Volume 13 Number 6 1985 Nucleic Acids Research

Cloning and nucleotide sequence of the aspartase gene of *Escherichia coli* W

Jun S.Takagi, Nobuo Ida, Masanobu Tokushige, Hiroshi Sakamoto* and Yoshiro Shimura*

Department of Chemistry and *Department of Biophysics, Faculty of Science, Kyoto University, Kyoto 606, Japan

Received 1 February 1985; Accepted 22 February 1985

ABSTRACT

The aspA gene of Escherichia coli W which encodes aspartase was cloned into the plasmid vector pBR322. The nucleotide sequences of aspA and its flanking regions were determined. The aspA gene encodes a protein with a molecular weight of 52,224 consisted of 477 amino acid residues. The amino acid sequence of the protein predicted from the nucleotide sequence was consistent with those of the NH_2- and COOH-terminal regions and also with the amino acid composition of the purified aspartase determined previously. Potential promoter and terminator sequences for aspA were also found in the determined sequence.

INTRODUCTION

Aspartase (L-aspartate ammonia-lyase, EC 4.3.1.1) catalyzes the reversible conversion of L-aspartate to fumarate and ammonia. The enzyme is present in various bacteria and a few higher plants, whereas little information is available for its existence in animals (1). The physiological role of aspartase in Escherichia coli cells has been studied by measuring the enzyme level in the cells grown under various conditions. When glucose was employed as the sole carbon source and amination was essential for growth, the aspartase activity was low. In contrast, when glutamate served as the carbon source and deamination was essential, the aspartase activity was high. These results suggest that aspartase plays primarily a catabolic role rather than a biosynthetic role (2). In fact, studies using mutants which lacked aspartase showed that the enzyme is essential for the catabolism of glutamate (3). Separate experiments also showed that cells grown anaerobically on glucose contained more aspartase activity than those grown aerobically on glucose. Under the anaerobic conditions, aspartase seems to participate in the formation of fumarate and succinate (4). The enzyme has been purified to homogeneity from E. coli B (5), E. coli W (6), and Pseudomonas fluorescens (7). Of the three enzymes, the E. coli W enzyme has been studied most extensively. It has a molecular weight of 193,000 as determined by

sedimentation equilibrium analysis (6) and is composed of four identical subunits arranged in D_2 symmetry (8). The amino acid composition (6), the NH_2-terminal amino acid residue (9), and the sequence of the COOH-terminal octapeptide (10) were determined. The enzyme exhibits a complex cooperative behavior in the reaction kinetics (6). Chemical modification studies showed that cysteine (11) and histidine residues (12) are essential for the activity of the enzyme, although their functions in the catalytic mechanism are not known. Information of the primary structure is required for better understanding of the catalytic and regulatory mechanism of this enzyme. Cloning of the structural gene of aspartase, aspA, and determination of the nucleotide sequence of the gene will reveal the amino acid sequence of the enzyme and may also provide some insights into the mode of regulation of the gene expression. In this communication, we wish to report the results of studies on cloning of the aspA gene of E. coli W and determination of the nucleotide sequences of the gene and its flanking regions.

MATERIALS AND METHODS

Bacterial Strains and Media

An E. coli K-12 strain C600r-m- and its derivatives, TK6 (glt) and TK237 (aspA glt), where the glt mutation leads to constitutive uptake of glutamate, (13) were kindly provided by Dr. M. Kisumi, Research Laboratory of Applied Biochemistry, Tanabe Seiyaku Co. JM101 was used as a host for M13 phages. E. coli W, provided by Professor T. Yura, has been maintained in our laboratory. LB medium (14) supplemented with 0.1% glucose was used as the rich medium. 2X YT medium was for propagation of JM101. The glutamate minimal medium contained 30 mM monosodium L-glutamate, 0.7% K_2HPO_4, 0.3% KH_2PO_4, 0.1% $(NH_4)_2SO_4$, and 0.01% $MgSO_4 \cdot 7H_2O$ supplemented with 2 μg/ml thiamine hydrochloride, 1 mM L-threonine, and 1 mM L-leucine. The modified ASP medium (13) consisted of 132 mM fumarate, 67 mM NH_4OH, 0.2% K_2HPO_4, 0.05% $MgSO_4 \cdot 7H_2O$, 2% polypeptone, and 2% yeast extract (pH 7.4).

