Vol. 133, No. 1, 1985 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
November 27, 1985 Pages 134-139

AROMATIC AMINO ACID AMINOTRANSFERASE OF *Escherichia coli*: NUCLEOTIDE SEQUENCE OF THE *tyrB* GENE

Seiki Kuramitsu, Katsura Inoue, Tomoko Ogawa*, Hideyuki Ogawa*, and Hiroyuki Kagamiyama

Department of Medical Chemistry, Osaka Medical College, 2-7 Daigakucho, Takatsuki-shi, Osaka 569, Japan

*Department of Biology, Faculty of Science, Osaka University, Toyonaka-shi, Osaka 560, Japan

Received October 11, 1985

SUMMARY The *tyrB* gene of *E. coli* K-12, which encodes aromatic amino acid aminotransferase (EC 2.6.1.57) was cloned. The nucleotide sequence of about 2 kilobase pairs containing the gene was determined. The coding region of the *tyrB* gene and the deduced amino acid sequence revealed that the aromatic amino acid aminotransferase of *E. coli* is homologous with the aspartate aminotransferase.

Rudman and Meister described a "transaminase A" in *E. coli*, which catalyzes transamination reactions utilizing aspartate and aromatic amino acids as the amino donor (1). Subsequently several reports (2-5) indicated that two aminotransferases with overlapping substrate specificity were separable from the extract of *E. coli*; the one is heat-stable and active predominantly on aspartate, and the other is heat-unstable, repressed by tyrosine and active predominantly on aromatic amino acids. Genetic studies of Gelfand and his coworkers (6, 7) showed that "transaminase A" was a mixture of the two enzymes; the tyrosine-repressible aromatic amino acid aminotransferase and the non-repressible aspartate aminotransferase, encoded by two distinct genes, *tyrB* and *aspC*, respectively, and that when both of the two genes were mutated, no aromatic amino acid or aspartate aminotransferase activity was found in the cell extract. They also demonstrated that the two enzymes showed overlapping specificities for tyrosine, phenylalanine and aspartate.

The nucleotide sequence of the *aspC* gene (8) and primary structure of aspartate aminotransferase (9) of *E. coli* has been reported. This paper

0006-291X/85 $1.50

describes the nucleotide sequence of the tyrB gene and the deduced amino acid sequence of aromatic amino acid aminotransferase.

MATERIALS AND METHODS

Bacterial Strains and Phage. E. coli DG44 (aspC, tyrB507, proA2, his-4, argE3, thi-1, hsdS14, hppT29, lacY1, galK2, xyl-5, mtl-1, rpsL31, tsx-33, supE44, recB21, recC22, sbcB15, λ^-) (6, 7) was obtained from the E. coli Genetic Stock Center (Yale University) and used for the selection of $tyrB^+$ plasmids. E. coli JM 103 (Δlacpro, supE, thi, strA, sbcB15, endA, hsdR4, F'traD36, proAB, $LacI^qZ\Delta M15$) (10) was used to assay the expression of the lac DNA of the M13mp18 or M13mp19 phage (11).

Chemicals. Restriction endonucleases were obtained from Takara Shuzo Co., Kyoto, Nippon Gene Co., Toyama, and Toyobo Co., Osaka. Ampicillin and 5-bromo-4-chloro-3-indolylgalactoside (Xgal) were obtained from Sigma Chemical Co., St. Louis, isopropylthiogalactoside (IPTG) from Bethesda Research Laboratory, Neu-Isenberg, a M13 phage sequence kit from Takara Shuzo Co., Kyoto, and [α-^{32}P] dCTP (400Ci/mmol) from Amersham.

Media. The media used for the bacterial growth and for the selection of the clones carrying the $tyrB^+$gene plasmids were the same as described previously (8).

Determination of Nucleotide Sequence. Nucleotide sequences of DNA fragments were determined by the method of Messing (12) as described previously (8).

Purification and N-Terminal Amino Acid Sequence Analysis of Aromatic Amino Acid Aminotransferase. Aromatic amino acid aminotransferase was purified to homogeneity from E. coli W3110 (K-12) by the consecutive use of chromatographies on DEAE-cellulose at pH 7.0 and 5.5, Sephacryl S-200, phenyl-Sepharose CL-4B, pyridoxal phosphate-Sepharose, and hydroxyapatite. Aspartate aminotransferase was separated completely from aromatic amino acid aminotransferase by the phenyl-Sepharose chromatography. The details of purification procedures and properties of the enzyme will be described elsewhere. The amino acid sequence of the N-terminal 10 residues of the purified enzyme was determined by the manual Edman method (13). Phenylthiohydantoin derivatives of amino acids were identified by high performance liquid chromatography on a LS-410K C-18 column (Toyo Soda Co.) (14).

