Nucleotide Sequence of the asd Gene of *Streptococcus mutans*

IDENTIFICATION OF THE PROMOTER REGION AND EVIDENCE FOR ATTENUATOR-LIKE SEQUENCES PRECEDING THE STRUCTURAL GENE

(Received for publication, October 1, 1986)

Guy A. Cardineau† and Roy Curtiss III‡

From the Department of Microbiology and Institute of Dental Research, University of Alabama in Birmingham, Birmingham, Alabama 35294

The complete nucleotide sequence of the *asd* gene of *Streptococcus mutans* encoding aspartate β-semialdehyde dehydrogenase (EC 1.2.1.11), an enzyme comprised of 357 amino acids, having an M₀ of 38,897 and active in the biosynthetic pathway of lysine, threonine, methionine, dianaminopimelic acid, and isoleucine, has been determined. In addition we report the 276 nucleotides upstream of the structural gene which contain a highly efficient promoter identified by both RNA polymerase binding and *in vitro* transcription analysis. A leader transcript which terminates at a fixed point immediately preceding the *asd* promoter region was identified in the DNA sequence and confirmed by *in vitro* transcription analysis as well. The close proximity of this transcript and its ρ-independent transcriptional terminator to the *asd* coding sequence suggests involvement in a mechanism of regulation. Message stability experiments indicate the half-life of *asd* specific messages to be comparable to that of *Escherichia coli* messages. Conditions of varying concentrations of lysine, threonine, and methionine exert no apparent control over expression of the *S. mutans* *asd* gene in *Escherichia coli* suggesting the requirement of an accessory regulatory element specific for the *S. mutans asd* gene.

Aspartate β-semialdehyde dehydrogenase (EC 1.2.1.11) is an enzyme common to the biosynthetic pathways of lysine, threonine, and methionine, as well as dianaminopimelic acid and isoleucine. The synthesis of this enzyme in *Escherichia coli* is independently repressible by the first three amino acids, although derepression is greatest upon lysine limitation (Haziza et al., 1982; Umbarger, 1978). We have previously shown that the *asd* gene of the Gram-positive bacterium *Streptococcus mutans*, a principal etiological agent in the production of dental caries (Hamada and Slade, 1980), when cloned in the plasmid vector pBR322, is expressed at a very high level in whole *E. coli* cells (Curtiss et al., 1982) and at the apparent expense of the pBR322-encoded genes for ampicillin and tetracycline resistance in minicells (Jagusztyn-Krynicka et al., 1982) producing a protein of ~40,000 M₀. Due to the limited concentrations of the components necessary for replication, transcription, and translation found in minicells (Frazer and Curtiss, 1975), this suggested that the *S. mutans asd* gene contained an efficient promoter which competed favorably with the promoters of pBR322 for available DNA-dependent RNA polymerase. Such competition has been reported among endogenous pBR322 promoters (Stuber and Bujard, 1981). Alternatively, or in addition, the mRNA transcript of the *S. mutans asd* gene might be unusually stable in *E. coli* or translated more efficiently than most *E. coli* mRNA species.

pYA575, a chimeric plasmid capable of producing a functional *asd* gene product, contains ~1330 bp of *S. mutans* DNA inserted between the EcoRI and HindIII sites of pBR322 (Jagusztyn-Krynicka et al., 1982). Subclones pYA576 and pYA577 containing deletions of the first 208 bp of the insert proximal to the *EcoRI* site and the last ~700 bp proximal to the HindIII site, respectively, were unable to complement *S. mutans asd* mutants and, therefore, conferred an *AasD⁻* phenotype. However, pYA577 produced a novel and apparently fused protein while no such protein was seen for pYA576. This suggested that putative regulatory sequences were located within the first 208 bp of the insert, proximal to the *EcoRI* site of pBR322 (Jagusztyn-Krynicka et al., 1982). pYA574 containing ~3000 nucleotides of additional contiguous *S. mutans* DNA upstream from these sequences showed approximately 2-fold higher levels of *asd* expression than did pYA575. The region required for elevated expression was narrowed to 700 bp in pYA631.

To facilitate study of the *asd* gene, we prepared a detailed restriction map of the *S. mutans* insert in pYA575. *E. coli* RNA polymerase filter binding assays were performed under conditions which enhance selective binding to promoter-containing DNA fragments (Strauss et al., 1981). The assays indicated preferential binding to an *EcoRI/HindII* fragment comprised of the first 139 bp of the insert, consistent with the subcloning data. The nucleotide sequence of this fragment contains regions of homology with those sequences involved in the promotion of RNA transcription in *E. coli* systems including five overlapping and tandem Pribnow boxlike sequences (Hawley and McClure, 1983; Pribnow, 1975a; Pribnow, 1975b; Rosenberg and Court, 1979; Siebenlist et al., 1980). A second *HindII/EcoRI* fragment of 636 bp downstream and internal to the *asd* structural gene also showed binding under conditions of reduced stringency.

Haziza et al. (1982) have reported the nucleotide sequence of the *asd* gene of *E. coli*. They had hoped to find evidence of
regulation by transcriptional termination at an attenuator (Yanofsky, 1981) but did not find the appropriate sequences. We report here the complete nucleotide sequence of the coding region for the S. mutans asd gene product containing about 1071 bp, capable of coding for a protein of 357 amino acids and Mₐ = 38,897. The sequence extending 276 bp upstream of the asd ATG start codon which contains, in addition to the promoter sequence described above, signals for both the promotion and termination of transcription (Rosenberg and Court, 1979) as well as those for ribosome binding (Gold et al., 1981) that would permit the production of a 44-amino acid polypeptide. Upon initial inspection, this sequence bears resemblance to the reported arrangement of attenuation signals in amino acid biosynthetic operons in Gram-negative bacteria (Kolter and Yanofsky, 1982). However, several differences were apparent. The $p$-independent termination sequences begin in the middle of the presumptive leader peptide coding region and are preceded by only three of a total of seven lysine, threonine, or methionine codons in the peptide. Of these seven, only the two lysine codons at the very end of the peptide are contiguous. No alternative secondary structure to promote formation of a transcription termination hairpin loop is apparent.