DNA Preparations

DNA was prepared from E. coli W cells by the method of Cosloy and Oishi (15) and purified by centrifugation in a CsCl-ethidium bromide gradient. Plasmids were prepared according to Maniatis et al. (14).

Nucleotide Sequence Analysis

Various restriction fragments were cloned into M13mp18 or M13mp19 (16) and sequenced by the dideoxy chain termination method (17).

Enzyme Assay

For measurement of aspartase activity, the cells were grown in the modified ASP medium containing ampicillin (50 μg/ml), where possible, for 16 h at 37°C with gentle shaking. Crude extracts were prepared by sonic disruption of the cells and aspartase activity was measured according to Suzuki et al. (6). The assay mixture contained 100 mM monosodium L-aspartate, 2 mM $MgCl_2$, and 50 mM N-tris(hydroxymethyl)methyl-3-aminopropanesulfonate-NaOH (pH 8.5). Protein concentration was determined by the method of Lowry et al. (18).

Materials

Restriction endonucleases were purchased from Takara Shuzo, New England Biolabs, Bethesda Research Laboratories, and P-L Biochemicals. T4 DNA ligase and M13 sequencing kit were from Takara Shuzo and bacterial alkaline phosphatase was from P-L Biochemicals. They were used as directed by the suppliers. [α-^{32}P]dCTP (410 Ci/mmol) was obtained from Amersham.

RESULTS AND DISCUSSION

Cloning of the aspA Gene

Aspartase activity is essential for growth of E. coli cells on glutamate as the sole source of carbon. However, aspA mutants which lack aspartase cannot be selected on glutamate minimal medium, because wild type strains themselves do not grow or grow very slowly on the medium due to severe repression of glutamate transport system. This situation necessitates the use of strains that express the transport system constitutively to select the aspA mutants (3). The same problems exist when one tries to clone the aspA gene by complementation. Since the glt mutation causes constitutive uptake of glutamate, TK6 (glt) is able to grow well on glutamate minimal medium. TK237 (aspA glt), which was derived from TK6, also expresses the glutamate transport system constitutively. However, having no aspartase activity TK237 is unable to grow on glutamate minimal medium. If the aspA mutation is complemented, TK237 should be able to grow well on glutamate minimal medium. Therefore, selection of the clone which harbored a hybrid plasmid carrying aspA was performed by using TK237 as a host and screening the transformants that grew well on glutamate minimal medium. DNA preparation from E. coli W was digested with various restriction enzymes and inserted into the relevant sites of pBR322. TK237 was transformed with these preparations and screened on glutamate minimal medium containing ampicillin. Initial attempts using the complete digests with BamHI, SalI, or BamHI plus SalI and the partial digests with Sau3AI were unsuccesful. Among other restriction enzymes tested, AvaI

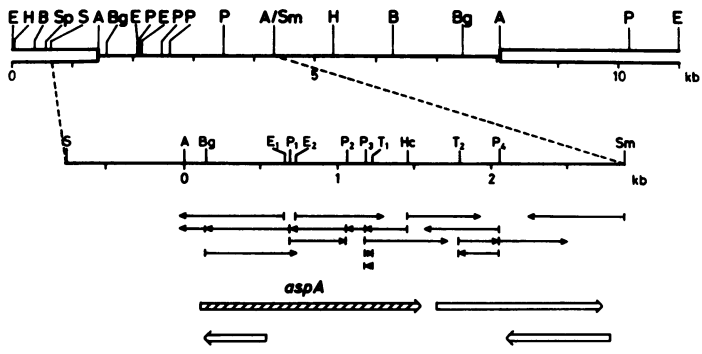

Figure 1. Restriction map of pA303 and sequencing strategy for the 2.9 kb AvaI-SmaI segment in the pA303 insert. Restriction map of pA303 is shown as a linearized form with the EcoRI site of pBR322 derived segment (open box) at both ends. The SalI-SmaI segment is shown in an expanded form with sequencing strategy. Only the restriction sites used for M13 cloning are shown in this segment. Arrowheads indicate the directions of sequences and the lines behind the arrowheads indicate portions that actually yielded satisfactory sequences. Open boxes with arrowheads represent open reading frames (longer than 400 nucleotides) and their directions. The open reading frame of aspA is hatched. Restriction enzyme sites are indicated by the following abbreviations: A, AvaI; B, BamHI; Bg, BglII; E, EcoRI; Hc, HincII; H, HindIII; P, PstI; S, SalI; Sm, SmaI; Sp, SphI; T, TaqI.