RESULTS AND DISCUSSION

Cloning of DNA Fragments Which Complement The tyrB Mutation. The chromosomal DNA of wild type E. coli W3110 was prepared as described by Clarke and Carbon (15), and partially digested with the restriction enzyme, EcoRI or HindIII. The DNA fragments not less than 1 kilobase pairs long were collected on a sucrose density gradient (16) and ligated with EcoRI- or HindIII-cleaved pMF3 (17). The transformants that restore prototrophy for tyrosine and aspartate were selected on minimal agar plates lacking the two amino acids, and then subjected to the ampicillin (20 μg/ml) selection. The plasmid DNAs were prepared according to the procedures described by Clewell and Heinsky (18). Two plasmids were found to confer the $tyrB^+$ phenotype. The one contained the EcoRI fragment with 11 kilobase pairs, and the other the HindIII fragment with

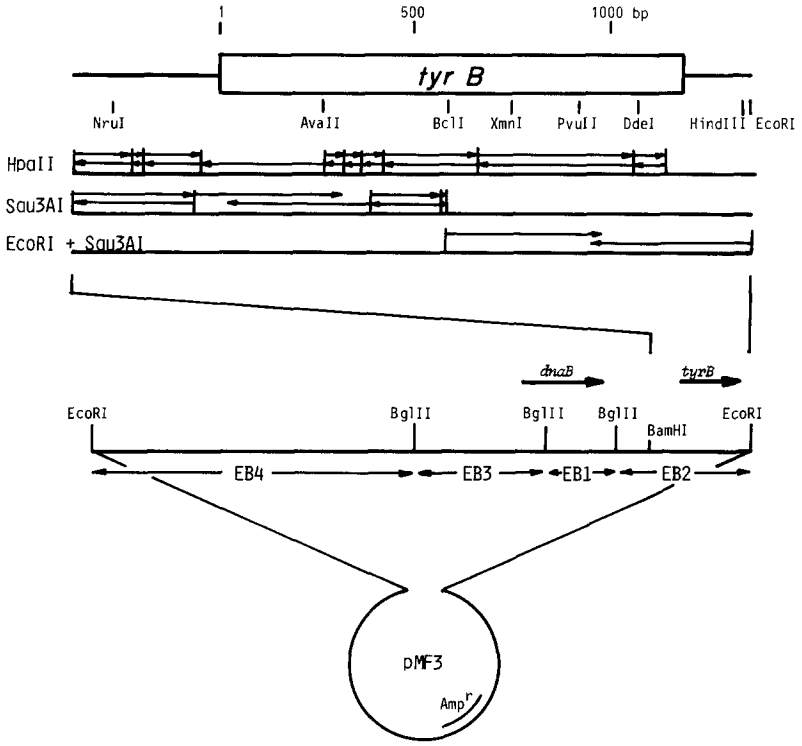

Figure 1. Cleavage maps and sequencing strategy for the tyrB gene. The direction and the extent of the sequence determination of the fragments are shown by arrows. An arrow indicates the 5' to 3' direction. Numbers at the top indicate the distance from the initiation site of the tyrB gene. bp, base pairs.

17 kilobase pairs. The shorter plasmid (pKYA-1) containing the EcoRI fragment was used for further studies.

Nucleotide Sequence of The tyrB Gene and Predicted Amino Acid Sequence of Aromatic Amino Acid Aminotransferase. The restriction map of the pKYA-1 plasmid was shown in Fig. 1. The EcoRI fragment with 11 kilobase pairs was digested with BglII. Four fragments with 5.5, 2.3, 2.2 and 1.2 kilobase pairs were isolated by agarose gel electrophoresis and were named EB1, EB2, EB3 and EB4, respectively. They were further digested with restriction endonuclease HpaII or Sau3AI, and the digests were inserted into AccI- or BamHI-cleaved M13mp19 phage for sequence analysis (12). The manual Edman degradation of the isolated aromatic amino acid aminotransferase established the amino acid sequence of the aminoterminal 10 residues of the enzyme as Met-Phe-Gln-Lys-Val-Asp-Ala-Tyr-Ala-Gly. A comparison of the deduced amino acid sequence from

the determined nucleotide sequences with the established aminoterminal amino acid sequence allowed us to identify the tyrB reading frame in EB2. The nucleotide sequence around the tyrB gene in EB2 was determined according to the strategy presented in Fig. 1. Since the 3'-terminal portion of the tyrB gene exists near the EcoRI site of EB2, the DNA fragments obtained by the Sau3AI digestion of EB2 were inserted into BamHI- and EcoRI-cleaved M13mp18 or 19 phage and the nucleotide sequences were determined.