We altered the growth conditions of $\chi_{1825}$, an E. coli K-12 strain deleted for asd, harboring either pYA574 or the smaller pYA575 to determine if expression was affected by different concentrations of the appropriate amino acids and thereby indicate whether regulation occurred in the presence of the upstream sequences in the larger chimaera. Very little difference was evident. In vitro run-off transcription analysis did demonstrate, however, that a leader transcript was produced and terminated, at ~50% efficiency, at the proposed $p$-independent termination sequence. Four transcripts were produced from the asd promoter region unless rifampicin was added to the reaction in which case only the shortest of the transcripts, which mapped to the fourth of the five -10 sequences preceding the asd structural gene, was retained. An additional transcript was found to initiate approximately 652 bp into the asd structural gene sequence which coincided with RNA polymerase binding assays and subcloning experiments (Jaguszyn-Krzywicka et al., 1982). Measurement of the rate of degradation of asd-specific messages by hybridization of pulse-labeled RNA to specific DNA probes bound to filters (Mostellar et al., 1970) demonstrated the half-life to be in the order of 2.5 min which is comparable to E. coli specific messages (Adesnik and Levinthal, 1970; Baker and Yanofsky, 1970; Geiduscheck and Haselkorn, 1969) indicating that message stability is not involved in higher levels of asd expression.

MATERIALS AND METHODS

Bacterial Strains—E. coli K-12, $\chi_{1825}$ (gen422 2, $\tau_{p-1}$ D29 [bioH-asd]), and $\chi_{1849}$ (tonA53 dapD8 minA1 purE41 [gal-ureB] $\lambda$-minB2 his-s3 gpy25 metC56 tte-1 D29 [bioH-asd]) ibo-277 cytB2 cytA1 ona-1 hsdR2) were used.

Growth Media—The complex medium was L broth (Lennox, 1955) containing per ml 50 $\mu$g of L-meso-diaminopimelic acid (DAP), 25 $\mu$g of ampicillin (Lederle), and 12.5 $\mu$g of tetracycline (Sigma) as necessary. The minimal medium was M9 salts (Miller, 1972) containing 0.5% glucose, 1 mM MgSO₄, 0.1 mM CaCl₂, 0.5% casamino acids (Difco), 0.5 $\mu$g of biotin/ml, and supplemented per ml with 50 $\mu$g of DAP and 40 $\mu$g of adenine as needed.

Enzymes—T4 DNA ligase and restriction enzymes were purchased from Bethesda Research Laboratories except for NorI, FnuDII, CiaI, Noel, EcoRV, RsaI, and Hhal which were purchased from New England Biolabs. Digestions were routinely carried out in 10 mM Tris, pH 7.5, 10 mM MgCl₂, 6 mM $\beta$-mercaptoethanol, 100 $\mu$g of bovine serum albumin/ml, and the appropriate concentration of NaCl or KCl as recommended by the supplier. Ligonases EcoRI and SmaI digestions were carried out in the buffers recommended by the supplier. E. coli RNA polymerase holoenzyme was the kind gift of Peter Chan, David Wood, and Jacob Lebowitz. E. coli alkaline phosphatase was from Worthington, calf intestinal alkaline phosphatase from Sigma, T4 polynucleotide kinase from Pharmacia P-L Biochemicals, and the large fragment (Klenow) of E. coli DNA polymerase I from New England Nuclear. The gene 6 minicore recombinant was prepared by the method of Kerr and Sadowski (1972a, 1972b).

Plasmids and Transformation—The plasmids pYA574 and pYA575 were as described (Jaguszyn-Krzywicka et al., 1982). pYA631 was derived from pYA574 digested with EcoRV and treated with calf intestinal alkaline phosphatase to dephosphorylate the 5' ends followed by ligation to EcoRV-digested pBR322. $\chi_{1825}$ was transformed by the method of Dagert and Ehrlich (1979) and the transformation mix plated on L agar (1.5% agar (w/v) in L broth) containing 12.5 $\mu$g of tetracycline/ml and no DAP. Transformants were screened for plasmid DNA by the minilysis technique of Birnbaum and Doly (1979), and verification of the desired construction was achieved by restriction analysis.

Plasmid Purification—To amplify plasmid DNA, 170 $\mu$g of chloramphenicol (Sigma) was added per ml of a culture grown in M9 minimal medium to an A₅₅₀ = 0.5–0.6, and growth was continued at 37°C for 16 h. Cell lysis was by the method of Guerry et al. (1973) modified by a freeze-thaw step after lysozyme digestion and the use of 3% sodium dodecyl sulfate (SDS). CaCl₂-ethidium bromide buoyant density gradient centrifugation, collection, and concentration of DNA have been described (Holt et al., 1982) except that Beckman VT.50 and VT.65 rotors were used at 45,000 rpm for 20 and 14 h, respectively.