digestion yielded 4 clones which grew well on glutamate minimal medium out of 10,000 transformants. Plasmids were prepared from these clones and analyzed by restriction mapping. Thus, all the plasmids were shown to contain the same insert in the same orientation. The restriction map of one of the plasmids, designated pA303, is shown in Fig. 1. Besides the two terminal sites, an additional AvaI site was present in the insert despite extensive digestion of the bacterial DNA with the enzyme. It is possibe that two independent AvaI fragments were ligated to form the 6.6 kb insert. The specific activity of aspartase in the crude extract of TK237 carrying pA303 was about three times that of E. coli W (Table 1). Complementation of the aspA mutation of TK237 and increase of the aspartase activity caused by pA303 strongly support that

Table 1. Aspartase activity of E. coli strains.

Strains	Specific activity (units/mg protein)
W	0.91
K-12 TK6	0.42
K-12 TK237	0
K-12 TK237/pA303	2.8

```
                                                -35                 -10
CTCGGGTATTCGGTCGATGCAGGGGATAATCGTCGGTCGAAAAACATTCGAAACCACATATATTCTGTG      69

TGTTTAAAGCAAATCATTGGCAGCTTGAAAAAGAAGGTTCACATGTCAAACAACATTCGTATCGAAGAA     138
                                        MetSerAsnAsnIleArgIleGluGlu         9

GATCTGTTGGGTACCAGGGAAGTTCCAGCTGATGCCTACTATGGTGTTCACACTCTGAGAGCGATTGAA     207
AspLeuLeuGlyThrArgGluValProAlaAspAlaTyrTyrGlyValHisThrLeuArgAlaIleGlu      32

AACTTCTATATCAGCAACAACAAAATCAGTGATATTCCTGAATTTGTTCGCGGTATGGTAATGGTTAAA     276
AsnPheTyrIleSerAsnAsnLysIleSerAspIleProGluPheValArgGlyMetValMetValLys      55

AAAGCCGCAGCTATGGCAAACAAAGAGCTGCAAACCATTCCTAAAAGTGTAGCGAATGCCATCATTGCC     345
LysAlaAlaAlaMetAlaAsnLysGluLeuGlnThrIleProLysSerValAlaAsnAlaIleIleAla      78

GCATGTGATGAAGTCCTGAACAACGGAAAATGCATGGATCAGTTCCCGGTAGACGTCTACCAGGGCGGC     414
AlaCysAspGluValLeuAsnAsnGlyLysCysMetAspGlnPheProValAspValTyrGlnGlyGly     101

GCAGGTACTTCCGTAAACATGAACACCAACGAAGTGCTGGCCAATATCGGTCTGGAACTGATGGGTCAC     483
AlaGlyThrSerValAsnMetAsnThrAsnGluValLeuAlaAsnIleGlyLeuGluLeuMetGlyHis     124

CAGAAAGGTGAATATCAGTACCTGAACCCGAACGACCATGTTAACAAATGTCAGTCCACTAACGACGCC     552
GlnLysGlyGluTyrGlnTyrLeuAsnProAsnAspHisValAsnLysCysGlnSerThrAsnAspAla     147

TACCCGACCGGTTTCCGTATCGCAGTTTACTCTTCTCTGATTAAGCTGGTAGATGCGATTAACCAACTG     621
TyrProThrGlyPheArgIleAlaValTyrSerSerLeuIleLysLeuValAspAlaIleAsnGlnLeu     170

