* strain LS853 (*E. coli* Genetic Stock Center, Dr. B. Bachman), which is a deletion mutation in *cya*, adenylate cyclase. Cells for experiments of phosphate repression of AMP nucleosidase were grown in medium (Neidhardt et al., 1974) containing 76 mM phosphate and thiamin (10 μg/mL). Cells were grown to OD_{600} = 0.5, rinsed twice in saline, and used to inoculate fresh culture containing different concentrations of phosphate. Cells were grown from an original OD_{600} of 0.1 to an OD_{600} of 0.5 in the test medium (Neidhardt et al., 1974) and then were harvested. Addition of cAMP was to cells at OD_{600} = 0.5 that had been treated as above. Cells were harvested 40 min after addition of cAMP. Cells were harvested by centrifugation and disrupted by sonication in 0.1 M Tris-HCl, pH 8, 0.1 M NaCl, 0.1 mM EDTA, 0.1 mM dithiothreitol, and 3 μM phenylmethanesulfonyl fluoride. The extracts were assayed for AMP nucleosidase (Leung & Schramm, 1984) and β-galactosidase

[1] Abbreviations: kb, kilobase(s); ss, single stranded; bp, base pair; SSPE, 0.15 M NaCl, 10 mM NaH_2PO_4, 1 mM EDTA, pH 7.4; SDS, sodium dodecyl sulfate.

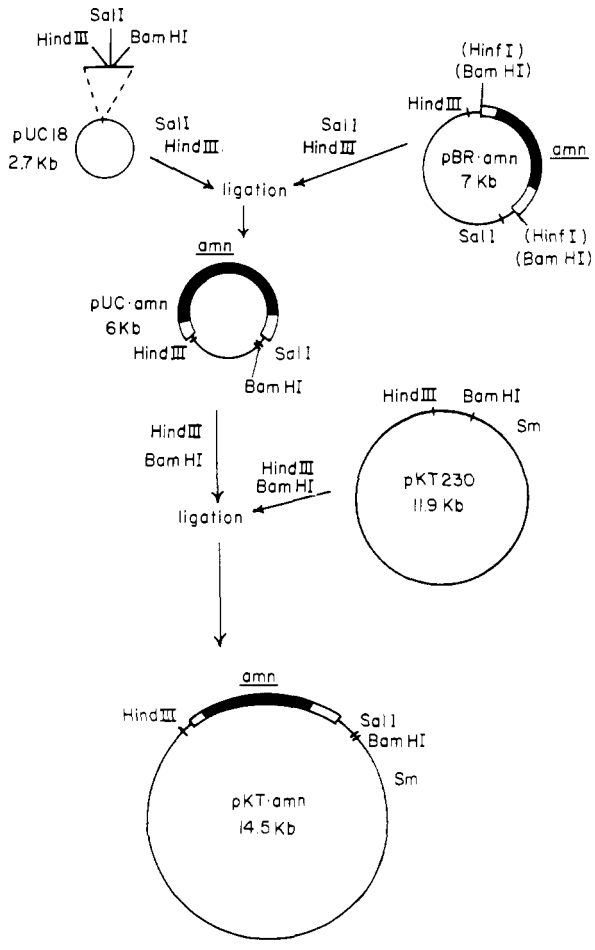

FIGURE 1: Construction of an *E. coli–A. vinelandii* shuttle vector containing AMP nucleosidase. The solid, heavy line indicates the coding and regulatory region for AMP nucleosidase and is labeled "*amn*". The flanking regions of the *E. coli* gene are indicated as open regions of the heavy lines. The insert in pBR·*amn* is the large *Hin*fI fragment shown in Figure 2, cloned into the *Bam*HI site of pBR322. The *Hin*fI and *Bam*HI sites in parentheses in pBR·*amn* were removed in the blunt-end ligation of the *Hin*fI fragment into the *Bam*HI site. The figures are not to exact scale. "Sm" indicates streptomycin resistance. Other details are given under Experimental Procedures.

(Maniatis et al., 1982) activities as previously described.

RNA Studies. Total RNA was isolated according to the method of Ausubel et al. (1989) from strain HL359 (Leung & Schramm, 1984) containing the plasmid construct pTZ18U·*amn* which was grown in 10-mL cultures of LB media (Miller, 1972) containing 60 μg/mL ampicillin. During the isolation, Inhibit-Ace (1 unit/30 μL of reaction volume) was added during the lysis step instead of diethyl pyrocarbonate. After the ethanol precipitation, the RNA was extracted with an equal volume of phenol and then with chloroform. The RNA was precipitated with ethanol and resuspended in 100 μL of sterile water containing Inhibit-Ace (3.3 units).

The 5′-end of the *amn* transcript was determined by primer extension using a synthetic oligonucleotide (5′-CAGCGCGTCGAGTTTATCCAG-3′) that is complementary to nucleotides 40–60 in the coding region of *amn*. The primer (2.5 ng) was added to the RNA (0.1, 1.0, 10, 50, 100 μg) in 10 mM Tris buffer, pH 8.0, containing 0.1 mM EDTA and 0.3 M NaCl to a total volume of 30 μL. The mixture was heated for 5 min at 75 °C and incubated at 55 °C for 1 h. The RNA was precipitated with ethanol, dried, and resuspended in 25 μL of 50 mM Tris buffer, pH 8.3, containing 50 mM KCl, 5 mM $MgCl_2$, 5 mM DTT, 4.8 μM dATP, and 40 μM dCTP, dGTP, and dTTP with 200 μCi (8.3 μM) of [α-^{35}S]dATPαS. M-MuLV-reverse transcriptase (50 units) was added, and the reaction was incubated for 90 min at 42 °C. After RNase digestion and phenol/chloroform extraction, the product was dissolved in 6 μL of formamide loading buffer. The products of primer extension and a sequencing ladder of the same primer annealed to ssM13·*amn* (the DNA strand identical with the mRNA strand) were loaded on a DNA sequencing gel. After autoradiography, the sites of extension termination were determined by comparison to the standard sequencing ladder.