Restriction Mapping and Gel Electrophoresis—DNA restriction fragments were mapped by single and multiple digestions and subjected to electrophoresis on 2, 3, or 4% (w/v) agarose gels (Bio-Rad and Sigma) in a horizontal electrophoresis apparatus in 100 mM Tris, pH 8.3, 100 mM boric acid, 2 mM Na₂EDTA (1 × TEB). Allul and HinfI digests of pBR322 served as markers. Restriction fragments were purified from acrylamide gels as described (Maxam and Gilbert, 1980) or from agarose by electrophoresing the fragment into a DEAE membrane (Schleicher and Schuell) and eluting as recommended by the manufacturer. RNA transcripts and DNA sequence reactions were both run on 6–8% polyacrylamide, 7 M urea gels (380 × 280 × 0.375 mm). The running buffer in all cases was 1 × TEB, pH 8.3.

RNA Polymerase Binding Assay—The method employed was essentially that of Chan et al. (1979) except that the KCl concentration was varied as indicated in the figure legend.

DNA Sequence Analysis—The DNA sequence was determined by the methods of Maxam and Gilbert (1980) and Sanger et al. (1977). In the case of the latter method, templates were prepared by digesting pYA574 with either ClaI or SsI, each of which cut once within the single site in the vector, pBR322, on opposite sides of the insert. Linearized molecules were then made single stranded by digestion with the 5' specific gene 6 exonuclease of T7 and eluted from DEAE membranes and purified by polyacrylamide gel electrophoresis. Sequencing controls. pYA631 was transformed with either pYA574 or pYA575 was grown under the same conditions as the vector, pBR322, on opposite sides of the insert. Linearized molecules were then made single stranded by digestion with the 5' specific gene 6 exonuclease of T7 (Kerr and Sadowski, 1972b; Kolter and Yanofsky, 1982). Primers were gel-purified restriction fragments.

Analysis of Repression/Depression of asd by Altering Concentrations of Lysine, Threonine, and Methionine—This experiment was carried out under conditions similar to those described by Hafzah et al. (1982). A 25-μl inoculum of $\chi_{1825}$, alone or transformed with pBR322, was added to 5 ml of M9 minimal salts, 0.5% glucose, 1 mM MgSO₄, 0.1 mM CaCl₂, and per ml 5 μg of biotin, 50 μg of DAP (0.26 mM), 80 μg of DL-threonine (0.67 mM), and 20 μg of L-methionine (0.134 mM) and allowed to grow at 37°C with shaking. $\chi_{1825}$ transformed with either pYA574 or pYA575 was grown under the same conditions except for substitution of 88 μg of L-lysine HCl/ml (0.48 mM) in place of DAP. These were considered normal conditions and the controls. $\chi_{1825}$ containing pYA574 or pYA575 was also grown in this medium except the concentrations per ml of lysine, threonine, and methionine were varied as follows: (i) none added; (ii) repression: 1.827 μg of L-lysine (10 mM), 0.595 μg of DL-threonine (5 mM), 0.298 μg of L-methionine (2 mM); (iii) lysine limitation: 0.88 μg of L-lysine (0.53 mM), 80 μg of DL-threonine (0.67 mM), 20 μg of L-methionine (0.134 mM); (iv) lysine excess: 1.827 μg of L-lysine (10 mM), 80 μg of DL-threonine (0.67 mM), 20 μg of L-methionine (0.134 mM). Cells were grown to an A₅₅₀ = 0.6 which required 9–10 h, and were plated on ice. After 5 min a 200-μl sample of each culture was removed and the cells pelleted in a 1.5-ml Eppendorf tube for 2 min at ~15,000 × g. The pellets were suspended in 25 μl of sample buffer (2.3% SDS, 5% $\beta$-mercaptoethanol, 10% glycerol, 0.1% bromphenol blue, 62.5 mM Tris, pH 6.8), heated at 100°C for 5 min, and
subjected to SDS-polyacrylamide gel electrophoresis by the method of Laemmli and Favre (1973) followed by staining with Coomassie Blue (Weber and Osborn, 1969). Low molecular weight protein markers were from Bio-Rad.

**In Vitro Transcription Analysis**—Various DNA restriction fragments were used in pYA575 and pYA631 contained within a 2170-bp EcoRV fragment, beginning 575 bp upstream from the start of the *asd* structural gene and continuing downstream 160 bp beyond the end of the insert into pBR322, were used as transcriptional templates. Transcription conditions were essentially as described by Maniatis et al. (1982). KCl concentrations and presence or absence of rifampicin are as indicated in figure legends. RNA markers were kindly provided by John Donahue (University of Alabama).

**Message Stability Determination**— Cultures of *S. mutans* alone or transformed with either pBR322 or pYA575 were grown with shaking at 37°C in supplemented M9 media to an *A<sub>600</sub>* of 0.6 at which time [*H]uridine (ICN, specific activity, 44 Ci/mmol) was added at 25 μCi/ml. Shaking was continued for an additional 2 min. Rifampicin (Sigma) and cold uridine were added at respective concentrations of 300 and 201 μg/ml. Zero time, 30-s, and 1-min samples were taken followed by samples every minute thereafter for a period of 7 min total time. Samples were immediately made 20 mM in sodium azide and poured over crushed ice. Cells were harvested, lysed, and the RNA extracted by the method of Summers (1970).

pBR322 and pYA575, linearized with HindIII, and the 1100-base pair EcoRI fragment of the *asd* insert were loaded onto nitrocellulose and baked for 20 h at 42°C (Maniatis et al., 1982). Subsequent washing of the filters, covering an area in pYA575 and pYA631 contained within a structural gene and continuing downstream 160 bp beyond the end of the insert into pBR322, were used as transcriptional templates. Transcription conditions were essentially as described by Turnbough et al. (1983). KCl concentrations and presence or absence of rifampicin are as indicated in figure legends. RNA markers were kindly provided by John Donahue (University of Alabama).