The nucleotide sequence of the tyrB gene and the deduced amino acid sequence are presented in Fig. 2. The GTG codon was used for initiation in the tyrB gene. Only one open frame from nucleotide +1 to +1,191 was found to specify a protein of a sufficient length of aromatic amino acid aminotransferase. The termination codon TAA was found at the position +1,192-+1,194. The polypeptide encoded by the gene consists of 397 amino acid residues and the predicted amino acid sequence well explains the previously estimated molecular weight and isoelectric point of aromatic amino acid aminotransferase (2-5). The codon usage of the tyrB gene was similar to those for the proteins in E. coli (19).

The tetranucleotide sequence from -13 to -10, 5'GGAG3' , probably serves as the ribosome binding site. Dyad symmetry was found in the regions from nucleotides -315 to -304, from -258 to -251, from -226 to -197, from -77 to -59, and from -48 to -30. The GC-rich palindrome from nucleotides -226 to -197 was followed by a stretch of consecutive AT base pairs, which is characteristic of rho-independent transcription termination (20). An open frame from nucleotide -127 to -35 may encode a leader peptide having phenylalanine-rich sequence at the C-terminal portion (Met-Tyr-Val-Cys-His-Glu-Ser-His-Ser-Val-Ala-Asn-Cys-Arg-Ser-Leu-Leu-Asn-Ile-His-Ser-Ile-Phe-Ala-Phe-Phe-Arg-Phe-Ile-Val-Phe). In the partially determined nucleotide sequences of EB3 and EB4, the coding region of the dnaB gene (21), which had been known to be close to the tyrB gene on the linkage map of E. coli (22), was found (Fig. 1, data not shown).

```
                                                                    GATCCGGTCATTTTA
TGGGGCGAAGGTTTGCCCGTAGAACGTATCGCTGAAATGACGAAAGTAAGCGCTTACGAACTTATTACGCGCCTGACTTCAAGGGTCGCG
ATGAAATACGTGGATTAATCGTTCTGTAATATTTGATTGTCTGTGCCGGATGCGGCGTGAATGCCTTATCCGGCCAATAAAATCCTAAAA
ATTCAATAAGTTGATGTTCTTTCATGCTCTTATAAAGGTCGTGCCTCTGGCGGATGTACGTTTGTCATGAGTCTCACTCTGTTGCTAATT
GCCGTTCGCTCCTGAACATCCACTCGATCTTCGCCTTCTTCCGGTTTATTGTGTTTTAACCACCTGCCCGTAAACCTGGAGAACCATCGC

        10         20         30         40         50         60         70         80         90
GTGTTTCAAAAAGTTGACGCCTACGCTGGCGACCCGATTCTTACGCTTATGGAGCGTTTTAAAGAAGACCCTCGCAGCGACAAAGTGAAT
M  F  Q  K  V  D  A  Y  A  G  D  P  I  L  T  L  M  E  R  F  K  E  D  P  R  S  D  K  V  N

       100        110        120        130        140        150        160        170        180
TTAAGTATCGGTCTGTACTACAACGAAGACGGAATTATTCCACAACTGCAAGCCGTGGCGGAGGCGGAAGCGCGCCTGAATGCGCAGCCT
L  S  I  G  L  Y  Y  N  E  D  G  I  I  P  Q  L  Q  A  V  A  E  A  E  A  R  L  N  A  Q  P

       190        200        210        220        230        240        250        260        270
CATGGCGCTTCGCTTTATTTACCGATGGAAGGGCTTAACTGCTATCGCCATGCCATTGCGCCGCTGCTGTTTGGTGCGGACCATCCGGTA
H  G  A  S  L  Y  L  P  M  E  G  L  N  C  Y  R  H  A  I  A  P  L  L  F  G  A  D  H  P  V

       280        290        300        310        320        330        340        350        360
CTGAAACAACAGCGCGTAGCAACCATTCAAACCCTTGGCGGCTCCGGGGCATTGAAAGTGGGCGCGGATTTCCTGAAACGCTACTTCCCG