To determine the size of the *amn* transcript, Northern analysis was carried out by electrophoresis of the total *E. coli* RNA (25 μg) on a 1.2% formaldehyde denaturing gel according to the method of Ausubel et al. (1987) except 6% formaldehyde was added to the running buffer and the buffer was recirculated during the electrophoresis. The fractionated RNA and RNA standards were transferred from the agarose gel to nylon membrane according to the manufacturer's instructions. The nylon membrane was prehybridized at 56 °C for 3 h with 3 mL of 5× SSPE (0.75 M NaCl, 50 mM NaH_2PO_4, 5 mM EDTA), 5× Denhardt's (Maniatis et al., 1982), 0.5% SDS, 7.5 mM sodium pyrophosphate, and 0.1 mg/mL denatured salmon sperm DNA. To the prehybridization solution, 50 ng of ^{32}P end labeled oligonucleotide (primer extension 21-mer) was added, and the solution was incubated overnight at 56 °C. The blot was exposed to film with a lightning plus (Du Pont) intensifying screen for 12–24 h at −70 °C. To visualize the ribosomal standards, a duplicate portion of the gel was stained with ethidium bromide (250 μg/500 mL of water) for 1 h and destained in water for 30 min. Also, the nylon membrane was soaked in 0.4 M sodium acetate containing 0.02% methylene blue for 10 min and then rinsed in 1× SSPE for 15 min to stain the ribosomal standards.

Computer Analysis of DNA and Protein Sequence. The BIONET National Computer Resource for Molecular Biology was used to translate the DNA sequence to the amino acid sequence and to analyze the amino acid sequence for secondary structure. Hydropathicity analysis used the window values from Kyte and Doolittle (1982) with the arithmetic window average at the *i*th residue calculated across six residues, i_{-3}–i_{+3} inclusive. Secondary structure analysis was predicted from the Chou and Fasman algorithm (1978). Sequence homology for *E. coli* AMP nucleosidase with yeast AMP deaminase used the region pairwise alignment function of BIONET with solution parameters set at amino residue length 2, deletion weight 1, length factor 0, matching weight 1, and spread factor 50. Similarities between AMP nucleosidase and other proteins were determined by a search of the transcribed DNA data bank available from BIONET. Search strategy for the BIONET data bank used IFIND with default values. Management of the DNA sequence was aided by the GENED, SEQ, and PEP programs of the BIONET system, which is funded by the National Institutes of Health, Grant P41 RR01685 (Smith et al., 1986).

RESULTS

Deoxynucleotide Sequence of amn. The restriction map of the *E. coli* genome in the neighborhood of the *amn* gene as compared to the *E. coli* genomic restriction map of the 2005.8–2006.7-kb region generated by Kohara et al. (1987) is given in Figure 2. This 9-kb genomic piece is contained within the 8C10 phage clone and maps between the *fla*A and *sbc*B genes (42.7 and 43.6 min). This also agrees with the region mapped by us (Leung & Schramm, 1984) that places *amn* at 43.3 min. The restriction map of the constructs

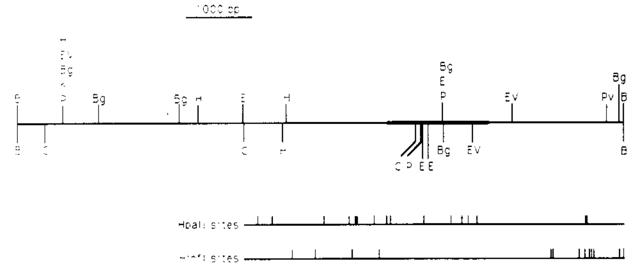

FIGURE 2: Restriction map for the *E. coli* gene for AMP nucleosidase. The restriction sites in the neighborhood of the AMP nucleosidase gene are shown. The abbreviations B, C, Bg, E, EV, K, P, and Pv indicate sites for *Bam*HI, *Cla*I, *Bgl*I, *Eco*I, *Eco*RV, *Hin*dIII, *Kpn*I, *Pst*I, and *Pvu*II, respectively. The heavy line denotes the *amn* gene. The direction of transcription is left to right. The restriction sites above the line of the 9-kb *Bam*HI fragment were measured from the *E. coli* genomic map generated by Kohara et al. (1987). The restriction sites below the line, the *Hpa*II and *Hin*fI sites, were determined in this paper and by Leung and Schramm (1984) from the *amn* constructs pHL150, pHL8, and pBR·*amn*.

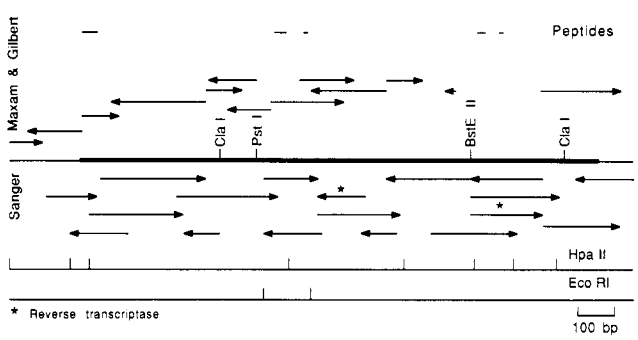

FIGURE 3: Deoxynucleotide sequencing strategy for *E. coli amn* gene. The lines labeled "Peptides" represent the sequences derived from tryptic fragments of the protein, and these sequences are underlined in Figure 4. The fragments labeled "Maxam & Gilbert" indicate the fragments, direction of sequencing, and extent of sequences determined by this method. Restriction sites for *Cla*I, *Pst*I, *Bst*EII, *Hpa*II, and *Eco*RI are indicated. The fragments labeled "Sanger" were sequenced by using either the M13 universal or synthetic primers to initiate at specific sites on DNA. The fragment marked with an asterisk were sequenced both by reverse transcriptase and by Klenow fragment DNA polymerase at elevated temperature to sequence G-C-rich regions of the gene. The thick line represents the coding region of AMP nucleosidase. The direction of transcription is from left to right.

pHL150 and pHL8 containing *amn* generated by Leung and Schramm (1984) and the construct pBR·*amn* (this paper) agrees well with the genomic *Bam*HI fragment. The sequencing strategy for determination of the DNA sequence of the AMP nucleosidase gene in *E. coli* is given by Figure 3. The initial alignment of restriction fragments used the label-transfer technique described under Experimental Procedures. Over 96% of the DNA was sequenced from both strands by using a combination of Sanger et al. (1977) and Maxam and Gilbert (1977) techniques. The interpretation of the DNA sequence was confirmed by the amino acid sequences of peptides obtained from Edman degradations of tryptic fragments and from the N-terminal sequence of the intact protein.