**Results**

Restriction maps of pYA575, pYA631, and pYA574 are presented in Fig. 1. Proteins solubilized from whole cells of *S. mutans* alone or containing pBR322, pYA574, pYA575, or pYA631 are shown in Fig. 2. The band corresponding to the *asd* gene product is indicated. The higher level of activity of the *asd* gene consistently seen in pYA574 is retained in pYA631 indicating that those sequences required for greater expression than seen in pYA631 are located within the additional 700 bp of *S. mutans* DNA found in pYA631 upstream from the *asd* coding region.

RNA Polymerase Binding Assay—Subcloning and minicell experiments (Jaguśtyń-Krynicka et al., 1982) had suggested that the *asd* transcription regulatory signals should be located within the first 208 bp of the insert. The likelihood of a second promoter, not involved in *asd* transcription but responsible for the retention of activity of the tetracycline resistance genes of pBR322 in pYA575, despite disruption of the tet promoter by insertion of DNA at the HindIII site, had also been suggested by the subcloning data (Jaguśtyń-Krynicka et al., 1982). Alternatively, fortuitous placement of -35-like sequences within 14 bases of the HindIII end of the insert could account for this activity. Based on these considerations, *HinII* was deemed the enzyme of choice to generate fragments for the RNA polymerase binding assay. Thus, both pBR322 and pYA575 were cut with EcoRI and *HinII* to excise the insert DNA in pYA575, followed by *HinII* digestion and exposure to RNA polymerase at 50, 100, or 200 mM KC1 concentrations. DNA fragments, complexed with RNA polymerase and bound to nitrocellulose filters, were eluted, concentrated, and run on a 6% polyacrylamide gel as seen in Fig. 3.

In the only pBR322 fragments selectively retained under the stringent conditions imposed by the high KC1 concentration (Belintsev et al., 1980; Strauss et al., 1980) were those of 1000 bp which contains the β-lactamase promoter and 517 bp which specifies the 104-nucleotide RNA associated with incompatibility and copy number control in plasmids which containColEl replication functions (Chan et al., 1973; Stubber and Bujal, 1981). The EcoRI/*HinII* fragment corresponding to the first 139 bp of insert DNA, as seen in Fig. 1, was also retained, further indicating the presence of an efficient promoter in this region. The 626-bp *HinII*/EcoRI fragment appears, faintly, as well. All of the other fragments generated from the insert DNA, except that of 124-bp *HinII*/*HinII* fragment, demonstrate binding under less stringent conditions.

**DNA Sequence**—The sequence of 140 nucleotides encompassing the *asd* gene is presented in Fig. 4. The majority of the sequence was generated by the chain-terminating method of Sanger et al. (1977). The sequencing strategy, indicating primer extension and the 295 bases determined by the method of Maxam and Gilbert (1980), is shown in Fig. 1. Of the sequence described, 84% was generated from both strands...
including regions involved in the initiation and termination of both transcription and translation.

The *asd* gene has an unusual promoter region located between nucleotides 205 and 259. It contains a 35-like RNA polymerase recognition sequence of TTGATA, preceded by a run of Ts and a frequently conserved A at a spacing of 8 nucleotides upstream from TTG (Hawley and McClure, 1983; Siebenlist et al., 1980). At a spacing of 6 nucleotides downstream from the -35 hexamer begins a series of five potential RNA polymerase binding sequences contained within a stretch of 24 nucleotides. All five contain a first and second position TA and a sixth position T. The second of these is a perfect Pribnow box consensus of TATAAT(G) (Pribnow, 1975a; Pribnow, 1975b). Each of these sequences is followed by one or more purines, in all cases an A, in some a G as well, at a spacing from the sixth position T of the -10 hexamer of 5–8 nucleotides which is appropriate for the site of initiation of transcription (Hawley and McClure, 1983). Spacing between the -35 and -10 hexamers in *E. coli* is optimally 17 ± 1 nucleotides (Hawley and McClure, 1983; Siebenlist et al., 1980) suggesting the fourth sequence of TATTAT, at a spacing of 18 nucleotides, as the most likely RNA polymerase binding site.

Slightly farther downstream is a 9-base Shine-Dalgarno ribosome binding sequence of TAAAGAGGT bearing strong complementarity to the 3′ terminus of *E. coli* 16S ribosomal RNA (Gold et al., 1981; Shine and Dalgarno, 1974) (calculated ∆G = -13.5 kcal (Salser, 1977)). This is followed at a spacing of 5 nucleotides by an ATG translation initiation codon. The entire coding sequence of the *asd* gene is contained within nucleotides 274–1344 which would permit the synthesis of a protein of 357 amino acids and a molecular weight of 38,897. This is in good agreement with the size estimated by SDS-polyacrylamide gel electrophoresis (Fig. 2).

Sequences upstream of the promoter region contain features which suggest involvement in regulation of *asd* expression. An area of dyad symmetry is found between nucleotides 160 and 177 which is followed by a run of Ts. Such a sequence has been implicated in the ρ-independent termination of transcription (Holmes et al., 1983; Rosenberg and Court, 1979), particularly with regard to the regulation of gene expression in the mechanism of attenuation in amino acid biosynthetic operons (Farnham and Platt, 1980; Kolter and Yanofsky, 1982). A ∆G = -10.7 kcal can be calculated for the most stable configuration of this putative hairpin (Salser, 1977). The sequences from 2 to 47 also bear strong homology to sequences found to be involved in the promotion of transcription (Hawley and McClure, 1983; Rosenberg and Court, 1979; Siebenlist et al., 1980) with a -35-like sequence at 11–16 and a -10-like sequence at 34–39. This would allow transcription initiation at the A at position 46 and would permit the synthesis of a transcript, if terminated after the putative hairpin, of 140–145 nucleotides.

