JOURNAL OF BACTERIOLOGY, Feb. 1990, p. 538–547
0021-9193/90/020538-10$02.00/0

Vol. 172, No. 2

Analysis and Sequence of the *speB* Gene Encoding Agmatine Ureohydrolase, a Putrescine Biosynthetic Enzyme in *Escherichia coli*

M. B. W. SZUMANSKI[1] AND S. M. BOYLE[2]*

Department of Biochemistry and Nutrition, College of Agriculture and Life Sciences,[1] and Department of Pathobiology, Virginia-Maryland Regional College of Veterinary Medicine,[2] Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061-0442

Received 12 June 1989/Accepted 21 October 1989

The *speB* gene of *Escherichia coli* encodes the enzyme agmatine ureohydrolase (AUH). AUH catalyzes the hydrolysis of agmatine to urea and putrescine in one of the two polyamine biosynthetic pathways in *E. coli*. Sequencing of a 2.97-kilobase-pair fragment of the *E. coli* chromosome containing *speB* revealed the presence of three intact open reading frames (ORFs), ORF1 and ORF2 on one strand and ORF3 on the opposite strand, as well as a truncated ORF, ORF4, which terminated 92 kilobase pairs upstream from ORF3. ORF3 contained the coding sequence of the *speB* gene, as confirmed by complementation analysis. Two ORF3 transcripts were detected: a shorter transcript that included only ORF3 and a longer transcript that included both ORF3 and ORF4. The short transcript was abundantly expressed when the ORF4 sequences were deleted, but when ORF4 and its upstream sequences were present, the polycistronic message predominated and the amount of the monocistronic message was drastically reduced. The promoter from which the shorter transcript was produced contained a TATACT sequence at position −12, but sequences upstream from the −12 position seemed to be irrelevant for promoter activity. The predicted amino acid sequence of AUH contained three regions of high homology to the arginases of yeasts, rats, and humans.

In *Escherichia coli* putrescine is synthesized either by decarboxylation of ornithine or by decarboxylation of arginine to agmatine followed by hydrolysis of agmatine to putrescine and urea (14). The last two reactions are catalyzed by the enzymes arginine decarboxylase and agmatine ureohydrolase (AUH), respectively. The AUH protein has previously been purified from an *E. coli* isolate that was transformed with the plasmid pKA5; the enzymatic properties of AUH have been characterized previously (9). The subunit size of AUH, as deduced from its mobility on a sodium dodecyl sulfate-polyacrylamide gel, is 38 kilodaltons. The expression of AUH activity is antagonistically regulated by cyclic AMP and agmatine. In the presence of the cyclic AMP receptor protein, cyclic AMP represses the expression of the *speB* gene, while agmatine induces it. These two modulators appear to act independently from each other (10); the mechanism of this differential regulation is unknown. The *speB* gene coding for AUH is located at approximately 63.5 min on the *E. coli* chromosome. Although a large chromosomal fragment corresponding to this region, including the *speB* gene, is present in the pKA5 plasmid, the exact location of *speB* was not previously established. Here we report the nucleotide sequence of the *speB* structural gene and identify the promoter responsible for transcription of one of two mRNAs encoding AUH. We also report the mapping of the mRNA resulting from transcription that initiated at this promoter.

MATERIALS AND METHODS

Bacterial hosts, media, and growth conditions. *E. coli* CB806(Δ*lacZ galK phoA8 rpsL thi recA56*) (11) was used for all experiments involving the promoter vector pCB267 and its derivatives. *E. coli* DH5α [F⁻ *endA1 hsdR17 supE44 thi-1 recA1 gyrA96 relA1*, φ80d *lacZ*ΔM15 Δ(*lacZYA-argF*)*4196*] was purchased from Bethesda Research Laboratories, Inc. (Gaithersburg, Md.) and was used as a host for all other plasmids.

The bacteria were grown in either LB (6) or TB (15) broth. A portion of the cells from a frozen (−80°C) stock was inoculated into 3 ml of medium containing 100 μg of ampicillin per ml and grown overnight in a 37°C shaking water bath. The overnight culture (0.5 ml) was inoculated into fresh medium (50 ml) and grown to a density of about 90 Klett units. This culture was then used for either RNA purification or for the preparation of cell extracts.

Vectors. Plasmid pBR322 was purchased from Bethesda Research Laboratories, and plasmid pGEM-3Z was purchased from Promega Biotech Corp. The promoter cloning vector pCB267 (11) was obtained from T. Larson of the Department of Biochemistry and Nutrition (Virginia Polytechnic Institute and State University).

Purification of plasmid DNA. An overnight bacterial culture (50 ml) was centrifuged at 4,000 × *g* for 20 min at 4°C. The cell pellet was washed with and then suspended in 5 ml of ice-cold STE (100 mM NaCl, 20 mM Tris hydrochloride [pH 7.5], 10 mM EDTA). A total of 5 ml of phenol-chloroform (1:1) was added to the cell suspension and incubated for 20 min on a rotary shaker. The emulsified suspension was centrifuged at 1,200 × *g* for 12 min at room temperature; 2.5 volumes of ethanol were added to the aqueous phase, and the nucleic acid was precipitated at −20°C for 1 H. The precipitate was collected by centrifugation at 13,400 × *g* for 1 h; the pellet was redissolved in TE (10 mM Tris hydrochloride [pH 8.0], 1 mM EDTA), and the plasmid DNA was purified by CsCl banding (6).

Purification of total cellular RNA. The bacterial culture was treated as described above for the DNA purification procedure, up to the first ethanol precipitation of the nucleic

* Corresponding author.

acids, except that 0.375 μl of β-mercaptoethanol per ml of solution was added at each step. After the addition of the phenol-chloroform, the suspension was shaken for 30 s on a Vortex mixer (The Vortex Manufacturing Co., Cleveland, Ohio). The nucleic acid pellet was dissolved in 4 ml of standard saline citrate (1× SSC, which is 0.15 M NaCl plus 0.015 M sodium citrate)–1 mM HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid) and 1.5 μl of β-mercaptoethanol. A total of 2.5 ml of 7.5 M ammonium acetate was added, and the RNA was precipitated for 0.5 to 1 h at 0°C. After centrifugation at 13,400 × *g* for 1 h at 4°C, the RNA pellet was redissolved in 1× SSC–1 mM HEPES. Sodium dodecyl sulfate was added to 0.5%, and the RNA was stored at −20°C.

