Biochem. J. (1990) 270, 651–657 (Printed in Great Britain)

The amino acid sequence of cytosolic aspartate aminotransferase from human liver

John M. DOYLE,* M. Eugenia SCHININÀ,† Francesco BOSSA† and Shawn DOONAN*‡

*Department of Biochemistry, University College, Cork, Ireland, and †Dipartimento di Scienze Biochimiche e C.N.R. Centro di Biologia Molecolare, Università La Sapienza, 00185 Roma, Italy

1. The cytosolic aspartate aminotransferase was purified from human liver. 2. The isoenzyme contains four cysteine residues, only one of which reacts with 5,5′-dithiobis-(2-nitrobenzoic acid) in the absence of denaturing agents. 3. The amino acid sequence of the isoenzyme is reported, as determined from peptides produced by digestion with trypsin and with CNBr, and from sub-digestion of some of these peptides with *Staphylococcus aureus* V8 proteinase. 4. The isoenzyme shares 48% identity of amino acid sequence with the mitochondrial form from human heart. 5. Comparisons of the amino acid sequences of all known mammalian cytosolic aspartate aminotransferases and of the same set of mitochondrial isoenzymes are reported. The results indicate that the cytosolic isoenzymes have evolved at about 1.3 times the rate of the mitochondrial forms. 6. The time elapsed since the cytosolic and mitochondrial isoenzymes diverged from a common ancestral protein is estimated to be 860×10^6 years. 7. Experimental details and confirmatory data for the results presented here are given in a supplementary paper that has been deposited as a Supplementary Publication SUP 50158 (25 pages) at the British Library Document Supply Centre, Boston Spa, Wetherby, West Yorkshire LS23 7BQ, U.K., from whom copies can be obtained on the terms indicated in Biochem. J. (1990) **265**, 5.

INTRODUCTION

The cytosolic and mitochondrial aspartate aminotransferases (EC 2.6.1.1) are coded for by different, but structurally related, genes. This was initially shown by amino acid sequence analysis of the two isoenzymes from pig heart [1–3]. Subsequently a considerable amount of work has been done on structure determination of aspartate aminotransferases from various sources with the use of both classical protein sequencing methods and, more recently, cDNA sequencing [4–12].

The results of this work are important in that they provide insights into the evolutionary processes that have given rise to the aspartate aminotransferases.

An interesting, but open, question in this context has been whether the two isoenzymes have evolved at equal rates, or whether the specific requirement for import of the mitochondrial isoenzyme from its site of synthesis in the cell cytosol into the organelles has imposed extra constraints on its structure and hence on its rate of evolution. Early indirect evidence based on immunochemical comparisons suggested that mammalian cytosolic aspartate aminotransferases have evolved at about twice the rate of their mitochondrial counterparts [13], whereas comparisons of limited sequence data indicated little, if any, difference in rate [14]. More recently [9] we presented a quantitative comparison of the sequence differences for all aspartate aminotransferases whose structures had been determined. These results suggested a nearly equal rate of evolution of the mammalian isoenzymes, but were not conclusive since the sequences of only two mammalian cytosolic aspartate aminotransferases were known at that time. Consequently, we decided to extend the list of known sequences by determining that of the cytosolic isoenzyme from human liver; the results are reported in the present paper. In the meanwhile, amino acid sequences of the cytosolic aspartate aminotransferases from mouse [10] and rat [11], based on cDNA sequences, have been published, allowing for a more refined comparison of sequence differences; these comparisons are also presented here.

MATERIALS AND METHODS

Materials

CNBr and fluoren-9-ylmethyloxycarbonyl chloride were from Aldrich Chemical Co., dansyl-amino acids were from Mann Research Laboratories, phenyl isothiocyanate and amino acid standards were from Pierce Chemical Co., h.p.l.c.-grade solvents, pyridine and trifluoroacetic acid were from Rathburn Chemical Co., and all proteolytic enzymes, amino acid phenylthiohydantoin derivatives and most other materials were from Sigma Chemical Co. Reagents used for gas-phase sequencing were supplied by Applied Biosystems.

Enzyme purification

Human liver was obtained at autopsy and stored at −20 °C until use. Aspartate aminotransferase was isolated by modification of the method previously reported [15], with final purification on Cibacron Blue–Sepharose CL-6B [16]. Details are given in Supplementary Publication SUP 50158.

Structure analysis

Experimental details of the following procedures are described in Supplementary Publication SUP 50158: carboxymethylation of the enzyme; cleavage of carboxymethylated protein with trypsin and with CNBr; peptide fractionation by gel filtration; peptide purification by reverse-phase h.p.l.c.; sub-digestion of purified peptides with *Staphylococcus aureus* V8 proteinase; *N*-terminal analysis by dansylation; *C*-terminal analysis by digestion with carboxypeptidase Y; amino acid analysis; sequence analysis by the dansyl–Edman method and by automated gas-phase methods.

Thiol group analysis

This was done by the method of Ellman [17].