The DNA sequence for *amn* with the 5′- and 3′-flanking regions is given in Figure 4. The open reading frame which corresponds to the N-terminal amino acid sequence obtained directly from Edman degradation of the intact protein begins with an ATG, numbered bases 1–3 in Figure 4. The open reading frame of 1449 base pairs codes for 483 amino acids with a peptide molecular weight of 53 848. The sequence ends with two adjacent stop codons (TAA TAA) at positions 1450–1455 following the last base of the open reading frame.

```
                                 -180                                        -150
      T TGC CGG GAA CTG GCA TAG CGT CGC CAT TTA TTT GCT GGC TGG ATG CCT GAT TTC CGC AAT
                   -120                                      -90
GAC CGC TTT GTT GAT GAA AGA CAG TCA ACG CGC TTG ATA GCC TGG CGA AGA TCA TCC GAT CTT CGC CTT

      n TTT CAn ATn TCT GTn Ac
                          nT GTG Ann Tnn n-n Cnn ATn T
      -60                 ============================              -20                 -1
ACA CTT TTG TTT CAC ATT TCT GTG ACA TAC TAT CGG ATG TGC GGT AAT TGT ATG GAA CAG GAG ACA CAC
         "-35"                     "-10"     •   •               •• ••          S.D.
1                                  30                                      60
ATG AAT AAT AAG GGC TCC GGT CTG ACC CCA GCT CAG GCA CTG GAT AAA CTC GAC GCG CTG TAT GAG CAA
Met Asn Asn Lys Gly Ser Gly Leu Thr Pro Ala Gln Ala Leu Asp Lys Leu Asp Ala Leu Tyr Glu Gln
                         90                             120
TCT GTA GTC GCA TTA CGC AAC GCC ATT GGC AAC TAT ATT ACA AGT GGC GAA TTA CCT GAT GAA AAC GCC
Ser Val Val Ala Leu Arg Asn Ala Ile Gly Asn Tyr Ile Thr Ser Gly Glu Leu Pro Asp Glu Asn Ala
            150                             180
CGC AAA CAA GGT CTT TTT GTC TAT CCA TCA CTG ACC GTA ACG TGG GAC GGT AGC ACA ACC AAT CCC CCC
Arg Lys Gln Gly Leu Phe Val Tyr Pro Ser Leu Thr Val Thr Trp Asp Gly Ser Thr Thr Asn Pro Pro
210                                240                                     270
AAA ACG CGC GCA TTT GGT CGC TTT ACC CAC GCA GGC AGC TAC ACC ACC ACG ATT ACT CGC CCT ACT CTC
Lys Thr Arg Ala Phe Val Arg Phe Thr His Ala Gly Ser Tyr Thr Thr Thr Ile Thr Arg Pro Thr Leu
                           300                             330
TTT CGT TCG TAT CTT AAT GAA CAA CTT ACG TTG CTG TAT CAG GAT TAT GGT GCG CAT ATC TCA GTG CAA
Phe Arg Ser Tyr Leu Asn Glu Gln Leu Thr Leu Leu Tyr Gln Asp Tyr Gly Ala His Ile Ser Val Gln
                360                             390
CCC TCG CAG CAT GAA ATC CCT TAT CCT TAT GTC ATC GAT GGC TCT GAA TTG ACA CTT GAT CGC TCA ATG
Pro Ser Gln His Glu Ile Pro Tyr Pro Tyr Val Ile Asp Gly Ser Glu Leu Thr Leu Asp Arg Ser Met
    420                                   450                                         480
AGC GCT GGG TTA ACT CGC TAC TTC CCG ACA ACA GAA CTG GCG CAA ATT GGC GAT GAA ACT GCA GAC GGC
Ser Ala Gly Leu Thr Arg Tyr Phe Pro Thr Thr Glu Leu Ala Gln Ile Gly Asp Glu Thr Ala Asp Gly
                              510                                     540
ATT TAT CAT CCA ACT GAA TTC TCC CCG CTA TCG CAT TTT GAT GCG CGC CGC GTC GAT TTT TCC CTC GCA
Ile Tyr His Pro Thr Glu Phe Ser Pro Leu Ser His Phe Asp Ala Arg Arg Val Asp Phe Ser Leu Ala
                    570                                 600
CGG TTG CGC CAT TAT ACC GGT ACG CCA GTT GAA CAT TTT CAG CCG TTC GTC TTG TTT ACC AAC TAC ACA
Arg Leu Arg His Tyr Thr Gly Thr Pro Val Glu His Phe Gln Pro Phe Val Leu Phe Thr Asn Tyr Thr
        630                                660                                         690
CGT TAT GTG GAT GAA TTC GTT CGT TGG GGA TGC AGC CAG ATC CTC GAT CCT GAT AGT CCC TAC ATT GCC
Arg Tyr Val Asp Glu Phe Val Arg Trp Gly Cys Ser Gln Ile Leu Asp Pro Asp Ser Pro Try Ile Ala
                                   720                                 750
CTT TCT TGT GCT GGC GGG AAC TGG ATC ACC GCC GAA ACC GAA GCG CCA GAA GAA GCC ATT TCC GAC CTT
Leu Ser Cys Ala Gly Gly Asn Trp Ile Thr Ala Glu Thr Glu Ala Pro Glu Glu Ala Ile Ser Asp Leu
                         780                                 810
GCA TGG AAA AAA CAT CAG ATG CCA GCA TGG CAT TTA ATT ACC GCC GAT GGT CAG GGT ATT ACT CTG GTG
Ala Trp Lys Lys His Gln Met Pro Ala Trp His Leu Ile Thr Ala Asp Gly Gln Gly Ile Thr Leu Val
            840                             870
AAT ATT GGC GTG GGA CCG TCA AAT GCT AAA ACC ATC TGC GAT CAT CTG GCA GTG CTA CGC CCG GAT GTC
Asn Ile Gly Val Gly Pro Ser Asn Ala Lys Thr Ile Cys Asp His Leu Ala Val Leu Arg Pro Asp Val
900                                930                                     960
TGT TAT ATT GGT CAC TGT GGC GGA TTA CGT AAA GTC AGG CCA TTG GCC GAT TAT GTA CTT GCA CAC GCT
Cys Tyr Ile Gly His Cys Gly Gly Leu Arg Lys Val Arg Pro Leu Ala Asp Tyr Val Leu Ala His Ala

                           990                                 1020
TAT TTA CGC GAT GAC CAC GTT CTT GAT GCG GTT CTG CCG CCC GAT ATT CCT ATT CCG AGC ATT GCT GAA
Tyr Leu Arg Asp Asp His Val Leu Asp Ala Val Leu Pro Pro Asp Ile Pro Ile Pro Ser Ile Ala Glu
               1050                            1080
GTG CAA CGT GCG CTT TAT GAC GCC ACC AAG CTG GTG AGC GGC AGG CCC GGT GAG GAA GTC AAA CAG CGG
Val Gln Arg Ala Leu Tyr Asp Ala Thr Lys Leu Val Ser Gly Arg Pro Gly Glu Glu Val Lys Gln Arg
   1110                                1140                                         1170
CTA CGT ACT GGT ACT GTG GTA ACC ACA GAT GAC AGG AAC TGG GAA TTA CGT TAC TCA GCT TCT GCA CTT
Leu Arg Thr Gly Thr Val Val Thr Thr Asp Asp Arg Asn Trp Glu Leu Arg Tyr Ser Ala Ser Ala Leu
                              1200                                1230
CGT TTT AAC TTA AGC CGG GCC GTA GCA ATT GAT ATG GAA AGT GCA ACC ATT GCC GCG CAA GGA TAT CGT
Arg Phe Asn Leu Ser Arg Ala Val Ala Ile Asp Met Glu Ser Ala Thr Ile Ala Ala Gln Gly Tyr Arg
                    1260                            1290
TTC CGC GTG CCA TAC GGG ACA CTA CTG TGT GTT TCA GAT AAA CCG TTG CAT GGC GAG ATT AAA CTT CCC
Phe Arg Val Pro Tyr Gly Thr Leu Leu Cys Val Ser Asp Lys Pro Leu His Gly Glu Ile Lys Leu Pro
       1320                                    1350                                    1380
GGC CAG GCT AAC CGT TTT TAT GAA GGC GCT ATT TCC GAA CAT CTG CAA ATT GGC ATT CGG GCG ATC GAT
Gly Gln Ala Asn Arg Phe Tyr Glu Gly Ala Ile Ser Glu His Leu Gln Ile Gly Ile Arg Ala Ile Asp
                                   1410                                    1440        1449
TTG CTG CGC GCA GAA GGC GAC CGA CTG CAT TCG CGT AAA TTA CGA ACC TTT AAT GAG CCG CCG TTC CGA
Leu Leu Arg Ala Glu Gly Asp Arg Leu His Ser Arg Lys Leu Arg Thr Phe Asn Glu Pro Pro Phe Arg
                       1470                                1500
TAA TAA GTT CTG AAC CTT CTT ATC AGA AAT AGA TGT GAA GGA GTA AGT AAG ACA CCT GGC AAA TAG CCT
                    -------                     ------                                 =====

             1530                                   1560
GCA ATA ATC GTG GGC TGT TTG CTT CCT TGG GCG GAT AAC GAG TTT TAT TAT CGT CTT AAT GAT TTC CAC
==          ========

               1600
ATA TTA AAA GCA AGT AT
```