A Shine-Dalgarno sequence (Shine and Dalgarno, 1974) involved in ribosome binding (Gold et al., 1981) is found at 82–90, within the putative leader transcript region. This sequence is very similar to that preceding the *asd* coding sequence and gives the same calculated ∆G = -13.2 kcal (Salser, 1977) for interaction with the 3′ end of *E. coli* 16S ribosomal RNA (Gold et al., 1981; Shine and Dalgarno, 1974). An ATG start codon occurs at a spacing of 4 nucleotides at positions 95–97 followed by an open reading frame through position 226 which would permit the synthesis of a leader peptide of 44 amino acids. Unlike leader peptides formed in other amino acid biosynthetic operons (Kolter and Yanofsky, 1982) few codons for the end product amino acids of the biosynthetic pathways in which aspartate β-semialdehyde dehydrogenase is involved (lysine, threonine, and methionine) are found. Three other features are significantly different as well. The ρ-independent transcription termination sequence occurs in the middle of the coding region for the peptide, no alternative secondary structure to block formation of the terminator hairpin is evident, and a completely separate set of sequences capable of promoting transcription of the *asd* structural gene is present.

**Effect of Lysine, Threonine, or Methionine Concentration on Expression of the S. mutans asd Gene**—Since the *asd* gene product is easily detectable by SDS-polyacrylamide gel electrophoresis (Fig. 2), alterations in the activity of the gene should be readily discernible. Samples of proteins extracted from whole cells of x1825, alone or transformed with pBR322, pYA631, pYA574, and pYA575. The *asd* gene product is indicated. The scale used to estimate molecular mass is shown at the right.
**In Vitro Transcription Analysis**—Transcription initiation and termination sites were identified by *in vitro* transcription of various restriction fragments covering the 2170 bp between the *EcoRV* sites of pYA631, as outlined in Figs. 6 and 7. Transcripts were subjected to electrophoresis and mapped according to their size.

The efficiency and specificity of *in vitro* transcription increases with increasing salt concentration, as can be seen in Fig. 6. A 1009-bp *EcoRV/PvuII* fragment containing the *asd* promoter region was used as a template. This specificity is further enhanced by the addition of rifampicin to a concentration of 4 μM, 30 s after initiating the reaction. The main transcripts run essentially as two doublet bands of approximately 235/237 and 242/246 nucleotides. The presence of rifampicin virtually eliminates all of these but the 235-nucleotide transcript which maps to a position 9 nucleotides after the fourth and most favorably spaced -10 sequence. The 237-nucleotide transcript also maps to this sequence while the 242- and 246-nucleotide transcripts would initiate at positions following the third and perhaps second -10-like sequences, respectively. This suggests that *in vitro*, and under the least stringent conditions, spacing constraints between the -35 and -10 hexamers may be overcome. The 143- and 442-nucleotide transcripts are initiated upstream of the aspartate B-semialdehyde dehydrogenase promoter as can be seen in Fig. 7.

The entire 2170-bp *EcoRV* fragment produced three major transcripts of approximately 1400, 725, and 143 nucleotides as seen in *Lane* 1 of Fig. 7. The 143-nucleotide transcript was localized to a 490-bp *EcoRI* fragment (*Lane* 4) immediately upstream of *asd* and was found in transcription of all fragments (*Lanes* 1–4) which contained this DNA segment. This suggests a fixed terminus for this transcript. Based on its synthesis from specified DNA fragments and its size, it was identified as the putative leader transcript suggested by the nucleotide sequence which also indicated the location of the p-independent termination site.

A 299-bp *EcoRI* to *PvuII* fragment (nucleotides 189–487 of the sequence) produced four transcripts that migrate essentially as two doublet bands of approximately 235, 237, 242, and 246 nucleotides (*Lane* 5). These same transcripts were produced by the 1009-bp *EcoRV* to *PvuII* fragment (*Lane* 3 and Fig. 6) which also produced the transcripts of approximately 442 and 143 nucleotides. The 442-nucleotide transcript may be produced by inefficient termination of the 143-nucleotide transcript since it is absent in both *Lanes* 4 and 5 of Fig. 7 and is the appropriate size.

Transcription of a 1335-bp *EcoRV* to *BamHI* fragment seen in *Lane* 2 of Fig. 7 supports this contention. Transcripts of approximately 143 and 563 nucleotides corresponding to the terminated leader transcript and that for *asd* are seen. Also evident is a transcript of approximately 768 nucleotides which, again, is appropriate for a read-through of the termination site for the 143-nucleotide transcript. ITT has been demonstrated to reduce efficient termination by interfering with the formation of stable stem-loop structures (Lee and Yanofsky, 1977; Miller, 1972). If ITT is substituted for GTP in the reaction, the 143-nucleotide transcript disappears; however, the 768-nucleotide transcript does not increase as anticipated (data not shown).

A third independent transcript seen in *Lanes* 7–9 of Fig. 7 is initiated within the 453-bp *BamHI* to *EcoRI* fragment producing transcripts of approximately 342 and 725 nucleotides when run out to the *EcoRI* and *EcoRV* sites, respectively. The initiation site of this transcript may be located at a position 6 nucleotides downstream from a possible -10 RNA polymerase binding site at positions 913–918 of the sequence (Fig. 3) which is preceded by a -35 sequence identical to that...
for the leader transcript at positions 889–894. Alternatively, a second possible promoter is located slightly downstream with a -35 sequence at positions 899–904 and a -10 sequence at 922–927. The location of a transcriptional initiation site within this region had been indicated by both RNA polymerase binding and subcloning experiments and probably serves as a promoter site in expression of the pBR322 tet gene in these constructs (Jagusztyn-Krynicka et al., 1982).