CGTGAAGGCTTTGAACGTAAAGCTGTCGAATTCCAGGACATCCTGAAAATGGGTCGTACCCAGCTGCAG     690
ArgGluGlyPheGluArgLysAlaValGluPheGlnAspIleLeuLysMetGlyArgThrGlnLeuGln     193

GACGCAGTACCGATGACCCTCGGTCAGGAATTCCGCGCTTTCAGCATCCTGCTGAAAGAAGAAGTGAAA     759
AspAlaValProMetThrLeuGlyGlnGluPheArgAlaPheSerIleLeuLeuLysGluGluValLys     216

AACATCCAACGTACCGCTGAACTGCTGCTGGAAGTTAACCTTGGCGCAACAGCAATCGGTACTGGTCTG     828
AsnIleGlnArgThrAlaGluLeuLeuLeuGluValAsnLeuGlyAlaThrAlaIleGlyThrGlyLeu     239

AACACGCCGAAAGAGTACTCTCCGCTGGCAGTGAAAAAACTGGCTGAAGTCACTGGCTTCCCATGCGTA     897
AsnThrProLysGluTyrSerProLeuAlaValLysLysLeuAlaGluValThrGlyPheProCysVal     262

CCGGCTGAAGACCTGATCGAAGCGACCTCTGACTGCGGCGCTTATGTTATGGTTCACGGCGCGCTGAAA     966
ProAlaGluAspLeuIleGluAlaThrSerAspCysGlyAlaTyrValMetValHisGlyAlaLeuLys     285

CGCCTGGCTGTGAAGATGTCCAAAATCTGTAACGACCTGCGCTTGCTCTCTTCTGGCCCACGTGCCGGC    1035
ArgLeuAlaValLysMetSerLysIleCysAsnAspLeuArgLeuLeuSerSerGlyProArgAlaGly     308

CTGAACGAGATCAACCTGCCGGAACTGCAGGCGGGCTCTTCCATCATGCCAGCTAAAGTAAACCCGGTT    1104
LeuAsnGluIleAsnLeuProGluLeuGlnAlaGlySerSerIleMetProAlaLysValAsnProVal     331

GTTCCGGAAGTGGTTAACCAGGTATGCTTCAAAGTCATCGGTAACGACACCACTGTTACCATGGCAGCA    1173
ValProGluValValAsnGlnValCysPheLysValIleGlyAsnAspThrThrValThrMetAlaAla     354

GAAGCAGGTCAGCTGCAGTTGAACGTTATGGAGCCGGTCATTGGCCAGGCTATGTTCGAATCCGTTCAC    1242
GluAlaGlyGlnLeuGlnLeuAsnValMetGluProValIleGlyGlnAlaMetPheGluSerValHis     377
```

```
ATTCTGACCAACGCTTGCTACAACCTGCTGGAAAAATGCATTAACGGCATCACTGCTAACAAAGAAGTG  1311
IleLeuThrAsnAlaCysTyrAsnLeuLeuGluLysCysIleAsnGlyIleThrAlaAsnLysGluVal   400

TGCGAAGGTTACGTTTACAACTCTATCGGTATCGTTACTTACCTGAACCCGTTCATCGGTCACCACAAC  1380
CysGluGlyTyrValTyrAsnSerIleGlyIleValThrTyrLeuAsnProPheIleGlyHisHisAsn   423

GGTGACATCGTGGGTAAAATCTGTGCCGAAACCGGTAAGAGTGTACGTGAAGTCGTTCTGGAACGCGGT  1449
GlyAspIleValGlyLysIleCysAlaGluThrGlyLysSerValArgGluValValLeuGluArgGly   446

CTGTTGACTGAAGCGGAACTTGACGATATTTTCTCCGTACAGAATCTGATGCACCCGGCTTACAAAGCA  1518
LeuLeuThrGluAlaGluLeuAspAspIlePheSerValGlnAsnLeuMetHisProAlaTyrLysAla   469

AAACGCTATACTGATGAAAGCGAACAGTAATCGTACAGGGTAGTACAAATAAAGAAGGCACGTCAGATG  1587
LysArgTyrThrAspGluSerGluGln                                        Met     1

ACGTGCCTTTTTTCTTGTGAGCAGTAACTTAAAAATAACAACCTAATATCAACTTGTTAAAAAACAAGG  1656
ThrCysLeuPheSerCysGluGln

AAGGCTAATATGCTAGTTGTAGAACTCATCATAGTTTTGCTGGCGATCTTCTTGGGCGCCAGATTGGGG  1725
        MetLeuValValGluLeuIleIleValLeuLeuAlaIlePheLeuGlyAlaArgLeuGly    20