FIGURE 4: Deoxynucleotide sequence and deduced amino acid sequence of *E. coli amn*. The *amn* nucleotide sequence of the mRNA-identical (plus) strand of 1449 bp plus 199 bp 5′ and 155 bp 3′ to the protein-coding region are shown. The coding region begins at base 1. The 5′-flanking region is indicated by negative numbers, and the 3′-flanking region begins with base 1450. In the 5′-flanking region, the transcriptional start sites are marked with underdots, the Shine–Delgarno sequence is underlined and labeled S.D., and the "Pribnow" or "−10" sequence and the "−35" sequences are underlined and labeled. The sequences of the underlined amino acids were determined directly by automated Edman degradation of peptides isolated from purified enzyme. The sequence that is double-overlined is the region similar to the consensus sequence for interaction with cAMP receptor protein (deCrombrugghe et al., 1984). The overlined sequence indicates the region similar to the consensus sequence for phosphate regulation (Makino et al., 1986). The bases that comprise the consensus sequences are shown above the matching regions, homologous bases are indicated in capital letters, bases labeled "n" are variable in the known consensus sequences, and the dash is an introduced space to allow maximum homology with the *amn* promoter. In the 3′-flanking region, the double-underlined bases could form a base-paired stem with a poly(A) region preceding and a poly(U) region following the proposed termination signal. Alternatively, the two dashed regions in the 3′-region could form a simple terminator signal.

Characterization of E. coli amn Coding Region. Tryptic peptides from citraconylated AMP nucleosidase [see Experimental Procedures of the following paper (Meyer et al., 1989)] were purified and sequenced to confirm that the open reading frame corresponded to the peptide sequence. In addition, intact

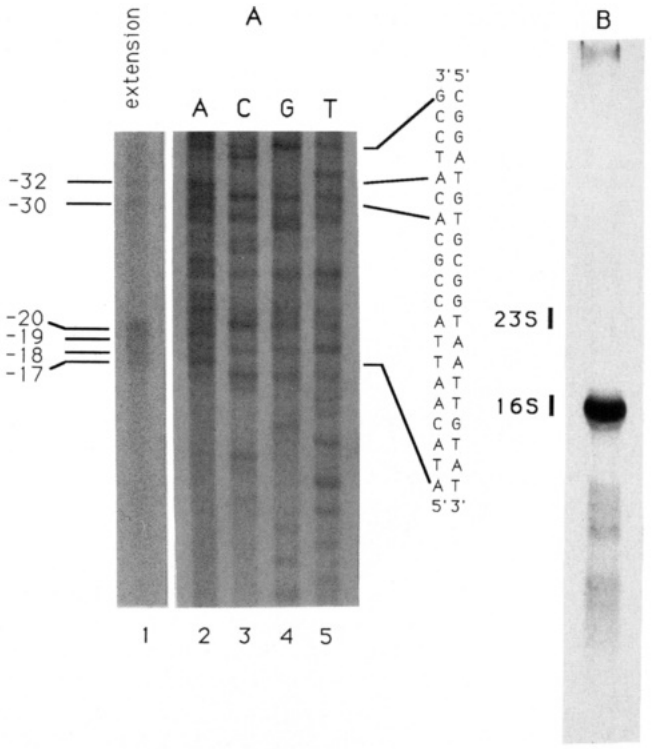

FIGURE 5: Primer extension and Northern blot analysis of *amn* transcript. (A) In lane 1, unlabeled synthesized oligonucleotide (21-mer) (2.5 ng) was annealed to total *E. coli* RNA (0.1 μg) and then extended with reverse transcriptase in the presence of [α-^{35}S]-dATPαS. The base numbers on the left margin denote the transcriptional start sites for RNA polymerase on *amn*. Lanes 2–5 represent the sequencing ladder of M13·*amn* ss DNA template annealed to the primer extension 21-mer. (B) A Northern blot of total *E. coli* RNA (25 μg) was hybridized to ^{32}P end labeled oligonucleotide (primer extension 21-mer). The positions of the stained ribosomal markers are indicated.