AUH assay. Cultures of 50 ml were grown to a density of about 90 Klett units and pelleted by centrifugation at 4,000 × *g* for 20 min at 4°C. The pellet was washed once with 5 ml and then suspended in 1 ml of ice-cold AUH reaction buffer (100 mM HEPES, 5 mM $MgCl_2$). The cells were disrupted by sonication with five 1-min exposures at 40% full scale by using a microtip probe (model 300 sonifier; Fisher Scientific Co., Pittsburgh, Pa.) and spaced with 1-min cooling on ice. Cell debris was removed by two successive 10-min centrifugations at 12,000 × *g* at 4°C. The supernatant constituted the cell extract that was used to assay the AUH activity. The extracts were preincubated for 3 min at 37°C; the reaction was initiated by the addition of agmatine sulfate to 25 mM. At 1-min intervals (over 5 min) 20-μl portions of the reaction mixture were transferred to 0.5 ml of solution of urease buffer reagent (no. 640-5; Sigma Chemical Co., St. Louis, Mo.) at 0°C; the AUH reaction was stopped by this treatment. The urea content in these fractions was measured by the diagnostics urea nitrogen determination procedure (no. 640; Sigma).

Phosphatase A assays. The phosphatase A assays were performed as described by Schneider and Beck (11).

Generation of recombinant clones. The recombinant clones were generated by standard protocols as described previously (4, 6, 13). The DNA modification enzymes were purchased from Bethesda Research Laboratories, Boehringer Mannheim Biochemicals (Indianapolis, Ind.), or Promega Biotech Corp.

Generation of the recombinant plasmids pKB1, pKB5, pKB2, pKB2S, pKB2H pKG1, and pKG2. The inserts of the deletion plasmids pKB1, pKB5, pKB2, pKB2S, and pKB2H are shown schematically in Fig. 1B. The indicated restriction fragments of plasmid pKA5 were excised and inserted into the pBR322 vector to produce plasmids pKB1, pKB5, pKB2, pKB2S, and pKB2H. The insert of plasmid pKB2S was subcloned into the vector pGem-3Z, producing plasmid pKG2 (described in the legend to Fig. 1). This subcloning was performed in two steps. The intermediate plasmid, pKG1, contained the *Sma*I-*Bam*HI fragment of the pKA5 insert.

Generation of unidirectional deletions in the pKG2 plasmid. Unidirectional deletions were made in the pKG2 insert in both directions by the *Exo*III-*Exo*VII method (7). The derivative plasmids are shown in Fig. 1C.

Plasmids pKG3, pKG4, and pB15N. The 2.95-kilobase-pair (kb) *Pst*I fragment of plasmid pKA5 was subcloned into the *Pst*I site of the pGEM-3Z vector, producing plasmid pKG3. In Fig. 1C, only the *Ava*I-*Pst*I fragment of plasmid pKG3 is shown. The 2.4-kb *Eco*RV-*Pst*I fragment of plasmid pKG3 was ligated into the pGEM-3Z vector that was cleaved with *Sma*I and *Pst*I (both *Eco*RV and *Sma*I produce blunt, mutually compatible ends), producing plasmid pKG4. Plasmid pB15N was constructed by deletion of the *Sma*I-*Nru*I fragment from plasmid B15 (which is depicted in Fig. 1C).

Plasmid pBB15N. In all the sequencing deletion clones designated B, a *Hin*dIII restriction site, derived from the pGEM-3Z vector, flanks the junction between the deleted end of the insert and the vector. Plasmid pBB15N was constructed in two steps. In the first step, the *Hin*dIII-*Nru*I fragment, containing the open reading frame (ORF) ORF3, was cloned into pBR322, which was cleaved with *Hin*dIII and *Nru*I; this resulted in substitution of approximately two thirds of the 5′ end of the *tetR* gene of the vector with the insert. The *Eco*RI-*Hin*dIII fragment of the resulting intermediate plasmid was deleted to remove the remaining portion of the *tetR* promoter.

Construction of plasmids pCO3B15P and pCO3B32P. Plasmids pCO3B15P and pCO3B32P contained the deletion proximal fragments of plasmids B15 and B32, respectively. The insert in plasmid B15 is shown in Fig. 1C. The insert in plasmid B32 was deleted for an additional 36 base pairs (bp). In both plasmids a *Hin*dIII site was located in the vector immediately adjacent to the right insert-vector border. The exact location of this junction is shown in Fig. 2E. Plasmids B15 and B32 were digested with *Hin*dIII and *Bam*HI. The approximately 850-bp fragments were ligated separately into the *Hin*dIII-*Bam*HI sites of the vector pCB267 (11), upstream from the promoterless *phoA* gene.

Sequencing. Sequencing was performed by the dideoxy method of Sanger et al. (8) by using the Sequenase protocol and reagents purchased from U.S. Biochemical Corp. Figure 1C shows the extent of sequence data that was derived from each of the deletion plasmids.

Primer extension and S1 nuclease assays. The primer extension and the S1 nuclease assays were performed as described previously (16). For the 5′-end mapping of the *speB* transcript, a 20-nucleotide-long synthetic oligonucleotide complementary to the 5′ region of the *speB* gene was used as a primer. The primer was synthesized by the Department of Microbiology and Immunology (Virginia Commonwealth University, Richmond, Va.). The exact location of the region to which the primer was complementary is shown in Fig. 2E. For use in the S1 nuclease assay the primer was extended on the B15 template; the primer was annealed to the B15 template as described in the Sequenase protocol, and the primer-template mixture was treated as in the labeling reaction, except for following modifications. A mixture of 0.5 mM each of dCTP, dGTP, and TTP was substituted for the Sequenase labeling mixture; 5 μl of [^{35}S]dATP (1,250 Ci/mmol, 12.5 mCi/ml, 10.4 mmol/ml), which was purchased from E. I. du Pont de Nemours & Co., Inc. (Wilmington, Del.), was used in each reaction; the reaction mixture was incubated at 37°C for 10 min, followed by the addition of 2 μl of 0.5 mM solution of all four deoxynucleoside triphosphates and an additional 5 min of incubation at 37°C.

For use in the S1 nuclease mapping of the 3′ end of the *speB* transcript, the T7 universal primer of the GemSeq K/RT sequencing system, purchased from Promega Biotech Corp., was annealed to the *Bgl*II-digested pB15N template and extended as described above.

Northern hybridization. Total cellular RNA from strains DH5α(pKA5) and DH5α(pBB15N) was electrophoresed in duplicate on an agarose-formaldehyde gel. Electrophoresis, Northern transfer, and hybridizations were performed as described by Selden (12). The *speA*-specific probe was complementary to nucleotides 462 to 605 of the *speB* gene, and the ORF4-specific probe was complementary to se-

quences 289 to 545 upstream from the 3′ end of ORF4. The probes were constructed by extension of the T7 universal primer (as described above for S1 nuclease mapping) by using *Pst*I-digested plasmids A53 and A42 (Fig. 1C) as templates, respectively.