GGAATAGGTATTGGTTTTGCAGGCGGATTGGGGGTGCTGGTTCTTGCCGCTATTGGCGTTAAACCCGGT  1794
GlyIleGlyIleGlyPheAlaGlyGlyLeuGlyValLeuValLeuAlaAlaIleGlyValLysProGly    43

AACATCCCGTTCGATGTCATTTCCATTATCATGGCGGTTATCGCCGCTATTTCTGCCATGCAGGTTGCT  1863
AsnIleProPheAspValIleSerIleIleMetAlaValIleAlaAlaIleSerAlaMetGlnValAla    66

GGCGGTCTGGACTATCTGGTTCATCAGACAGAAAAGCTGCTGCGCCGTAACCCGAAATACATCACGATC  1932
GlyGlyLeuAspTyrLeuValHisGlnThrGluLysLeuLeuArgArgAsnProLysTyrIleThrIle    89

CTCGCACCGATCGTGACCTATTTCCTGACTATCTTTGCTGGTACTGGCAACATCTCTCTGGCGACACTG  2001
LeuAlaProIleValThrTyrPheLeuThrIlePheAlaGlyThrGlyAsnIleSerLeuAlaThrLeu   112

CCAGTTATCGCTGAAGTTGCGAAGGAACAAGGCGTCAAACCTTGCCGTCCGCTGTCTACTGCAGTGGTA  2070
ProValIleAlaGluValAlaLysGluGlnGlyValLysProCysArgProLeuSerThrAlaValVal   135

TCCGCGCAGATTGCGATCACCGCATCGCCAATCTCAGCGGCAGTGGTTTACATGTCTTCCGTGATGGAA  2139
SerAlaGlnIleAlaIleThrAlaSerProIleSerAlaAlaValValTyrMetSerSerValMetGlu   158

GGTCATGGCATCAGCTACCTCCATCTGCTCTCCGTGGTCATCCCGTCCACCCTGCTGGCGGTTCTGGTG  2208
GlyHisGlyIleSerTyrLeuHisLeuLeuSerValValIleProSerThrLeuLeuAlaValLeuVal   181

ATGTCCTTTCTGGTCACTATGCTGTTCAACTCCAAACTCTCTGACGATCCGATTTATCGCAAGCGTCTG  2277
MetSerPheLeuValThrMetLeuPheAsnSerLysLeuSerAspAspProIleTyrArgLysArgLeu   204

GAAGAGGGCCTGGTTGAACTGCGCGGTGAAAAGCAGATTGAAATCAAATCCGGTGCAAAAACGTCCGTC  2346
GluGluGlyLeuValGluLeuArgGlyGluLysGlnIleGluIleLysSerGlyAlaLysThrSerVal   227

TGGCTGTTCCTGCTGGGCGTAGTTGGCGTGGTTATCTATGCAATCATCAACAGCCCAAGCATGGGTCTG  2415
TrpLeuPheLeuLeuGlyValValGlyValValIleTyrAlaIleIleAsnSerProSerMetGlyLeu   250

GTTGAAAAACCGCTGATGAACACCACCAACGCAATCCTGATCATCATGCTCAGCGTTGCAACTCTGACC  2484
ValGluLysProLeuMetAsnThrThrAsnAlaIleLeuIleIleMetLeuSerValAlaThrLeuThr   273

ACCGTTATCTGTAAAGTCGATACCGACAACATCCTCAACTCCAGCACCTTCAAAGCAGGTATGAGCGCC  2553
ThrValIleCysLysValAspThrAspAsnIleLeuAsnSerSerThrPheLysAlaGlyMetSerAla   296
```

```
TGTATTTGTATCCTGGGTGTTGCGTGGCTGGGCGATACTTTCGTTTCCAACAACATCGACTGGATCAAA  2622
CysIleCysIleLeuGlyValAlaTrpLeuGlyAspThrPheValSerAsnAsnIleAspTrpIleLys   319

GATACCGCTGGTGAAGTGATTCAGGGTCATCCGTGGCTGCTGGCCGTCATCTTCTTCTTTGCTTCTGCT  2691
AspThrAlaGlyGluValIleGlnGlyHisProTrpLeuLeuAlaValIlePhePhePheAlaSerAla   342

CTGCTGTACTCTCAGGCTGCAACCGCAAAAGCACTGATCGCGATGGCTCTGGCACTGAACGTTTCTCGC  2760
LeuLeuTyrSerGlnAlaAlaThrAlaLysAlaLeuIleAlaMetAlaLeuAlaLeuAsnValSerArg   365

TGACGCTGTTGCTTCTTTCGCTGCGGTGTCTGGTCTGTTCATTCTGCCGACCTACCCGACGCTGGTTGC  2829

TGCGGTACAGATGGATGACACGGGTACTACCCGTATCGGTAAATTCGTCTTCAACCATCCGTTCTTCAT  2898

CCC  2901
```