AMP nucleosidase was subjected to automated Edman degradation to determine the amino acid sequence of the N-terminal region. The tryptic-digested peptides were separated by HPLC on a Vydac column, and four peptides were selected for further purification and sequence analysis. The amino acid sequences obtained from the selected peaks are underlined in Figure 4. All of the peptides and the N-terminal sequences were located in the observed reading frame of the DNA. Total amino acid composition of AMP nucleosidase protein was determined following acid hydrolysis (data not shown) and is in good agreement with the amino acid sequence deduced from the deoxynucleotide sequence.

Characterization of 5′- and 3′-Flanking Regions of E. coli amn. To determine the control region of *amn*, the 5′-end of the mRNA was identified by primer extension. A 21-mer oligonucleotide primer designed to anneal to nucleotides 40–60 in the coding region of *amn* was hybridized to total *E. coli* RNA isolated from strain HL359 containing the plasmid pTZ18·*amn*. After primer extension with reverse transcriptase, the sizes of the extension products were determined by electrophoresis on a DNA sequencing gel and compared to a sequencing ladder of the extension primer annealed to ssM13·*amn* (the same strand as the m-RNA). Six extension products of 77, 78, 79, 80, 90, and 92 nucleotides in length were found (Figure 5A). These products correspond to transcriptional start sites at bases −17 to −20, −30, and −32 upstream of the ATG translational initiation codon. A ladder of bands was observed for the extended products when the oligonucleotide was annealed to 10, 50, and 100 μg of total RNA and when 50 μg RNA (no primer) was annealed (data not shown). No ladders, only the extension products, were seen with the primer–RNA (0.1 and 1 μg) hybridizations. Presumably, the ladder is a result of secondary structure in the mRNA from which priming can occur during the extension. No other bands were observed within 500 bases of the translational initiation codon.

To determine the approximate size of the *amn* transcript, Northern blot analysis was conducted with the same total RNA preparation. When the ^{32}P end labeled oligonucleotide (primer extension 21-mer) was hybridized to the nylon membrane at T_D −10 °C, a single band was observed at 1.7 ± 0.2 kb on the basis of comparison to ribosomal RNA standards (both *E. coli* and yeast).

Analysis of the nucleotide sequence upstream of the translational initiation codon revealed a Shine and Delgarno (1974) ribosome binding sequence (AGGAG) which is 6 bases from the initiation codon and contains 5 nucleotides complementary to the 3′-end of 16S RNA (Gold & Stormo, 1987). Additionally, analysis upstream of the transcription start site revealed the −10 and −35 promoter elements (15 bp apart) with 4 of 6 bases of each element homologous to the consensus sequence of *E. coli* promoter sequences (Hawley & McClure, 1983; Harley & Reynolds, 1987). These regions are underlined and labeled in Figure 4.

The 5′-flanking region was also analyzed for any regulatory regions. The region from −44 to −61 shows much similarity to sequences found in promoters of genes whose expression is sensitive to phosphate levels (14 of 18 bases) (Makino et al., 1986). This region is overlined in Figure 4, and the proposed consensus sequence is shown. A sequence that overlaps with the 3′-end of the phosphate regulation region is found at −45 to −49 (AGTGT), which is similar to the major recognition site of the consensus sequence for cAMP receptor protein. This region is double-overlined in Figure 4, and the consensus sequence is given.

Expression of E. coli AMP Nucleosidase in A. vinelandii. The coding region for AMP nucleosidase contained in pKT·*amn* (Figure 1) was tested for the ability to encode for catalytically functional enzyme by cloning into a strain of *A. vinelandii* that produces catalytically defective AMP nucleosidase (Leung & Schramm, 1981). Transformation of *A. vinelandii* with the pKT·*amn* resulted in synthesis of the enzyme. The catalytic activities of AMP nucleosidase in extracts of wild-type, mutant, and transformed *A. vinelandii* are given in Table I. The construct pTZ18·*amn* was transformed into the *E. coli* strains DH5α (*rec*A$^-$ strain) and HL359 (*amn* mutant strain) (Leung & Schramm, 1984). The activities of AMP nucleosidase in cell extracts are shown in Table I. The AMP nucleosidases from *E. coli* and *A. vinelandii* are immunochemically distinct as well as having different allosteric regulatory mechanisms (Leung & Schramm, 1980).

Effect of Phosphate on Expression of the AMP Nucleosidase. Purified AMP nucleosidase from *E. coli* was identified on the two-dimensional electrophoresis maps of Neidhardt et al. (1983) and designated to position G52.0. AMP nucleosidase was one of the proteins reported to be induced when phosphate is limited in the culture medium. Growth of *E. coli* with 76 mM inorganic phosphate resulted in a decrease of the cell extract activity of AMP nucleosidase to a basal level of 6 nmol min^{-1} mg^{-1} in *E. coli* containing wild-type adenylate cyclase (Table II). A slight increase in activity was observed at 1.32 mM phosphate, and the activity increased 7.5-fold over the basal level on 53 μM phosphate, growth conditions that limit both the rate and extent of cell division. The increase in AMP nucleosidase activity seen when *E. coli* are grown in

Table I: Expression of *E. coli* AMP Nucleosidase in *E. coli* and *A. vinelandii*

source	AMP nucleosidase activity [nmol min^{-1} (mg of protein)$^{-1}$]
E. coli[a]	
FP4102[b]	2
HL359[c]	$<2 \times 10^{-5}$
FP4102 [pBR·*amn*]	300
HL359 [pTZ18·*amn*]	1500
DH5α [pTZ18·*amn*]	8000
A. vinelandii[d]	
OP (wild type)	25
LAVM[e]	0.3
LAVM [pKT230][f]	0.3
LAVM [pKT·*amn*][g]	16

[a] The values were obtained when *E. coli* strains were grown in the presence of phosphate. [b] The genotype of this strain is described in Leung and Schramm (1984). A variety of *E. coli* strains give the same enzymatic activity. [c] The *amn* mutant *E. coli* strain is described in Leung and Schramm (1984). [d] Value is taken from *A. vinelandii* grown in minimal medium with atmospheric N_2 as the sole nitrogen source. Similar values are obtained with growth in complex media. [e] LAVM is a catalytically defective AMP nucleosidase strain that is described in Leung and Schramm (1981). [f] LAVM [pKT230] is the organism defective in *amn* and transformed to kanamycin and streptomycin resistance with pKT230 containing no DNA insert. [g] LAVM [pKT·*amn*] contains pKT230 with the complete *amn* gene and was constructed as indicated in Figure 1.