L  K  Q  Q  R  V  A  T  I  Q  T  L  G  G  S  G  A  L  K  V  G  A  D  F  L  K  R  Y  F  P

       370        380        390        400        410        420        430        440        450
GAATCAGGCGTCTGGGTCAGCGATCCTACCTGGGAAAACCACGTAGCAATATTCGCCGGGGCTGGATTCGAAGTGAGTACTTACCCCTGG
E  S  G  V  W  V  S  D  P  T  W  E  N  H  V  A  I  F  A  G  A  G  F  E  V  S  T  Y  P  W

       460        470        480        490        500        510        520        530        540
TATGACGAAGCGACTAACGGCGTGCGCTTTAATGACCTGTTGGCGACGCTGAAAACATTACCTGCCCGCAGTATTGTGTTGCTGCATCCA
Y  D  E  A  T  N  G  V  R  F  N  D  L  L  A  T  L  K  T  L  P  A  R  S  I  V  L  L  H  P

       550        560        570        580        590        600        610        620        630
TGTTGCCACAACCCAACGGGTGCCGATCTCACTAATGATCAGTGGGATGCGGTGATTGAAATTCTCAAAGCCCGCGAGCTTATTCCATTC
C  C  H  N  P  T  G  A  D  L  T  N  D  Q  W  D  A  V  I  E  I  L  K  A  R  E  L  I  P  F

       640        650        660        670        680        690        700        710        720
CTCGATATTGCCTATCAAGGATTTGGTGCCGGTATGGAAGAGGATGCCTACGCTATTCGCGCCATTGCCAGCGCTGGATTACCCGCTCTG
L  D  I  A  Y  Q  G  F  G  A  G  M  E  E  D  A  Y  A  I  R  A  I  A  S  A  G  L  P  A  L

       730        740        750        760        770        780        790        800        810
GTGAGCAATTCGTTCTCGAAAATTTTCTCCCTTTACGGCGAGCGCGTCGGCGGACTTTCTGTTATGTGTGAAGATGCCGAAGCCGCTGGC
V  S  N  S  F  S  K  I  F  S  L  Y  G  E  R  V  G  G  L  S  V  M  C  E  D  A  E  A  A  G

       820        830        840        850        860        870        880        890        900
CGCGTACTGGGGCAATTGAAAGCAACAGTTCGCCGCAACTACTCCAGCCCGCCGAATTTTGGTGCGCAGGTGGTGGCTGCAGTGCTGAAT
R  V  L  G  Q  L  K  A  T  V  R  R  N  Y  S  S  P  P  N  F  G  A  Q  V  V  A  A  V  L  N

       910        920        930        940        950        960        970        980        990
GACGAGGCATTGAAAGCCAGCTGGCTGGCGGAAGTAGAAGAGATGCGTACTCGCATTCTGGCAATGCGTCAGGAATTGGTGAAGGTATTA
D  E  A  L  K  A  S  W  L  A  E  V  E  E  M  R  T  R  I  L  A  M  R  Q  E  L  V  K  V  L

      1000       1010       1020       1030       1040       1050       1060       1070       1080
AGCACAGAGATGCCAGAACGCAATTTCGATTATCTGCTTAATCAGCGCGGCATGTTCAGTTATACCGGTTTAAGTGCCGCTCAGGTTGAC
S  T  E  M  P  E  R  N  F  D  Y  L  L  N  Q  R  G  M  F  S  Y  T  G  L  S  A  A  Q  V  D

      1090       1100       1110       1120       1130       1140       1150       1160       1170
CGACTACGTGAAGAATTTGGTGTCTATCTCATCGCCAGCGGTCGCATGTGTGTCGCCGGGTTAAATACGGCAAATGTACAACGTGTGGCA
R  L  R  E  E  F  G  V  Y  L  I  A  S  G  R  M  C  V  A  G  L  N  T  A  N  V  Q  R  V  A

      1180       1190
AAGGCGTTTGCTGCGGTGATGTAATGCAGGAAAGCAGGCTGGAGCTACCCAGCCTGCAGTGAAATTAAACTGTCGTCGCTTTCACTCTTT
K  A  F  A  A  V  M  *

CTTTATAGATGATTTTTTTGATGCCATCGTTCTACGTGAGAGATAATAAACGTTGTTAGTTCTTTTATTGTTAAGCTTATCCCAATTATC

TGGAATTC
```

Figure 2. Nucleotide sequence of the tyrB gene and its neighbourhood, and amino acid sequence of the encoded protein. The DNA sequence is numbered beginning at the initiation site of the tyrB gene. *, termination codon.