RESULTS

Location of the sequences necessary for AUH expression. Plasmid pKA5 contains a 7.5-kb insert, which was derived from an *E. coli* genomic library (2), cloned into the *Eco*RI site of plasmid pBR322. This insert was shown to contain the genes *speB*, *speA*, and *metK*, encoding AUH, arginine decarboxylase, and methionine adenosyltransferase, respectively (1). We updated the restriction map of plasmid pKA5 (Fig. 1A). In order to localize the restriction fragment that contains the intact *speB* gene, we constructed a series of deletions and assayed the deletion clones for the AUH activity (Fig. 1B). The clones bearing plasmids pKB1 and pKB5 were negative for AUH activity. When the inserts from these plasmids were reconstituted in the original order in plasmid pKB2, AUH activity was restored, indicating that the internal *Bam*HI site interrupts the *speB* gene. Further deletions in the rightward direction from the *Eco*RI site revealed that the 460-bp *Sma*I-*Hin*dIII fragment is necessary for *speB* expression. Previous results indicated that removal of sequences to the left or to the right of the unique *Bal*I site also inactivates the *speB* gene (C. Satishchandran, Ph.D. thesis, Memorial University of Newfoundland, St. John's, Newfoundland, Canada, 1985). Therefore, we concluded that the shortest available restriction fragment that contains an intact *speB* gene is the one present in plasmid pKB2S.

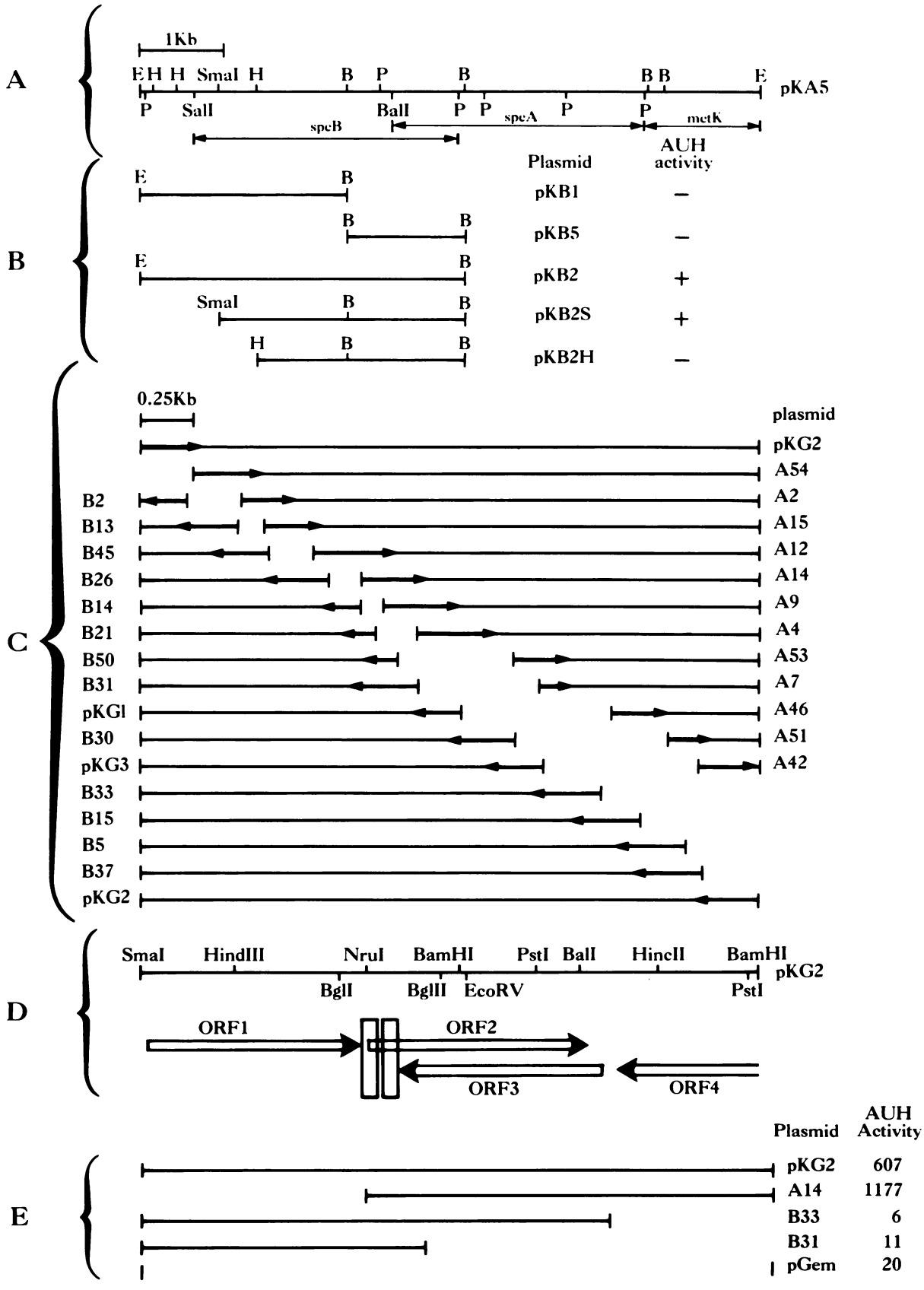
Nucleotide sequence of the pKB2S insert. We subcloned the insert of plasmid pKB2S into the vector pGEM-3Z, creating plasmid pKG2. A chance existed that during the two-step cloning procedure we may have left behind a small, unnoticed *Bam*HI fragment. To ensure that this was not the case, we later subcloned the fragment *Pst*I-*Eco*RV from plasmid pKA5 into the pGEM-3Z vector (plasmid pKG4) and sequenced it through the intact *Bam*HI site, thus validating the reconstitution. Figure 1C shows the strategy that was used for sequencing the 2.97-kb insert of plasmid pKG2. Sequence analysis (Fig. 1D) revealed the presence of three complete ORFs (ORF1, ORF2, and ORF3) and the 3′ end of another ORF (ORF4) that terminated 92 nucleotides upstream from ORF3. ORF1 and ORF2 had the same orientation and were separated by 31 nucleotides. ORF2 and ORF3 were divergently oriented and overlapped by 969 nucleotides. A sequence of two 82-bp-long imperfect tandem repeats, with 70 bp of homology, extended between the diverging 3′ ends of ORF1 and ORF3 and overlapped the 5′ end of ORF2. The two repeats were 31 bp apart; the region of repeats was very GC-rich and contained four palindromic sequences, three of which overlapped (see Fig. 4). These palindromes are strongly homologous to the repetitive extragenic palindromic sequences (3).