Table II: Effect of Phosphate and cAMP on Expression of AMP Nucleosidase

[phosphate] in medium (mM)	adenylate cyclase	[cAMP] (mM)	AMP nucleosidase activity[a] (nmol min^{-1} mg^{-1})	β-galactosidase activity[a] (nmol min^{-1} mg^{-1})
0.053	wild type	0	45	
1.32		0	9	
76.0		0	6	
0.053	deficient	0	16	3.4
0.053		5	50	14.0
1.32	deficient	0	17	2.5
1.32		5	15	10.4
76.0	deficient	0	2	1.4
76.0		5	2	10.4

[a] The activities of AMP nucleosidase and β-galactosidase were measured in triplicate from cell extracts of *E. coli* treated with variable concentrations of phosphate and cAMP. The standard errors are less than 25% of the indicated values. See Experimental Procedures for detailed procedures.

53 μM phosphate suggests that the *amn* gene is induced by low phosphate concentrations. This is consistent with the observation of a highly conserved phosphate regulation sequence at −44 to −61 in the 5′-flanking region of *amn* (Figure 4). This sequence has been observed in other *E. coli* genes that are responsive to inorganic phosphate levels (Makino et al., 1986).

Effect of cAMP on Expression of the AMP Nucleosidase Gene. In addition to the sequence proposed for phosphate regulation of *amn*, a DNA sequence from −30 to −50 is similar to the proposed consensus sequence for binding of the cAMP receptor protein (deCrombrugghe et al., 1984). Growth of wild-type *E. coli* in the presence of added cAMP had no effect on AMP nucleosidase activity in cell extracts. Strains of *E. coli* deficient in adenylate cyclase gave low AMP nucleosidase activity when grown in medium containing 76 mM phosphate. Addition of cAMP at concentrations that induced *lacZ* (β-galactosidase) had no effect on the activity of AMP nucleosidase at high or intermediate concentrations of inorganic phosphate. However, at limiting phosphate, AMP nucleosidase activity was detected at levels similar to that of wild type only in the presence of added cAMP. Control experiments measured the effect of cAMP on induction of *lacZ*. β-Galactosidase activity was elevated by addition of cAMP at all concentrations of phosphate. Consistent with the presence of a similar DNA sequence for regulation by cAMP receptor protein, these results suggest that the *amn* gene is induced by cAMP at limiting phosphate concentrations.

Discussion

Sequence of amn. The open reading frame for *amn*, the AMP nucleosidase gene from *E. coli*, corresponds to the structure of the protein. A subunit molecular weight of 52000 for AMP nucleosidase from denaturing gel electrophoresis (Leung & Schramm, 1980) is in good agreement with the molecular weight of 53848 deduced from the sequence of the gene. The amino acid sequence of the N-terminus of the purified protein corresponds to the open reading frame of *amn*, as do four tryptic peptides derived from citraconylated AMP nucleosidase. The experimental amino acid composition agrees well with the amino acid composition deduced from the deoxynucleotide sequence of the gene. The hydropathicity profile of AMP nucleosidase (not shown) demonstrates well-ordered regions of hydrophobic and hydrophilic structure in the protein. The protein contains an excess of acidic amino acids with 55 (Asp + Glu) and 47 (Arg + Lys). The pK_a of 5.2 for the protein is consistent with this observation and makes AMP nucleosidase one of the more acidic proteins from *E. coli* on isoelectric focusing gels (e.g., [e.g., Neidhardt et al. (1983)]. The protein has low content of Met, Cys, and Trp, in common with many proteins from *E. coli* (de Boer & Kastelein, 1986).

Adenylate Binding Sites. AMP nucleosidase binds AMP at the catalytic site, MgATP at an allosteric regulatory site, and inorganic phosphate also at a regulatory site. The proposed "A-type" phosphoryl binding sequence for nucleotide binding proteins, Gly-X-X-X-X-Gly-Lys (Walker et al., 1982; Moller & Amons, 1985), is not found in the sequence, nor are any Gly-Lys pairs. The proposed "B-type" nucleotide binding sequence Arg-X-Gly-X_3-(hydrophobic)-Asp is also not found; however, the sequence Arg-Thr-Gly-Thr-Val-Val-Thr-Thr-Asp at nucleotides 1108–1134 of the reading frame and near the C-terminus of the protein (amino acids 370–378) has some resemblance to a B-type sequence (Walker et al., 1982; Moller & Amons, 1985). Comparison of the protein sequence indicated a region of similarity with the outer membrane pore protein precursor from *E. coli*, *pho*E. Amino acid residues 126–134 of AMP nucleosidase, *Val-Ile-Asp-Gly*-Ser-Glu-*Leu-Thr-Leu*, correspond to the sequence *Val-Ile-Asp-Gly*-Leu-Asn-*Leu-Thr-Leu* in residues 162–170 of the *pho*E precursor (Shinagawa et al., 1983). Since both proteins bind inorganic phosphate, it is possible that this sequence in AMP nucleosidase is part of the allosteric inhibitory site for inorganic phosphate. Additional searches for homology with the BIONET database detected no proteins with overall homology to AMP nucleosidase. The enzyme is not closely related to known proteins and has atypical nucleotide binding sites. Few regulatory ATP sites have been characterized, and the catalytic site for AMP nucleosidase apparently differs from the nucleotide binding sites in kinases. Site-directed mutagenesis experiments and/or the X-ray crystal structure (Giranda et al., 1986) will be necessary to locate these regions.