Results in this and the previous papers (8, 9) indicate that aromatic amino acid aminotransferase and aspartate aminotransferase of E. coli are

homologous proteins with the sequence homologies of 51 % (nucleotide) and 44 % (amino acid). Interestingly, most of the amino acid residues that was predicted to be important for the catalysis of aspartate aminotransferase (23-25) are detected at the corresponding positions in the aromatic amino acid aminotransferase. These findings may explain the overlapping substrate specificity of the two enzymes.

ACKNOWLEDGEMENT

We are grateful to Dr. Y. Yamamoto of Hyogo Medical College for his valuable advice and technical help. Thia work was supported in part by Grants for Scientific Research from the Ministry of Education, Science and Culture of Japan.

REFERENCES

1. Rudman, D., and Meister, A. (1953) J. Biol. Chem., 200, 591-604
2. Collier, R. H., and Kohlhaw, G. (1972) J. Bacteriol., 112, 365-371
3. Mavrides, C., and Orr, W. (1974) Biochim. Biophys. Acta, 336, 70-78
4. Mavrides, C., and Orr, W. (1975) J. Biol. Chem., 250, 4128-4133
5. Powell, J. T., and Morrison, F. (1978) Eur. J. Biochem., 87, 391-400
6. Gelfand, D. H., and Steinberg, R. A. (1977) J. Bacteriol., 130, 429-440
7. Gelfand, D. H., and Rudo, N. (1977) J. Bacteriol., 130, 441-444
8. Kuramitsu, S., Okuno, S., Ogawa, T., Ogawa, H., and Kagamiyama, H. (1985) J. Biochem., 97, 1259-1262
9. Kondo, K., Wakabayashi, S, Yagi, T., and Kagamiyama, H. (1984) Biochem. Biophys. Res. Commun., 122, 62-67
10. Messing, J., Crea, R., and Seeburg, P. H. (1981) Nucl. Acids Res., 9, 309-321
11. Norrander, L., Kempe, T., and Messing, J. (1983) Gene, 26, 101-106
12. Messing, J. (1983) Methods Enzymol., 101, 20-78
13. Edman, P., and Henschen, A. (1975) Protein Sequence Determination (Needleman, S. B., ed.) 2nd ed., pp.232-262, Springer-Verlag, Heidelberg
14. Zimmerman, C. L., Appella, E., and Pisano, J. J. (1977) Anal. Biochem., 77, 569-573
15. Clarke, L., and Carbon, J. (1979) Methods Enzymol., 68, 396-408
16. Ogawa, T., and Tomizawa, J. (1967) J. Mol. Biol., 23, 265-276
17. Manis, J. J., and Kline, B. C. (1977) Molec. Gen. Genet., 152, 175-182
18. Clewell, D. B., and Helinsky, D. R. (1969) Proc. Natl. Acad. Sci. USA, 62, 1159-1166
19. Guoy, M., and Gautier, C. (1982) Nucl. Acids Res., 22, 429-444
20. Rosenberg, M., and Court, D. (1979) Ann. Rev. Genet., 13, 319-353
21. Nakayama, N., Arai, N., Bond, M., Kaziro, Y., and Arai, K. (1984) J. Biol. Chem., 259, 97-101
22. Bachmann, B. J. (1983) Microbiol. Rev., 47, 180-230
23. Jansonius, J. N., Eichele, G., Ford, G. C., Picot, D., Thaller, C., and Vincent, M. G. (1985) Transaminases (Christen, P., and Metzler, D. E., eds.) pp. 110-138, John-Wiley and Sons, New York
24. Arnone, A., Rogers, P. H., Hyde, C. C., Briley, P. D., Metzler, C. M., and Metzler, D. E. (1985) ibid, pp.138-155
25. Harutyunyan, E. G., Malashkevich, V. N., Kochkina, V. M., and Torchinsky, Y. M. (1985) ibid, pp.164-173