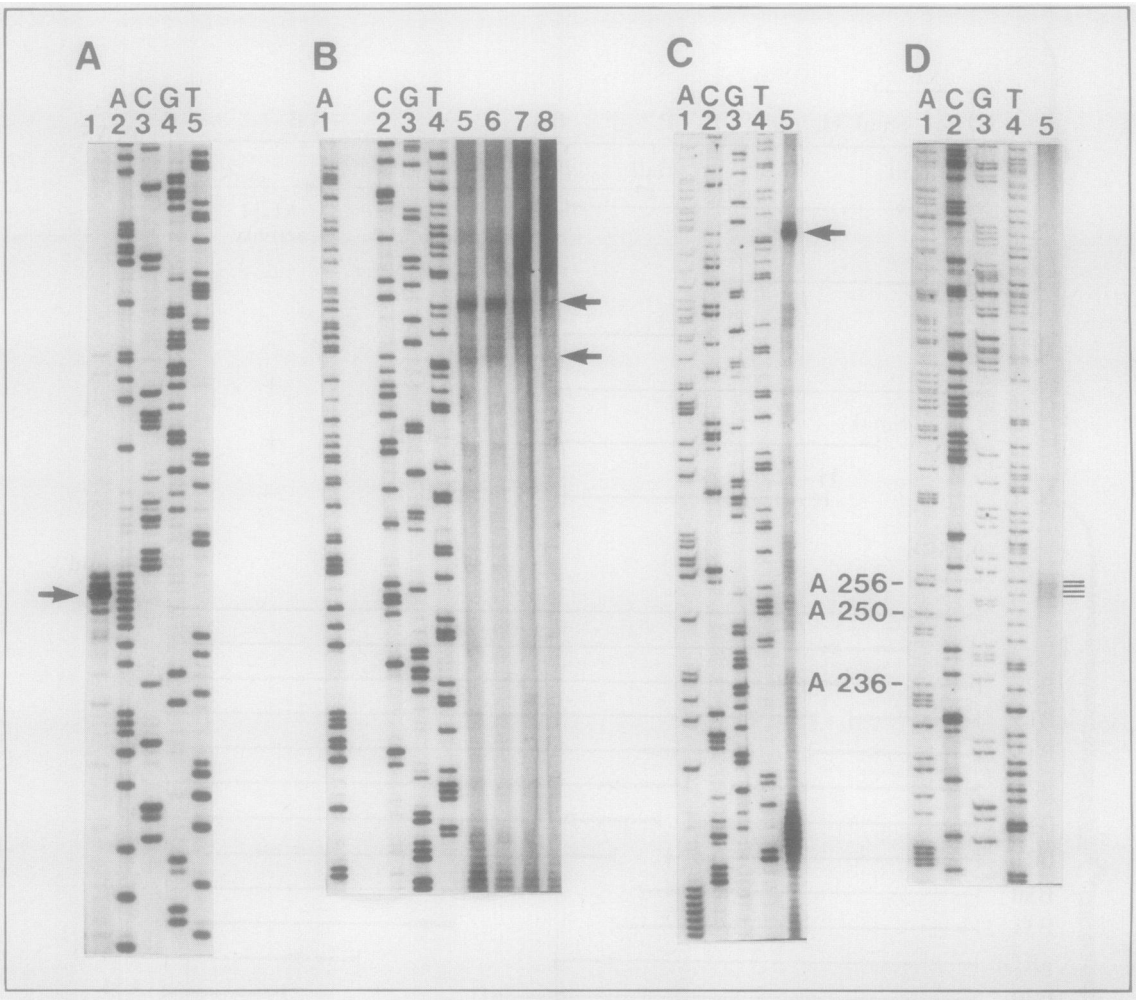
A 37-nucleotide-long palindromic structure was located at the 5′ end of ORF3. It had the potential of forming a loop with a 15-nucleotide-long stem, including one mismatch and a 7-nucleotide-long "bubble." The palindrome overlapped the start of ORF3 by 15 nucleotides and was terminated in the ORF3 direction by five T residues.

Identification of the *speB* gene. To establish which of the three ORFs encoded AUH, we assayed three of the deletion clones for AUH activity (Fig. 1E). Only clone A14, containing ORF2 and ORF3, overexpressed AUH activity. Since neither clone B31, containing ORF1, nor clone B33, containing ORF1 and ORF2, exhibited increased AUH activity, ORF3 must contain the coding sequence of the *speB* gene. We also performed S1 nuclease analysis on the mRNA that was purified from the A14 clone using a primer complementary to the sequences upstream from ORF3 (data not shown). This experiment revealed that most of the *speB* transcripts are initiated at a site 92 bp beyond the insert-vector junction. This nucleotide is the transcription start point for the α peptide of β-galactosidase in pGEM-3Z. Thus, the *speB* gene in the A14 clone is transcribed from the *lacZ* promoter of the vector. At 31 bp downstream from the start of ORF3 was a Shine-Dalgarno consensus sequence that was followed 7 bp downstream by a methionine codon (Fig. 2E; see Fig. 4). Both of these sites were deleted in clone B33 (Fig. 2E), which did not express AUH activity. Together with the observation that *speB* expression in the A14 clone was directed from the *lacZ* promoter of the vector, this constituted a strong indication that the methionine codon mentioned above must be the initiation codon for *speB* translation.

Mapping of the 5′ end of the *speB* transcript and the location of the *speB* promoter. We used the primer extension and S1 nuclease methods to identify the transcription start of the *speB* gene. By the primer extension method (Fig. 2A), a major band appeared 1 nucleotide prior to the start of the 37-nucleotide-long palindromic sequence (Fig. 2E; see Fig. 4). The S1 nuclease assays, which were performed on the total cellular RNAs derived from clones bearing plasmids A14 (data not shown) or pKA5 or pBB15N (Fig. 2A, B, and C) revealed strong bands at positions corresponding to the 1-nucleotide mismatch in the stem. These results strongly indicate the presence of a cruciform structure within the DNA-RNA hybrid prior to the S1 nuclease digestion. Upstream from this sequence two bands were visible on the RNA derived from the pKA5-bearing clone: a strong band at position 113 bp and a very weak band at position 102 bp upstream from the start of ORF3. The position 113 bp upstream from ORF3 marked the insert-vector junction of plasmid B15 (see Fig. 2E for the exact limit of the B15 deletion clone), which was the template for the extension of the primer used in this S1 nuclease analysis. Thus, the sequences of the primer that extended beyond this point were complementary to the template but not to the upstream pKA5 sequences. Therefore, a band at this position signifies an mRNA species that is initiated upstream from this point.

FIG. 1. Mapping of the location of the *speB* gene. (A) Restriction map of the *E. coli* chromosome fragment present in the plasmid pKA5. Abbreviations: E, *Eco*RI; H, *Hin*dIII; P, *Pst*I; B, *Bam*HI. The approximate positions of the *speB*, *speA*, and *metK* genes are indicated. (B) Subclones of the pKA5 insert used to determine the location of the *speB* gene. The ability of the plasmids to express AUH activity was determined by enzyme assays; the vector was pBR322. (C) Deletion clones used for sequencing of the pKG2 insert. The short vertical lines represent the junction with the vector (pGEM-3Z). Arrows represent the extent and direction of sequencing. (D) Top, Restriction sites deduced from the sequence data; bottom, location of the ORFs, as deduced from the sequencing data. Boxes represent the region of long tandem repeats. (E) Ability of the deletion plasmids from panel C to express AUH activity in *E. coli* DH5α. The AUH specific activity is expressed in nanomoles per minute per milligram of protein.