Figure 2. Nucleotide sequences of the aspA gene and its flanking regions and predicted primary structure of aspartase. The nucleotide sequence of the 2.9 kb AvaI-SmaI segment is shown. The predicted amino acid sequence of aspartase is presented below the nucleotide sequence by three-letter designations. The nucleotide positions are assigned relative to the 5'-terminal base of the AvaI recognition sequence. Amino acid residues are numbered relative to the initiating methionine residue. The potential ribosome binding site for aspA is underlined and the putative -10 and -35 promoter sequences for aspA are boxed. Regions of dyad symmetry are shown by arrows above the nucleotide sequence. Sequences of the NH_2- and COOH-terminal regions which are identical to the sequence data of the purified aspartase are underlined. The predicted amino acid sequences of the open reading frames identified downstream of the aspA gene are also shown.

the hybrid plasmid contains the aspA gene.

Location of the aspA Gene on the pA303 Insert

In view of the fact that the molecular weight of aspartase subunit was determined to be 48,500 by sodium dodecylsulfate-polyacrylamide gel electrophoresis (6), the size of the structural gene encoding this protein is expected to be around 1.3 kb. Since the cloned fragment was much larger than the expected size, the location of aspA on this fragment was examined. Plasmids which lacked the 5.3 kb HindIII-HindIII, 5.9 kb BamHI-BamHI, or 2.1 kb EcoRI-EcoRI-EcoRI fragments of pA303 were constructed in vitro. TK237 carrying these plasmids did not grow well on glutamate minimal medium and had no aspartase activity. From these results and the expected size of the gene, we inferred that aspA is located within the AvaI-AvaI/SmaI segment of the insert of pA303.

Nucleotide Sequence of the aspA Gene

pA303 was digested with SmaI plus SalI and a 3.7 kb fragment generated was purified by agarose gel electrophoresis. This fragment was subjected to

nucleotide sequence analysis. The sequencing strategy is shown in Fig. 1. The sequence of AvaI-$PstI_4$ segment, within which aspA lies (see below), was fully overlapped and determined with at least two independent clones. More than 90% of the sequence of this fragment was derived from both strands. Figure 2 shows the complete nucleotide sequence of the AvaI-SmaI segment. Four open reading frames containing more than 400 nucleotides are found (Fig. 1). Among them, only the frame between positions 112 and 1545 which encodes a protein of 477 amino acid residues with a calculated molecular weight of 52,224 (excluding the initiating methionine residue) is compatible with the reported molecular weight of the aspartase subunit. Residues 2-25 of the predicted amino acid sequence of this reading frame are identical with the result obtained from Edman degradation of the purified aspartase protein (N. Yumoto and M. Tokushige, unpublished results; residues 6, 15, 18, 19, and 21-23 were unidentified and amides were not assigned) (Fig. 2). Probably the initiating formylmethionine residue is removed by posttranslational processing. In addition, the COOH-terminal sequence (residues 471-478) is also identical with the sequence data of the COOH-terminal peptide which is released from the purified enzyme protein upon digestion with subtilisin BPN' (10) (Fig. 2). Although the residue preceding the subtilisin BPN' cleavage site was previously shown to be serine by carboxypeptidase Y digestion of the resulting large polypeptide (10), the present study showed that the residue is lysine in the predicted sequence. This discrepancy may be associated with the fact that COOH-terminal lysine residue is not a good substrate for carboxypeptidase Y (19). It is possible that serine was released from some internal sequence due to contaminating endopeptidase activity or from, if any, other subtilisin BPN' cleavage site. Therefore, the predicted amino acid residue is sopposed to be more reliable. Amino acid composition of the predicted sequence fits well to that of the purified aspartase (6) (Table 2). These results led us to conclude that this open reading frame represents the aspA gene.