Stability of asd-specific mRNA—To determine if efficient expression of the S. mutans asd gene in E. coli is a result of slow turnover of the asd-specific message, stability experiments were undertaken by hybridizing [3H]uridine pulsed labeled mRNA to specific DNA probes bound to filters. As seen in Fig. 8, the half-life for mRNA purified from strain x1849 containing pYA575 and hybridized to the 1100-bp EcoRI fragment (EcoA), containing the majority of the coding sequence of ad, was approximately equal to that of the same messages hybridized to the vector, pBR322. The half-life determinations for pYA575/pYA575 and pBR322/pBR322 RNA/DNA hybridizations were essentially the same. Of possible significance is that in both hybridizations involving pYA575 RNA and DNA containing asd-specific sequences, the initial rate of degradation was slightly more rapid. Also, asd-specific messages are produced at a level approximately 2-fold higher than total messages specific for pBR322 sequences when the template for transcription is pYA575.

DISCUSSION

Southern blot hybridization analysis (Southern, 1975) of the original S. mutans asd clones indicated a lack of homology with the E. coli genome (Jagusztyn-Krynicka et al., 1982). Comparison of the nucleotide sequences of the S. mutans and E. coli asd genes confirms this absence of homology both in the coding region and in that containing the signals involved in the initiation of transcription and translation. Despite the fact that the asd gene of S. mutans efficiently complements a deletion of the corresponding gene in E. coli, little homology exists between the deduced amino acid sequences of either aspartate P-semialdehyde dehydrogenase protein. Codon utilization is very different, as can be seen in Table I. The S. mutans coding region encompasses 1071 nucleotides and would permit the synthesis of a protein of 357 amino acids.
for aspartic acid and cysteine for histidine at residues 134 and 135. A cysteine residue at the active site is more in line with that seen for other dehydrogenases (Harris and Waters, 1976). Two of the three cysteine residues found in the *S. mutans* deduced amino acid sequence occur at positions 125 and 128. If one assumes the cysteine residue at position 128 to be the active site of both molecules, the sequence can be overlapped with that of *E. coli* so that 16 of 42 surrounding amino acids coincide, the longest run of homology being 9 of 13 residues between positions 156 and 168 in the *E. coli* sequence and 149 and 161 of the *S. mutans* sequence. This indicates some degree of sequence conservation about what may be the active site of both molecules.

Attenuation mechanisms (Kolter and Yanofsky, 1982; Yanofsky, 1981) have been implicated in the regulation of several amino acid biosynthetic operons (Barnes, 1978; Bertrand *et al.*, 1975; DiNocera *et al.*, 1978; Gardner, 1979; Lawther and Hatfield, 1980; Lee and Yanofsky, 1977; Nargang *et al.*, 1980; Zurawski *et al.*, 1978) as well as in the regulation of resistance to erythromycin (Horinouchi and Weisblum, 1980) and pyrimidine biosynthesis (Turnbough *et al.*, 1983). In general, in the amino acid operons the mechanism involves a transcribed leader region, prior to the coding sequence for the first structural gene, which encodes a leader peptide containing several residues corresponding to the amino acid end product of the operon. Two regions of dyad symmetry occur in the transcript, one in the coding region for the leader peptide followed by a second, closely resembling a ρ-independent transcription terminator (Rosenberg and Court, 1979), preceding a run of uridine nucleotides. These regions in turn can interact to form a single base-paired structure. Termination or read-through of transcription is dependent on the availability of the appropriate charged tRNAs for translation which mediates which secondary structure will form.

In determining the nucleotide sequence of the *asd* gene of *S. mutans* we were somewhat surprised to find a sequence resembling an attenuator preceding the coding region for aspartate β-semialdehyde dehydrogenase, since no evidence of an attenuator had been identified in the *E. coli* system (Haziza *et al.*, 1982). The fact that the *E. coli* gene is regulated by levels of lysine principally but also threonine and methionine whereas no such regulation could be demonstrated by the *S. mutans* gene, at least in *E. coli*, indicates that this attenuator-like sequence does not function in that capacity. However, in *vitro* transcription indicates that a leader transcript is synthesized and terminated with about 50% efficiency. A putative peptide is encoded by the leader transcript, but it spans the terminator and continues for 13 amino acid residues beyond the point of transcription termination. Seven codons specifying amino acids that are ultimate products of the action of the *asd* gene product occur in the peptide; except for two lysines these amino acids are noncontiguous, and four are found subsequent to the transcription termination sequence. It is possible that this transcription termination site serves as simply a pause site to permit better strand separation about the downstream promoter immediately preceding the *asd* structural gene. This would allow greater efficiency of transcription initiation and hence higher levels of aspartate

![Fig. 5. Protein profiles of x1825 containing pYA574 or pYA575 grown under conditions of varying concentrations of lysine, threonine, and methionine. Cells harboring the *asd*-containing plasmids pYA574 or pYA575 were grown under conditions of repression (R), lysine limitation (LL), lysine excess (LE), or with no lysine, threonine, or methionine in the media (NA). x1825 alone and transformed with pBR322, pYA574, or pYA575 were grown in complete minimal media as controls. The *asd* gene product is indicated.](image)