Homology with Other Proteins. The following paper (Meyer et al., 1989) demonstrates that only one small region of similarity exists between *E. coli* AMP nucleosidase and yeast AMP deaminase. Some similarity is to be expected since both enzymes have catalytic sites for AMP and allosteric sites for ATP and inorganic phosphate. A section of 54 amino acids

Table III: Proteins with Similarity to Amino Acids 122–149 of AMP Nucleosidase

Protein	first amino acid number and amino acid sequence[a]
AMP nucleosidase, *E. coli*	122 P Y P Y V I D G S E L T L D R S M S A G L T R Y F P T T
AMP deaminase, yeast	388 E K V I F R D G K L L T L D E V F R S L H L T G Y D L S
Nodulation protein B	117 Y E A P Y G V W S E E A L T E S A S A G L T A I H W S A
Cytochrome c oxidase	81 I Y P Y V I Q E L R P T L N E L G I S T P E E
T4 gene 57 B protein	10 G L Y V A A K F S E L T L D A L E E L Q R S L R
K-*ras* protein, human[b]	41 R K Q V V I D G E T C L L D I L D T A G Q E E Y S
Outer membrane *phoE* precursor	158 D F F G V I D G L N L T L Q Y Q G K N E N R D
IgG-binding protein	282 E K P E V I D A S E L T P A V T I Y K L V I
Tryptophanase, *E. coli*	113 L I K K R E Q E K G L D E S K M V A F S N Y F F T T

[a] The amino acids that are identical with AMP nucleosidase are underlined. [b] The same similarity is found for a variety of K-*ras*, H-*ras*, and N-*ras* proteins.

of the *E. coli* enzyme, centered on the region of similarity with yeast AMP deaminase, was compared to sequences in the protein data bank. The results indicate similarity of this region with a variety of proteins (Table III). The variety in the list suggests that this region may play a common structural role. In cytochrome *c* oxidase, residue 81 is the start of a transition from the heme binding pocket to a turn that leads to an α-helix (Dickerson & Timkovich, 1975). The family of *ras* genes is highly conserved in this region, which has been interpreted as the "phosphoryl group" region (Barbacid, 1987). The homologous Ala^{59} of *ras* proteins becomes the phosphoryl acceptor site when mutated to Thr^{59}. Thus, this region appears to function at the phosphoryl-transfer site. None of the nonfunctional *ras* domains are in the region shown in Table III.

Codon Use and Expression of amn. The low expression of the gene for *E. coli* AMP nucleosidase, even in conditions that were intended to maximize production of AMP nucleosidase protein, is reflected in the codon usage. For example, the first three amino acids following initiation at ATG are Asn-Asn-Lys. These amino acids are encoded by AAT for both Asn's and AAG for Lys. Highly expressed genes in *E. coli* use the AAC codon for Asn with a preference of 13 to 1, and AAA is used to encode Lys with a preference of 3 to 1 (de Boer & Kastelein, 1986). Similar patterns are observed (data not shown) for codon use of many amino acids of AMP nucleosidase. The codon bias index (Bennetzen & Hall, 1982) for AMP nucleosidase is 0.40, where a value of 1.0 indicates use of only preferred codons. A codon bias index of 0.8 corresponds to 10^5–10^6 protein molecules/cell, and a value of 0.5 corresponds to approximately 10^4 molecules/cell. At the basal level of AMP nucleosidase, approximately 10^2 molecules are present per cell, in agreement with the codon bias index. Despite the unfavorable codon bias, from 1% to 20% of total protein of *E. coli* can be produced as AMP nucleosidase in cells containing *amn* on pBR322 and pTZ18U, respectively (Leung & Schramm, 1984; Table I). Expression of the AMP nucleosidase gene thus appears to be a function of plasmid copy number, and the depletion of the cellular phosphate–repressor protein complex.

Flanking Regions of amn. The open reading frame for AMP nucleosidase begins with the ATG initiation codon and ends with two TAA stop codons. When the 3′-flanking region was analyzed for a termination signal, a region of dyad symmetry of 10 bases 3′ to the stop codon was observed (CCTTC at 1464–1468 and GAAGG at 1486–1490). The series of uridine residues that typically follows these simple terminators is not observed for *amn* (Platt, 1986). Another region from bases 1514–1520 to 1529–1535 could form a stem and loop structure that is preceded by a poly(A)-rich region and followed by a poly(U)-rich region. Similar structures have been shown to act as termination signals in *E. coli* (Dunn & Studier, 1980; Platt, 1985). This structure has a calculated $\Delta G°_{37}$ = −7.7 kcal/mol, which is in agreement with its putative formation (Freier et al., 1986).

In the 5′-flanking region overlapping putative regulatory regions for cAMP and inorganic phosphate are located between sequences −61 and −30. The proposed consensus sequences are given in Figure 4. Evidence that these are functional regulatory elements in vivo comes from observed changes in AMP nucleosidase activity in cells grown on limiting phosphate and excess cAMP. Phosphate represses the expression of *amn*, while cAMP induces expression, but only in the presence of low phosphate. The cAMP effect was only observed in *E. coli* strains deficient in adenylate cyclase. A regulatory mechanism consistent with these observations is that a phosphate binding repressor protein binds to the regulatory region when cellular phosphate is abundant, preventing the cAMP receptor protein from modulating expression. At low phosphate in normal cells, the phosphate–repressor complex is depleted, and cAMP–cyclic receptor protein can bind to the regulatory region to promote transcription. The consensus sequence for the "phosphate box", a regulatory region responsive to inorganic phosphate levels, has been proposed recently. The products of the *pho*B and *pho*R genes are proposed to regulate expression of *pho*A (alkaline phosphatase), *pho*S (phosphate binding protein), *pho*E (outer membrane binding protein), porinE, and *ugp*B (*sn*-glycerol 3-phosphate binding protein). The *pho*B gene product is a positive regulator at low phosphate, and the *pho*R gene product is a negative regulator at high phosphate. Although phosphate and cAMP regulation of AMP nucleosidase expression is consistent with the activity results and sequence data, additional studies to show changes in *amn* transcript levels when cells are grown under the different conditions are needed to verify this hypothesis. Studies of AMP nucleosidase expression in mutants of the *pho* and *crp* regulatory proteins will be required to understand the regulation of AMP nucleosidase.