Putaive Expression Signals of the aspA Gene

A potential promoter for the aspA gene is found at positions 37-42 (TCGAAA) and positions 59-64 (TATATT) which resemble the consensus -35 and -10 sequences, respectively (20). If these sequences function as the promoter for aspA, transcription may start with G at position 71. It is worth noting that the putative -10 sequence is flanked by a sequence of dyad symmetry (Fig. 2). Significance of this structure is not known. In view of the fact that the level of aspartase in E. coli cells is repressed by glucose and elevated under

Table 2. Amino acid composition of aspartase

Amino acid residues	Number of residue / subunit	
	Predicted	Analyzed[a]
Asp	19	56.8[b]
Asn	36	
Thr	24	23.3
Ser	21	19.8
Glu	38	65.2[c]
Gln	19	
Pro	19	20.6
Gly	34	35.2
Ala	42	42.5
Cys	11	10.3
Val	43	41.9
Met	15	13.3
Ile	32	30.8
Leu	44	43.0
Tyr	16	15.4
Phe	13	13.5
His	8	8.1
Lys	27	25.7
Arg	16	17.9
Trp	0	1.1

[a] The values are calculated based on the data of ref. 6 so as to give a molecular weight of 52,224, the value calculated from the predicted sequence.
[b] Value of Asp+Asn
[c] Value of Glu+Gln

anaerobic conditions, it is likely that the aspA gene has some signal(s) related to the regulation of gene expression, possibly in the 5'-flanking region. It remains to be examined, however, whether the signal(s), if any, is included within the sequence analyzed in the present study or in further upstream sequences not cloned in pA303. Five nucleotides upstream from the initiation codon AUG (position 112), there is a stretch of purine-rich sequence (positions 100-107) which is almost complementary (7 out of 8 nucleotides) to the 3'-terminal sequence of E. coli 16S rRNA (21). This region is consistent with the Shine-Dalgarno sequence and may serve as a ribosome binding site when transcribed (21). Downstream of the stop codon UAA (position 1546) there is a sequence of dyad symmetry (positions 1573-1596) which is followed by a T cluster. This structure is characteristic of rho-independent transcriptional terminator (22).

Other Open Reading Frames in the 2.9 kb AvaI-SmaI Segment

Downstream of the putative transcription terminator of the aspA gene, there is a short sequence of dyad symmetry (positions 1623-1646) followed by

an A cluster. Following this region, an open reading frame starts at position 1666 without changing the frame of aspA and is terminated at position 2760. Whether or not this open reading frame is expressed remains to be examined. However, it is noteworthy that the frame is preceded by a purine-rich sequence (positions 1653-1659) which could be a good ribosome binding site. In addition, we also note the presence of a short open reading frame (from AUG codon at position 1585 to UAA codon at position 1612) which is overlapped partially with the postulated terminator for aspA and preceded by a potential ribosome binding site (positions 1570-1576). Provided that this potential gene is expressed, there is a possibility that the preceding two dyad symmetry regions are involved in regulation of its expression. There are also two long open reading frames with opposite polarity. One starts with AUG codon (position 2802) and is terminated by UGA codon (position 2106). A dyad symmetry (positions 1917-1956) exists downstream of this frame. The other is found between AUG (position 522) and UGA codon (position 117) located within the aspA gene sequence. Following this frame, a short region of dyad symmetry exists which corresponds to the putative promoter region for aspA. Transcription study will provide information on whether or not these frames are expressed and the structures described above have any functional significance.

Recently the aspA genes of E. coli K-12 (23, T. Taniguchi, S. Komatsubara, and M. Kisumi, unpublished results) and Serratia marcescens (13) have been cloned. The restriction map of the aspA gene and its flanking regions of E. coli K-12 appears to be grossly similar to that of E. coli W determined in this work. In the case of E. coli K-12, subcloning experiment revealed that the aspA gene locates within a 2.8 kb BclI fragment containing a single EcoRI site (23), which may correspond to either of the two EcoRI sites present in the aspA gene of E. coli W. The subcloned fragment extends from the EcoRI site for 1.7 kb in the direction of the AvaI/SmaI site and for 1.1 kb in the opposite direction (23). In the course of this study, we tried to subclone the 2.9 kb AvaI-SmaI fragment into pBR322. However, the expected hybrid plasmid could not be readily obtained. pBR322-derived plasmids containing E. coli K-12 aspA were found unstable probably due to overproduction of aspartase (T. Taniguchi, S. Komatsubara, and M. Kisumi, unpublished results). Another possibility is that amplification of some sequence(s) other than aspA carried by these plasmids is responsible for the unstability as is suggested in the case of S. marcescens (13). This problem remains to be clarified.