A
1Kb
E H H SmaI H B P B B B E
pKA5
P SalI BalI P P speA P P metK
speB
AUH
Plasmid activity
B
E B pKB1 –
B B pKB5 –
E B pKB2 +
SmaI B B pKB2S +
H B B pKB2H –
C
0.25Kb
plasmid
pKG2
A54
B2 A2
B13 A15
B45 A12
B26 A14
B14 A9
B21 A4
B50 A53
B31 A7
pKG1 A46
B30 A51
pKG3 A42
B33
B15
B5
B37
pKG2
D
SmaI HindIII NruI BamHI PstI BalI HincII BamHI
pKG2
BglI BglII EcoRV PstI
ORF1
ORF2
ORF3
ORF4
Plasmid AUH Activity
E
pKG2 607
A14 1177
B33 6
B31 11
pGem 20



E

228

3'-CCTTCAAGTAGCCGTCCGCATTTTTTGGTTTCCGTAACCTTTGGTCA

B33

CTCAATAGCATAACTACTGGATTCCACGAGTACCTACCTGAGGATA

ATGTTTCAGCTATATTTTTTTGGGAAGGCGCAGCAATCCGGGCTGC

ORF3

TCCTTCCCAACCTAAACAGTGTTATTTAACACCGCCTAATAGTGGC

GATTTAATTCGCCTAAGTTCATTGTGTCCTGAATGAGTAGAAGTTC

B32

ORF4

TATTCATATTGGCATGTTTGGACGGAGCTTGAGAAGTTCCTTGACA

B15

ACAACGTCAAGTCGTAG-5'

-65

F

plasmid	Phosphatase A specific activity	plasmid	AUH specific activity
	(u)		(u/mg)
pCO3B15P	5.970	pBB15N	478
pCO3B32P	0.281	pKA5	924
pCB267	0.054	pBR322	11

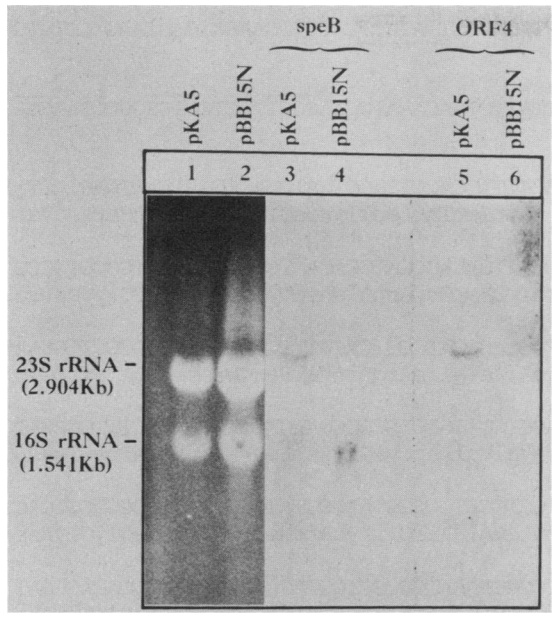


FIG. 3. Northern hybridization of RNA derived from *E. coli* DH5α bearing plasmid pKA5 or pBB15N. Lanes 1 and 2, Ethidium bromide-stained gel before transfer; lanes 3 through 6, autoradiogram of the hybridized membranes. The plasmid that was used is indicated above the lane, and the probe is indicated above the bracket.

When the S1 nuclease analysis was repeated with a primer that was extended on a template in which homology to the transcript ended at position 160 bp upstream from the start of ORF3 (data not shown), two bands still appeared on the autoradiogram; the lower band was at the same position as described above, but the higher band was at a position 160 bp upstream of ORF3. Thus, the longer *speB* transcript must be initiated more then 160 bp upstream from ORF3. The faint lower band, which was visible in this experiment, signified a transcript starting 102 bp upstream from ORF3. Twelve bp upstream from the nucleotide equivalent to this band was a TATACT sequence which was strongly homologous to the *E. coli* −10 TATAAT promoter consensus sequence. Since this sequence marked the end of the B15 insert (Fig. 2E), we used this deletion to map the location of the promoter initiating the shorter *speB* transcript. We isolated the *speB* gene fragment, which was present in the deletion plasmid B15, by subcloning it into the pBR322 vector. The strain harboring the resulting plasmid pBB15N exhibited elevated AUH activity (Fig. 2F). When mRNA derived from this clone was used in S1 nuclease analysis (Fig. 2C), a strong band appeared at the same location as the weak lower band from the analysis of the pKA5 RNA (Fig. 2B), while the upper band disappeared. These results confirm that two *speB* transcripts are present and suggest that the shorter transcript is initiated from a promoter that does not require any specific sequences upstream from position −12 for activation. In our construction of plasmid pBB15N, we removed the entire sequence containing the *tetR* promoter of the vector to avoid any possibility of producing a substitution for a putative −35 region necessary for *speB* expression. Still, a small chance existed that the presence of some unidentified sequences upstream from the vector-insert junction might have resulted in a coincidental complementation of the truncated *speB* promoter in the insert. To exclude this possibility, we constructed plasmid pCO3B15P. This plasmid contained the upstream fragment of the *speB* gene that was identical to the one present in pBB15N, which was joined in a transcriptional fusion to the promoterless *phoA* gene within the pCB267 vector. Expression of the *phoA* gene was activated in this construct (Fig. 2F). When an additional 36 bp was deleted, as in plasmid pCO3B32P, which represented a derivative of deletion clone B32, the promoter activity was drastically reduced.

Establishing the origin of the longer *speB* transcript. To establish whether the longer *speB* transcript coded both for AUH and for ORF4, the RNAs from pKA5- and pBB15N-bearing strains were hybridized to a probe specific for either *speB* or ORF4 (Fig. 3). The autoradiograph revealed that a *speB* probe hybridized only to a 3.2-kb band in the RNA from pKA5 and to a 1.5-kb band in the RNA from pBB15N. The sizes of these transcripts conformed to the molecular weights for the polycistronic and monostronic messages expected from our RNA mapping and ORF length analyses, respectively. This confirmed the observation from the S1 nuclease analyses (Fig. 2B and C) that the shorter transcript is produced by a strain bearing pKA5 in trace amounts compared with the amount of the longer transcript that is produced in this strain; in contrast, the strain harboring pBB15N produced comparatively high quantities of the shorter transcript. On the membrane that was hybridized to the ORF4 probe, the band representing the longer transcript was visible in pKA5-transformed cells, while the shorter transcript was no longer detected in pBB15N-transformed