![Fig. 6. *In vitro* transcription under conditions of varied salt concentration in the absence or presence of rifampicin. A 1009-bp *EcoRV/PvuII* fragment of pYA631 served as a transcriptional template in reactions containing either 50, 100, or 150 mM KCl and with or without the addition of rifampicin to a final concentration of 4 μM, 30 s after commencement of the reaction. Specificity of transcription increased with the increase in KCl concentration. The addition of rifampicin eliminated the 242- and 246-nucleotide transcripts and greatly reduced that of 237 nucleotides.](image)
sequences when pYA575 is the transcriptional template. This includes both the β-lactamase gene and the sequences coding for the 104-nucleotide RNA which have been reported to have strong promoters (Chan et al., 1979; Stuber and Bujard, 1981). Furthermore, a fusion of the asd promoter and translational signals with the E. coli lacZ gene on the plasmid vector pMLB1034 (Berman, 1983) is capable of producing over 8000 units of β-galactosidase activity² (Miller, 1972). We have also used this promoter region to construct an expression vector, pYA626, which has been successfully employed by members of our laboratory to express Mycobacterium leprae genes in E. coli (Clark-Curtiss et al., 1985; Jacobs et al., 1986). These characteristics are comparable to those described by Gentz and Bujard (1985) for "efficient unregulated promoters in the E. coli system."

We had anticipated that the S. mutans asd gene, constitutively expressed in E. coli on the plasmid pYA575, would respond to excess or limiting concentrations of lysine when upstream sequences coding for the leader transcript were supplied in the plasmid pYA574. The fact that, regardless of lysine concentration, the pYA574-carried gene specifies higher levels of aspartate β-semialdehyde dehydrogenase may be alternatively explained by the availability of the read-through leader transcripts for translation. However, this seems inadequate and, of course, is only based on in vitro and not in vivo studies of transcripts formed. We firmly believe the presence of the leader region is more than coincidence and plays some role in expression of the asd gene. Haziza et al. (1982) proposed the involvement of a regulatory protein mediating the effects of lysine, threonine, and methionine concentration. If such a molecule recognized sequences specific to the E. coli gene it is unlikely it would act upon that of S. mutans due to the lack of homology between the sequences. To determine a mechanism of regulation would then require

² G. A. Cardineau, unpublished data.
returning the asd gene to S. mutans. To this end, we have constructed a shuttle vector, pYA629, capable of being transformed into and replicated by both E. coli and S. mutans (Murchison et al., 1986). Transformation of S. mutans with pYA629 carrying the asd/lacZ fusion described above will provide us with a readily assayable gene product which will facilitate the study of regulation of the gene and its level of activity in its normal host.

Acknowledgments—We are grateful to Peter Chan and Jacob Lembowitz for instruction in both RNA polymerase binding and Maxam/Gilbert DNA sequencing techniques, John Donahue and Charles Turnbough for instruction in *in vitro* transcription analysis and critical discussion, and to members of our laboratory for their support and assistance. We also thank Pat Pierce, Amy Papian, and Nancy Burkhart for their contributions toward typing the manuscript.

REFERENCES

Adesnik, M., and Levinthal, C. (1970) Cold Spring Harbor Symp. Quant. Biol. 35, 451-459

Baker, R., and Yanofsky, C. (1970) Cold Spring Harbor Symp. Quant. Biol. 35, 467-470

Barnes, W. M. (1978) *Proc. Natl. Acad. Sci. U. S. A.* 75, 4281-4285

Belitsky, B. N., Zavriev, S. K., and Shemyakin, M. P. (1980) Nucleic Acids Res. 8, 1391-1404

Berman, M. L. (1983) *Biotecniques* 1, 178-183

Bertrand, K., Korn, L., Lee, F., Platt, T., Squires, C. L., Squires, C., and Yanofsky, C. (1975) *Science* 189, 22-26

Bieilmann, J. F., Edl, P., Hirth, C., and Jornvall, H. (1980) *Eu. J. Biochem.* 104, 59-64

Birnboim, H. C., and Doly, J. (1979) Nucleic Acids Res. 7, 1513-1523

Bever, K., and Szybalski, W. (1971) *Methods Enzymol.* 21, 350-383

Chen, P. T., Lebowitz, J. W., and Bastia, D. (1979) *Nucleic Acids Res.* 7, 1247-1262

Clark-Curtiss, J. E., Jacobs, W. R., Docherty, M. A., Ritchie, L. R., and Curtiss, R. III (1985) *Bacteriol. Rev.* 161, 1093-1102

Curtiss, R. III, Jagusztyn-Krynicka, E. K., Hansen, J. B., Smorawinska, M., Abiko, Y., and Cardineau, G. (1982) in *Microbial Drug Resistance* (Mitsushashi, S., ed.) Vol. 3, pp. 15-27, Japan Scientific Press, Tokyo and University Park Press, Baltimore

Dagert, M., and Ehrlich, S. D. (1979) *Gene* (Amst.) 6, 23-28

Davis, R., and Vapnek, D. (1976) *J. Bacteriol.* 1105, 1113-1120

DiNocera, P. F., Blesi, F., DiLauro, R., Frunzio, R., and Bruni, C. B. (1978) *Proc. Natl. Acad. Sci. U. S. A.* 75, 4270-4290

Farnham, P. J., and Platt, T. (1980) *Cell* 20, 739-748

Frazer, A. C., and Curtiss, R. III (1975) *Curr. Top. Microbiol. Immunol.* 69, 3-84

Gardner, J. F. (1979) *Proc. Natl. Acad. Sci. U. S. A.* 76, 1706-1710

Geiduscheck, E. P., and Haselkorn, R. (1969) *Annu. Rev. Biochem.* 38, 647-676

Gents, R., and Bujard, H. (1985) *J. Bacteriol.* 164, 70-77

Gold, L., Pribnow, D., Schneider, T., Shindling, S., Singer, B. S., and Stormo, G. (1981) *Annu. Rev. Microbiol.* 35, 365-403