As our and other previous studies showed, aspartase exhibits a complex cooperative behavior in the reaction kinetics and is considered to be an allosteric enzyme (5,6,24). Nevertheless, little is known about the reaction mechanism of the enzyme. The amino acid sequence of aspartase from E. coli W predicted from the nucleotide sequence of the gene will provide a clue to identify the locations of the amino acid residues which are involved in the catalytic and regulatory properties of the enzyme. Site-directed mutagenesis of the cloned gene obtained in this work will be a useful tool to investigate the function-structure relationship of aspartase.

ACKNOWLEDGEMENTS

The authors wish to express their sincere gratitude to Professor H. Katsuki, Drs. K. Izui, N. Fujita, and T. Kodaki of the Department of Chemistry for encouragement and valuable discussions during this investigation. We are grateful to Dr. M. Kisumi, Director, Research Laboratory of Applied Biochemistry, Tanabe Seiyaku Co., for providing us the mutants and communicating his unpublished results and to Professor T. Murachi, Department of Clinical Science, Faculty of Medicine, Kyoto University, for giving us the opportunity to use a protein sequencer for NH_2-terminal sequence analysis. This investigation was supported in part by Grants-in-Aid from the Ministry of Education, Science and Culture of Japan.

REFERENCES

1. Williams, V.R. and Lartigue, D.J. (1969) Methods Enzymol. 13, 354-361
2. Halpern, Y.S. and Umbarger, H.E. (1960) J. Bacteriol. 80, 285-288
3. Marcus, M. and Halpern, Y.S. (1969) Biochim. Biophys. Acta 177, 314-320
4. Courtright, J.B. and Henning, U. (1970) J. Bacteriol. 102, 722-728
5. Rudolph, F.B. and Fromm, H.J. (1971) Arch. Biochem. Biophys. 147, 92-98
6. Suzuki, S., Yamaguchi, J., and Tokushige, M. (1973) Biochim. Biophys. Acta 321, 369-381
7. Takagi, J.S., Fukunaga, R., Tokushige, M., and Katsuki, H. (1984) J. Biochem. 96, 545-552
8. Watanabe, Y., Iwakura, M., Tokushige, M., and Eguchi, G. (1981) Biochim. Biophys. Acta 661, 261-266
9. Yumoto, N., Tokushige, M., and Hayashi, R. (1980) Biochim. Biophys. Acta 616, 319-328
10. Yumoto, N., Mizuta, K., Tokushige, M., and Hayashi, R. (1982) Physiol. Chem. Phys. 14, 391-397
11. Mizuta, K. and Tokushige, M. (1975) Biochim. Biophys. Acta 403, 221-231
12. Ida, N. and Tokushige, M. (1984) J. Biochem. 96, 1315-1321
13. Takagi, T. and Kisumi, M. (1985) J. Bacteriol. 161, in press
14. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) Molecular Cloning, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York
15. Cosloy, S.D. and Oishi, M. (1973) Mol. Gen. Genet. 124, 1-10
16. Norrander, J., Kempe, T., and Messing, J. (1983) Gene 26, 101-106

17. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467
18. Lowry, O.H., Rosebrough, N.J., Farr, A.L., and Randall, R.J. (1951) J. Biol. Chem. 193, 265-275
19. Hayashi, R. (1976) Methods Enzymol. 45, 568-587
20. Hawley, D.K. and McClure, W.R. (1983) Nucleic Acids Res. 11, 2237-2255
21. Shine, J. and Dalgarno, L. (1974) Proc. Natl. Acad. Sci. USA 71, 1342-1346
22. von Hippel, P.H., Bear, D.G., Morgan, W.D., and McSwiggen, J.A. (1984) Annu. Rev. Biochem. 53, 389-446
23. Guest, J.R., Roberts, R.E., and Wilde, R.J. (1984) J. Gen. Microbiol. 130, 1271-1278
24. Williams, V.R. and Lartigue, D.J. (1967) J. Biol. Chem. 242, 2973-2978