FIG. 2. Mapping of the *speB* transcript and the *speB* promoter. All sequencing reactions and the 5′-end mapping experiments (A, B, and C) were performed with the primer depicted in panel E. (A) Primer extension mapping of the 5′ end of the *speB* transcript. Lane 1, RNA derived from clone A14; lanes 2 through 5, sequencing reactions on the A14 template. (B) S1 nuclease assay on RNA derived from clone pKA5. Lanes 1 through 4, sequencing reactions on B15 template; lanes 5 through 8, S1 nuclease assays on 250, 200, 150, and 100 μg of RNA, respectively. Hybridizations were performed at 39°C. (C) S1 nuclease assay on RNA derived form the pBB15N clone. Lanes 1 through 4 are as described for panel B; lane 5, S1 nuclease assay on 200 μg of RNA. Hybridization was performed at 45°C. (D) S1 nuclease mapping of the 3′ end of the *speB* transcript. Lanes 1 through 4, Sequencing reactions on plasmid B32; lane 5, S1 nuclease assay on 200 mg of pBB15N RNA. The hybridization was performed at 50°C. Arrows and short lines point to the indicated bands. The A followed by a number represents the size of DNA at the indicated band in the A sequencing lane. (E) Upstream sequences of the coding strand of the *speB* gene; the orientation is as in Fig. 1D. The bracket over the first 20 nucleotides indicates the region of complementarity to the oligonucleotide primer used in the 5′-mapping experiments (panels A, B, and C). Small arrows with clone names to the right indicate the end of the inserts in the named deletion plasmids. The methionine initiation codon as well as the sequences with homology to the ribosome-binding site and the −10 promoter consensus are boxed. The asterisk indicates the center of symmetry in the palindromic sequence that is marked by an arrow over the participating nucleotides. The start of ORF3 (*speB*) and the end of ORF4 are indicated. The leader sequence that precedes ORF3 is underlined. (F) Ability of the DNA fragments derived from the deletion clones indicated in panel E to promote transcription of the promoterless *phoA* gene or to express AUH activity. One unit of phosphatase A specific activity is the optical density at 410 nm (10^3) per volume of culture (in milliliters) times time (in minutes) times cell density (Klett units). One unit of AUH activity is 1 ng/min.

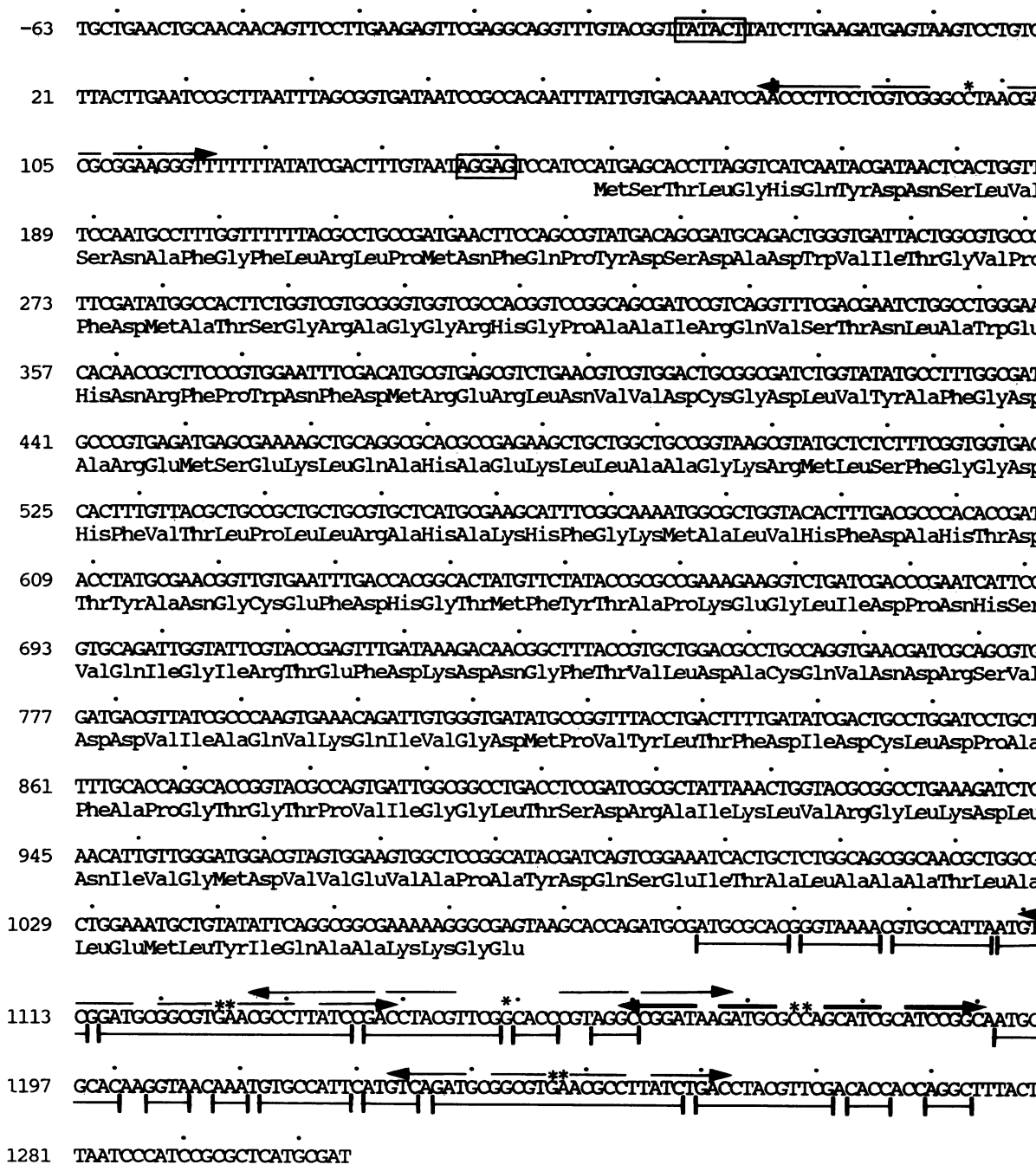


FIG. 4. Nucleotide and amino acid sequence of the *speB* gene. The amino acid sequence of the *speB* gene product, starting with the initiating methionine codon, as deduced from the sequence, is given below the codons. The promoter consensus sequence and the ribosome-binding site are boxed. The start of transcription is assigned the number 1. The homologous nucleotides within the long tandem repeat sequence are underlined. The palindromic sequences are indicated by arrows with asterisks at the center of symmetry. The bold arrow indicates the palindrome that terminates the *speB* transcript.

cells. These results indicate strongly that the longer transcript is a polycistronic message encoding both the *speB* and the ORF4 sequences. Additionally, since no other RNA cross-reacted with the ORF4 probe, this longer transcript must be the major product of the genes represented by ORF4 and *speB*.