Guerry, P., LeBlanc, D. J., and Fallow, S. (1973) *J. Bacteriol.* 116, 1064-1066

Hamada, S., and Slade, H. D. (1980) *Microbiol. Rev.* 44, 331-384

Harris, J. I., and Waters, M. (1976) in *The Enzymes* (Boyer, P. D., ed.) 3rd Ed., Vol. 15, pp. 1-49, Academic Press, Orlando, FL

Hawley, D. K., and McClure, W. R. (1986) *Nucleic Acids Res.* 11, 2237-2255

Hawza, C., Stragier, P., and Patte, J.-C. (1982) *EMBO* J. 1, 379-384

Holingshead, S. K., Fischetti, V. A., and Scott, J. R. (1986) *J. Biol. Chem.* 261, 1677-1686

Holmes, W. M., Platt, T., and Rosenberg, M. (1983) *Cell* 32, 1029-1032

Holt, R. G., Abiko, Y., Saito, S., Smorawinska, M., Hansen, J. B., and Curtiss, R. III (1982) *Infect. Immun.* 38, 147-156

Horinoouchi, S., and Weisblum, B. (1980) *Proc. Natl. Acad. Sci. U. S. A.* 77, 7079-7083

Jacobs, W. R., Docherty, M. A., Curtiss, R. III, and Clark-Curtiss, J. E. (1986) *Proc. Natl. Acad. Sci. U. S. A.* 83, 1926-1930

Jagusztyn-Krynicka, E. K., Smorawinska, M., and Curtiss, R. III (1982) *J. Gen. Microbiol.* 128, 1313-1314

Kerr, C., and Sadowski, P. D. (1972a) *J. Biol. Chem.* 247, 305-310

Kerr, C., and Sadowski, P. D. (1972b) *J. Biol. Chem.* 247, 311-318

Kolter, R., and Yanofsky, C. (1982) *Annu. Rev. Genet.* 16, 113-134

Laemmli, U. K., and Favre, M. (1973) *J. Mol. Biol.* 80, 575-599

Lawther, R. P., and Hatfield, G. W. (1980) *Proc. Natl. Acad. Sci. U. S. A.* 77, 1892-1896

Lee, F., and Yanofsky, C. (1977) *Proc. Natl. Acad. Sci. U. S. A.* 74, 4365-4369

Lennox, E. S. (1955) *Virology* 1, 190-206

Makite, R., Roe, B., and Ferretti, J. J. (1985) *Gene* (Amst.) 35, 357-362

Maniatis, T., Fritsch, E. F., and Sambrook, J. C. (1982) *Molecular Cloning: A Laboratory Manual*, p. 324, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Mannarelli, B., Baignedes, T. S., Greenberg, B., Springhorn, S. S., and Lacks, S. A. (1986) *Proc. Natl. Acad. Sci. U. S. A.* 82, 4468-4472

Maxam, A. M., and Gilbert, W. (1980) *Methods Enzymol.* 65, 499-560

Miller, J. H. (1972) *Experiments in Molecular Genetics*, p. 431, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY

Mostellar, R. D., Rosen, J. K., and Yanofsky, C. (1970) *Cold Spring Harbor Symp. Quant. Biol.* 35, 461-466

Murchison, H. H., Barrett, J. F., Cardineau, G. A., and Curtiss, R. III (1986) *Infect. Immun.* 54, 273-282
Nargang, F. E., Srahmanym, C. S., and Umbarger, H. E. (1980) Proc. Natl. Acad. Sci. U. S. A. 77, 1823–1827
Pribnow, D. (1975a) Proc. Natl. Acad. Sci. U. S. A. 72, 784–788
Pribnow, D. (1975b) J. Mol. Biol. 99, 419–443
Reznikoff, W. S., Siegle, D. A., Cowing, D. W., and Gross, C. A. (1985) Annu. Rev. Genet. 19, 355–387
Rosenberg, M., and Court, D. (1979) Annu. Rev. Genet. 13, 319–353
Salser, W. (1977) Cold Spring Harbor Symp. Quant. Biol. 42, 985–1002
Sanger, F., Nicklen, S., and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. U. S. A. 74, 5463–5467
Shine, J., and Dalgarno, L. (1974) Proc. Natl. Acad. Sci. U. S. A. 71, 1342–1346
Siebenlist, U., Simpson, R. B., and Gilbert, W. (1980) Cell 20, 269–281
Southern, E. M. (1975) J. Mol. Biol. 98, 503–517
Stassi, D. L., Dunn, J. J., and Lacks, S. A. (1982) Gene 20, 359–366
Strauss, H. S., Boston, R. S., Record, M. T., Jr., and Burgess, R. R. (1981) Gene (Amst.) 13, 75–87
Strauss, H. S., Burgess, R. R., and Record, M. T., Jr. (1980) Biochemistry 19, 3496–3503
Stuber, D., and Bujard, H. (1981) Proc. Natl. Acad. Sci. U. S. A. 78, 167–171
Summers, W. C. (1970) Anal. Biochem. 33, 459–463
Turnbough, C. L., Hicks, K. L., and Donahue, J. P. (1983) Proc. Natl. Acad. Sci. U. S. A. 80, 368–372
Umbarger, H. E. (1978) Annu. Rev. Biochem. 47, 533–606
Weber, K., and Osborn, M. (1969) J. Biol. Chem. 244, 4406–4412
Yamamoto, T., Yamagata, S., Horii, K., and Yamagishi, S. (1982) J. Bacteriol. 150, 269–276
Yanofsky, C. (1981) Nature 289, 751–758
Zurawski, G., Brown, K., Killingly, D., and Yanofsky, C. (1978) Proc. Natl. Acad. Sci. U. S. A. 75, 4271–4275