The presence of multiple transcriptional initiation sites for the *amn* promoter may be due to the presence of the two putative antagonist regulatory elements (cAMP and phosphate). Consistent with the hypothesis, one or two initiation sites are seen with the *E. coli gal* operon which is cAMP dependent. When the cells are grown at high cAMP levels, transcription initiates at one site, whereas at low cAMP levels, initiation proceeds from both sites (Irani et al., 1989). Investigation of the *E. coli* K-12 *pho*E gene by S_1 nuclease mapping and primer extension analysis revealed multiple putative initiation sites depending upon the phosphate level of the medium from which the mRNA was isolated (Tommassen et al., 1987). These studies suggest that the two transcriptional start sites at −17 to −20 and −30 to −32 from the ATG codon for *amn* may represent distinct sites depending on the cellular regulation of the gene. The presence of multiple-length transcripts at each of these sites could conceivably arise by 5′ exonucleolytic activity (Deutscher, 1985), causing multiple, closely related RNA transcripts. It is thus unlikely that all of the putative transcription sites function in vivo, with the longest transcripts from each of the regions representing the functional messages. The primer extension analysis demonstrates that both of the transcripts are located close to the

putative −10 and −35 promoter elements.

ACKNOWLEDGMENTS

Gel electrophoresis and assignment of the AMP nucleosidase position on two-dimensional gels were kindly provided by Teresa Phillips in the laboratory of Dr. F. C. Neidhardt, University of Michigan. Matthew Davis provided technical assistance in development of the nearest-neighbor fragment analysis. We thank the DNA Synthesis Facility, Albert Einstein College of Medicine, for synthesizing oligonucleotides used as probes and DNA sequencing primers. We also thank Grace Sullivan for her excellent typing of the manuscript.

REFERENCES

Atassi, M. Z., & Habeeb, A. F. S. A. (1972) *Methods Enzymol. 25*, 546–553.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Smith, J. A., Seidman, J. G., & Struhl, K. (1989) *Current Protocols in Molecular Biology*, Greene Publishing and Wiley-Interscience, New York.

Bagdasarian, M., Lurz, R., Ruckert, B., Franklin, F. C. H., Bagdasarian, M. M., Frey, J., & Timmis, K. N. (1981) *Gene 16*, 237–247.

Barbacid, M. (1987) *Annu. Rev. Biochem. 56*, 779–827.

Bennetzen, J. L., & Hall, B. D. (1982) *J. Biol. Chem. 257*, 3026–3031.

Biggin, M. D., Gibson, T. J., & Hong, G. F. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3963–3965.

Chen, E., & Seeburg, P. H. (1985) *DNA 4*, 165–170.

Chou, P. Y., & Fasman, G. D. (1978) *Annu. Rev. Biochem. 47*, 251–276.

David, M., Tronchet, M., & Dênariê, J. L. (1981) *J. Bacteriol. 146*, 1154–1157.

de Boer, H. A., & Kastelein, R. A. (1986) *in Maximizing Gene Expression* (Reznikoff, W., & Gold, L., Eds.) pp 225–285, Butterworth Publishers, Boston, MA.

deCrombrugghe, B., Busby, S., & Buc, H. (1984) *Science 224*, 831–838.

Deutscher, M. P. (1985) *Cell 40*, 731–732.

Dickerson, R. E., & Timkovich, R. (1975) *Enzymes (3rd Ed.) 11*, 397–547.

Dunn, J. J., & Studier, F. W. (1980) *Nucleic Acids Res. 8*, 2119–2132.

Freier, S. M., Kierzek, R., Jaeger, J. A., Sugimoto, N., Caruthers, M. H., Neilson, T., & Turner, D. H. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 9373–9377.

Giranda, V. L., Berman, H. M., & Schramm, V. L. (1986) *J. Biol. Chem. 261*, 15307–15309.

Gold, L., & Stormo, G. (1987) in *Escherichia Coli and Salmonella Typhimurium* (Neidhardt, F. C., Ed.) Vol. 2, pp 1302–1307, American Society for Microbiology, Washington, DC.

Harley, C. B., & Reynolds, R. P. (1987) *Nucleic Acids Res. 15*, 2343–2361.

Hawley, D. K., & McClure, W. R. (1983) *Nucleic Acids Res. 11*, 2237–2255.

Hermodson, M. A., & Mahoney, W. C. (1983) *Methods Enzymol. 91*, 352–359.

Irani, M., Musso, R., & Adhya, S. (1989) *J. Bacteriol. 171*, 1623–1630.

Kohara, Y., Akiyama, K., & Isono, K. (1987) *Cell 50*, 495–508.

Konigsberg, W., & Godson, G. N. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 687–691.

Kyte, J., & Doolittle, R. F. (1982) *J. Mol. Biol. 157*, 105–132.

Leung, H. B., & Schramm, V. L. (1980) *J. Biol. Chem. 255*, 10867–10874.

Leung, H. B., & Schramm, V. L. (1981) *J. Biol. Chem. 256*, 12823–12829.

Leung, H. B., & Schramm, V. L. (1984) *J. Biol. Chem. 259*, 6972–6978.

Makino, K., Shinagawa, H., Amemura, M., & Nakata, A. (1986) *J. Mol. Biol. 190*, 37–44.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Marquetant, R., Sabina, R. L., & Holmes, E. W. (1989) *Biochemistry* (third of three papers in this issue).

Maxam, A. M., & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 560–564.

McClure, W. R. (1985) *Annu. Rev. Biochem. 54*, 171–204.

Meyer, S. L., Kvalnes-Krick, K. L., & Schramm, V. L. (1989) *Biochemistry* (second of three papers in this issue).

Miller, J. H. (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor, Laboratory, Cold Spring Harbor, NY.

Moller, W., & Amons, R. (1985) *FEBS Lett. 186*, 1–7.

Neidhardt, F. C., Bloch, P. L., & Smith, D. T. (1974) *J. Bacteriol. 119*, 736–747.

Neidhardt, F. C., Vaughn, V., Phillips, T. A., & Bloch, P. L. (1983) *Microbiol. Rev. 47*, 231–284.

Parkin, D. W., & Schramm, V. L. (1984) *J. Biol. Chem. 259*, 9418–9425.

Platt, T. (1985) in *Genetics of Bacteria* (Scaife, J., Leach, D., & Galizzi, A., Eds.) pp 255–278, Academic Press, London.

Platt, T. (1986) *Annu. Rev. Biochem. 55*, 339–372.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.

Shinagawa, H., Makino, K., & Nakata, A. (1983) *J. Mol. Biol. 168*, 477–488.

Shine, J., & Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 1342–1346.

Smith, D. H., Brutlag, D., Friedland, P., & Kedes, L. (1986) *Nucleic Acids Res. 14*, 17–20.

Tommassen, J., Koster, M., & Overduin, P. (1987) *J. Mol. Biol. 198*, 633–641.

Walker, J. E., Saraste, M., Runswick, M. J., & Goy, N. J. (1982) *EMBO J. 1*, 945–951.