Mapping of the 3′ end of the *speB* transcript. The 3′ end of the *speB* transcript was mapped by S1 nuclease analysis of the total cellular RNA derived from clone pBB15N. The primer was homologous to the sequences flanking and overlapping the 3′ end of ORF3 (nucleotides 1304 to 939, as depicted in Fig. 4). Four bands of equal intensity appeared on the autoradiogram (Fig. 2D); they were positioned at a distance of 254 to 257 nucleotides from the 3′ end of the primer. This corresponds to positions 1193 to 1196 of the *speB* transcript (Fig. 4). This position marks the sequences

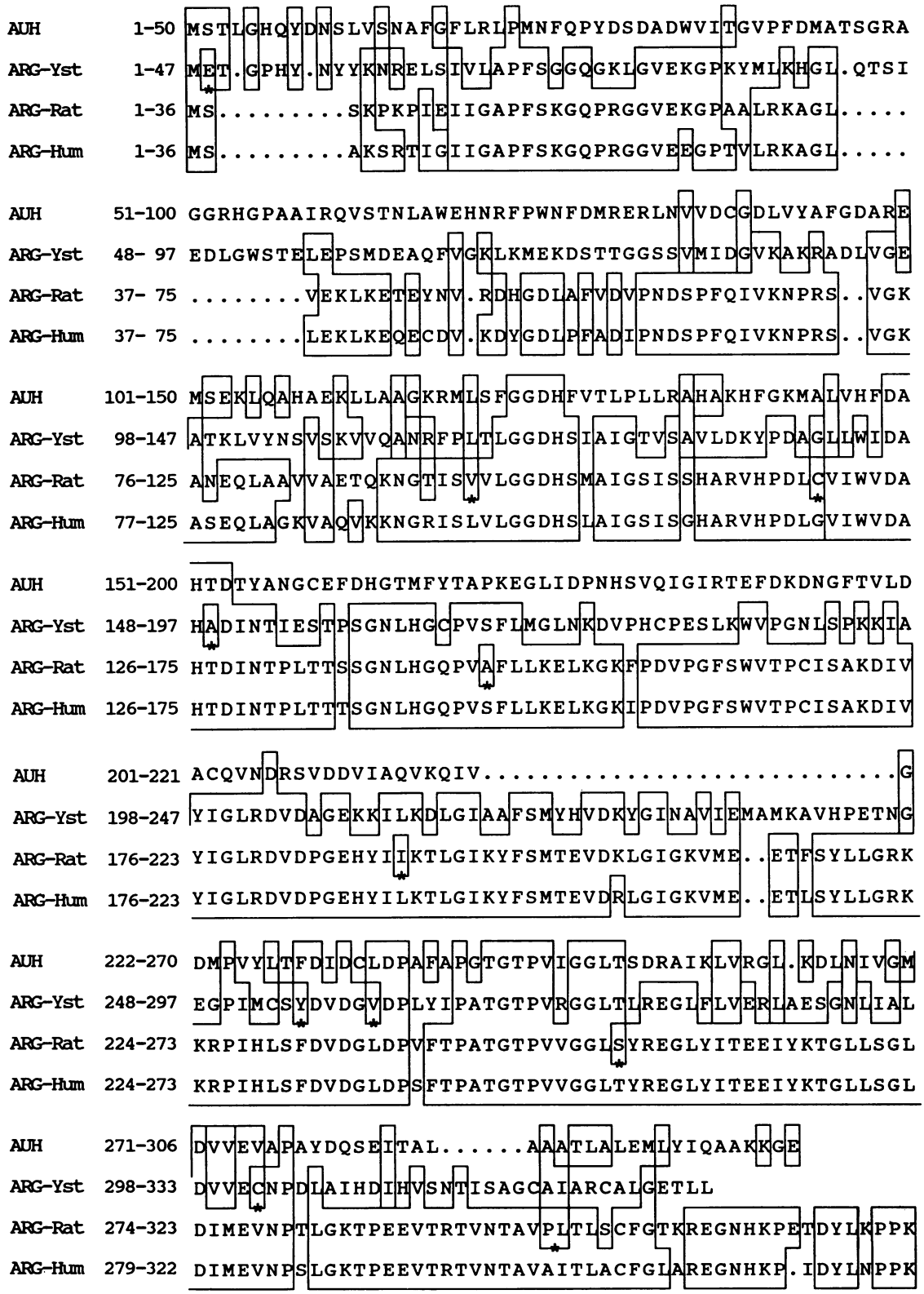


FIG. 5. Comparison of amino acid sequence of AUH with those of yeast (ARG-Yst), rat (ARG-Rat), and human (ARG-Hum) arginases. The amino acid sequences of these eucaryotic arginases are from Haraguchi et al. (5). Gaps were introduced to increase the similarities; the matching amino acids are boxed; in some cases, nonmatching amino acids are boxed and indicated with asterisks.

directly downstream from the last of the three overlapping repetitive extragenic palindromes present within the direct repeat region (Fig. 1D and 4). Thus, this palindrome is the transcription termination signal in the transcription of *speB* mRNA.

Sequence comparison of AUH and arginases. A comparison of the nucleotide sequence of AUH with those in the EMBL and GenBank data bases revealed that AUH has a predicted amino acid sequence which is similar to the amino acid sequences of the arginases from yeasts, rats, and humans (5). The amino acid sequence of AUH was aligned with these three eucaryotic arginases (Fig. 5). To increase the similarity to these previously aligned arginases (5), three gaps were introduced in AUH; and two additional gaps were introduced in the rat, yeast, and human arginases. There were three highly homologous regions corresponding to residues 118 to 126, 149 to 156, and 238 to 251 which were approximately 50, 63, and 78% identical to all three of the arginases, respectively. Overall, there was approximately 14% homology of AUH to the three eucaryotic arginases (47 of 333 residues or 48 of 322 residues in the gapped alignment).

DISCUSSION

Within the 7.5-kb insert of plasmid pKA5 we identified and sequenced the DNA necessary for the expression of AUH activity. Sequence analysis revealed the presence of three ORFs: ORF1, ORF2, and ORF3. ORF1 and ORF2 were arranged in tandem and were separated by 31 nucleotides, while ORF3 was on the opposite strand and overlapped ORF2 by 864 nucleotides. Among deletion clones harboring the individual ORFs, only the clone with ORF3 overexpressed AUH activity. Thus, ORF3 represents the coding region of the *speB* gene. The molecular mass of AUH, deduced from its sequence, was 33,409 daltons. The molecular mass previously established from the mobility of the purified enzyme on a sodium dodecyl sulfate-polyacrylamide gel was 38 kilodaltons (9). We do not know the reason for the 4.6-kilodalton discrepancy.

It is worth noting that the higher regions of homology between AUH and the arginases were in the central and carboxy regions (e.g., residues 238 to 251). This may have been related to conservation of the active sites in the two types of enzymes, since both produce urea as one of their end products. Because of the 40% homology at either the nucleotide or the amino acid level among the eucaryotic arginases, it has been suggested that they have a common origin (5). The homology of AUH to these arginases leaves open the possibility that AUH represents that common origin.

Two imperfect (86% identity) tandem repeats of 82 and 72 nucleotides were located between the 3′ ends of the converging ORF1 and ORF3 sequences (Fig. 1D). Four palindromes, which were strongly homologous to repetitive extragenic palindromic sequences, were present within this region; three of them overlapped (Fig. 4). We showed that the *speB* transcript terminates at the end of the third overlapping palindrome distal from the *speB* gene.

S1 nuclease mapping of the 5′ end of the *speB* transcript revealed that two species of mRNA are involved in the synthesis of the AUH protein. The start point and the endpoint of the shorter transcript, as well as the location of the promoter from which it was initiated, were mapped. An unusual feature of this promoter was that while it contained a sequence, TATACT, at position −12 which differed by only 1 nucleotide from the TATAAT −10 consensus sequence for the σ^{70}-recognized promoters, there were no upstream sequences resembling the −35 consensus sequence (TTGACA). Furthermore, *speB* promoter activity was retained in two fusion plasmids (pBB15N and pCO3B15P), in which the region upstream from position −12 was substituted by DNA sequences (derived from two different sources) which differed from each other and from the native sequence and contained no identifiable −35 consensus sequence. These results indicate that sequences upstream from the Pribnow box are irrelevant for the ability of the RNA polymerase to bind to and initiate transcription from the *speB* promoter. In contrast, deletion of the −12 consensus sequence (in plasmid pCO3B32P) abolished *speB* promoter activity.

The presence of a longer transcript containing the *speB* message was evident from the results of the S1 nuclease mapping experiments (Fig. 2B and C). Since the promoter of the *speB* gene overlapped the sequence of ORF4 (which represented the *speA* gene, coding for arginine decarboxylase; Robert Moore, Virginia Polytechnic Institute and State University, personal communication), this longer transcript could originate either from a second promoter within the *speA* gene or from the *speA* promoter itself. The results of the Northern hybridization experiment (Fig. 3) confirmed the latter hypothesis. Thus, AUH in *E. coli* is encoded by one ORF but is synthesized from two transcripts: one monocistronic, which is directed from the *speB* promoter, and one polycistronic, which is directed from the (*speA*) ORF4 promoter. From the relative intensities of the bands representing these two transcripts in the S1 nuclease analyses and in the Northern hybridization experiment, the transcription of the monocistronic *speB* message appears to be repressed when the polycistronic transcript is produced. The cellular environment in strain DH5α(pKA5), in which *speB* was expressed primarily as a polycistronic message, differed only from the one in strain DH5α(pBB15N), in which *speB* was expressed as a monocistronic message, by the presence of increased amounts of proteins (or RNA) coded within the pKA5 insert. Consequently, the switch between polycistronic and monocistronic expression must be mediated by products of genes flanking *speB* rather than by an intervention of a gene product encoded elsewhere on the chromosome.

In the course of our analysis of the 5′ end of the *speB* transcript, we discovered that a 37-bp, GC-rich palindromic sequence present at the start of ORF3 forms a loop in both the RNA and the DNA. This structure has a 15-bp stem (including one bubble caused by a 1-bp mismatch) and a 7-bp single-stranded loop. This cruciform structure that was formed within the RNA-DNA heteroduplex survived at the highest stringency of prehybridization conditions used in our S1 nuclease analyses. It seems likely, therefore, that it might also form in vivo within RNA, DNA, or both. The palindrome ended with a track of seven T residues in the *speA* (ORF4)-coding direction, resembling a *rho*-independent terminator structure. However, it did not stop the transcription from the *speB* promoter initiated 79 bp upstream, nor did it prevent the readthrough from *speA* into *speB*, resulting in the polycistronic transcript. Nevertheless, it is possible that this structure might be involved in the regulation of *speB* gene expression.

ACKNOWLEDGMENTS

We thank Timothy Larson for *E. coli* CB806 and plasmid pCB267, John Johnson for contributing ideas that were useful toward the development of our RNA purification method and S1 nuclease

assay, Bob Moore and Dennis Dean for helpful discussions, and Lisa Barroso for performing the primer extension assay on the 5′ end of the *speB* transcript.

This work was supported by grant DMB-8508917 from the National Science Foundation (to S.M.B.) and by a graduate fellowship from the Department of Biochemistry and Nutrition of the Virginia Polytechnic Institute and State University (to M.B.W.S.).

LITERATURE CITED

1. **Boyle, S. M., G. D. Markham, E. W. Hafner, J. M. Wright, H. Tabor, and C. W. Tabor.** 1984. Expression of the cloned genes encoding the putrescine biosynthetic enzymes and methionine adenosyltransferase of *Escherichia coli* (*speA*, *speB*, *speC* and *metK*). Gene **30**:129–136.
2. **Clarke, L., and J. Carbon.** 1976. A colony bank containing synthetic Col E1 hybrid plasmids representative of the entire *E. coli* genome. Cell **9**:91–99.
3. **Gilson, E., J.-M. Clement, D. Brutlag, and M. Hofnung.** 1984. A family of dispersed repetitive extragenic palindromic DNA sequences in *E. coli*. EMBO J. **3**:1417–1421.
4. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166**:557–580.
5. **Haraguchi, Y., M. Takiguchi, Y. Amaya, and S. Kawamoto.** 1987. Molecular cloning and nucleotide sequence of cDNA for human liver arginase. Proc. Natl. Acad. Sci. USA **84**:412–415.
6. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
7. **Ozkaynak, E., and S. D. Putney.** 1987. A unidirectional deletion technique for the generation of clones for sequencing. BioTechniques **5**:770–773.
8. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
9. **Satishchandran, C., and S. M. Boyle.** 1986. Purification and properties of agmatine ureohydrolase, a putrescine biosynthetic enzyme in *Escherichia coli*. J. Bacteriol. **165**:843–848.
10. **Satishchandran, C., and S. M. Boyle.** 1984. Antagonistic transcriptional regulation of putrescine biosynthetic enzyme agmatine ureohydrolase by cyclic AMP and agmatine in *Escherichia coli*. J. Bacteriol. **157**:552–559.
11. **Schneider, K., and C. F. Beck.** 1987. New expression vectors for identifying and testing signal structures for initiation and termination of transcription. Methods Enzymol. **153**:452–461.
12. **Selden, R. F.** 1987. Analysis of RNA by Northern hybridization, p. 491–498. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current protocols in molecular biology. John Wiley & Sons, Inc., New York.
13. **Struhl, K.** 1986. A rapid method for creating recombinant DNA molecules. BioTechniques **3**:452–453.
14. **Tabor, C. W., and H. Tabor.** 1985. Polyamines in microorganisms. Microbiol. Rev. **49**:81–99.
15. **Tartof, K. D., and C. A. Hobbs.** 1987. Improved media for growing plasmid and cosmid clones. Focus **9**:12.
16. **Wang, S.-Z., J.-S. Chen, and J. L. Johnson.** 1988. The presence of five *nifH*-like sequences in *Clostridium pasteurianum*: sequence divergence and transcription properties. Nucleic Acids Res. **16**:439–454.