Native enzyme (2 mg) was incubated at 25 °C for 2 h in 0.1 M-sodium phosphate buffer, pH 8.0, containing 2.66 mM-

‡ To whom correspondence should be addressed.

dithiothreitol and 1.3 mM-EDTA. Reagents were removed by gel filtration through a Sephadex G-25 (superfine grade) column (1.5 cm × 11 cm) equilibrated in 0.1 M-sodium phosphate buffer, pH 8.0. To the fraction containing enzyme was added 0.4 mg of 5,5′-dithiobis-(2-nitrobenzoic acid), and the absorbance change was measured at 412 nm. Then 300 μl of 10% (w/v) SDS was added, and the further absorbance change was measured.

The protein concentration was measured by the method of Lowry *et al.* [18], with BSA as standard.

Sequence comparisons

Alignment of amino acid sequences was done by using the computer program of Lipman & Pearson [19]. Quantitative comparisons of sequences were made in terms of the parameter K_{aa} as defined by Kimura [20], which provides an estimate of the average number of substitutions per site between two sequences.

RESULTS AND DISCUSSION

Purification of cytosolic aspartate aminotransferase from human liver

Results of a typical purification are given in Table 1, starting from 550 g of liver. The yield is based on activity after the heat-treatment step, since the initial homogenate also contained the mitochondrial isoenzyme, which is heat-labile [15]. In fact, not all of the mitochondrial isoenzyme was destroyed by heat treatment. When the CM-cellulose column was washed with 1.0 M-NaCl after elution of cytosolic isoenzyme with 0.1 M-NaCl, then a further 2000 units of activity were eluted; this activity was due to the mitochondrial isoenzyme and accounts for the relatively poor yield from this step.

An overall yield of 15 mg of protein was obtained. The product was pure as judged by SDS/PAGE and by cellulose acetate electrophoresis at pH 8.6 (results not shown). In addition, the enzyme was subjected to *N*-terminal analysis by the dansyl method; a single *N*-terminal residue, alanine, was observed.

Hence the purification method described yielded a cytosolic aspartate aminotransferase in sufficient quantity and state of purity for sequence analysis.

Amino acid analysis

The amino acid composition of oxidized enzyme was determined after hydrolysis for 24, 48 and 72 h. The numbers of residues found per subunit (M_r 46000) are given in Table 2. For most residues the values are the means of three determinations. In the cases of serine and threonine the values are those extrapolated back to zero time of hydrolysis. The value for methionine was determined on an unoxidized protein sample. Tryptophan was not determined.

Table 1. Purification of cytosolic aspartate aminotransferases from human liver

Step	Activity (units)	Protein (mg)	Specific activity (units/mg)	Yield (%)
Homogenate	22450	n.d.	n.d.	–
Heat step	10743	n.d.	n.d.	(100)
$(NH_4)_2SO_4$ fractionation	9684	n.d.	n.d.	90
CM-cellulose CM23	5850	1955	3.0	55
DEAE-cellulose DE52	4433	526	8.4	41
CM-Sephadex C-50	3831	75	51.1	36
Blue Sepharose CL-6B	2774	15	185.0	26

The values are generally in acceptable agreement with those calculated from the amino acid sequence (see below).

Thiol group analysis

The native enzyme was reacted with 5,5′-dithiobis-(2-nitrobenzoic acid) in the presence and in the absence of SDS. In the absence of SDS, somewhat less than one residue of cysteine could be titrated. This is similar to the behaviour with the bovine enzyme [21], but differs from that of the enzymes from pig and horse [22], where two residues react rapidly with the reagent. The rapidly reacting residues in the pig enzyme are Cys-45 and Cys-82 [23], both of which are present in the enzyme from human liver (see below). The reason why one of these residues is unreactive in the human isoenzyme is not known.

In the presence of SDS 3.3 residues of cysteine could be titrated, whereas the number of cysteine residues determined from sequence analysis is 4. The cytosolic aspartate aminotransferases from pig and horse contain five cysteine residues, one of which, Cys-390, shows the interesting property of increased reactivity in the presence of substrates [22]. It is notable that position 390 is occupied by a serine residue in the isoenzyme from human liver (see below), emphasizing that Cys-390 in other cytosolic aspartate aminotransferases plays no special part in catalysis.

Peptides from digestion with trypsin

Carboxymethylated protein (35 mg) was digested with trypsin, and the peptides produced were partially purified by gel filtration through Sephadex G-25 (superfine grade). Fractions were pooled on the basis of *N*-terminal analysis and analytical t.l.c. on cellulose plates with the solvent system butan-1-ol/acetic acid/

Table 2. Amino acid composition of human liver cytosolic aspartate aminotransferase

The analyses were performed on oxidized samples of protein except for methionine, where an unoxidized sample was used. Values are the means (range ±10% or better) of the three determinations after hydrolysis for 24, 48 and 72 h except for serine and threonine, for which values were obtained by extrapolation back to zero time of hydrolysis. Tryptophan was not determined (N.B.).

	Amino acid composition (mol/mol of protein)	
Amino acid	Amino acid analysis	Sequence
Cysteine	5	4
Arginine	21	24
Serine	27	27
Aspartic acid/asparagine	37	44
Glutamic acid/glutamine	44	41
Threonine	23	21
Glycine	32	28
Alanine	35	34
Proline	21	24
Methionine	4	5
Valine	28	31
Phenylalanine	20	23
Isoleucine	18	19
Leucine	39	37
Histidine	9	8
Lysine	20	21
Tyrosine	11	12
Tryptophan	N.D.	9
Total		412

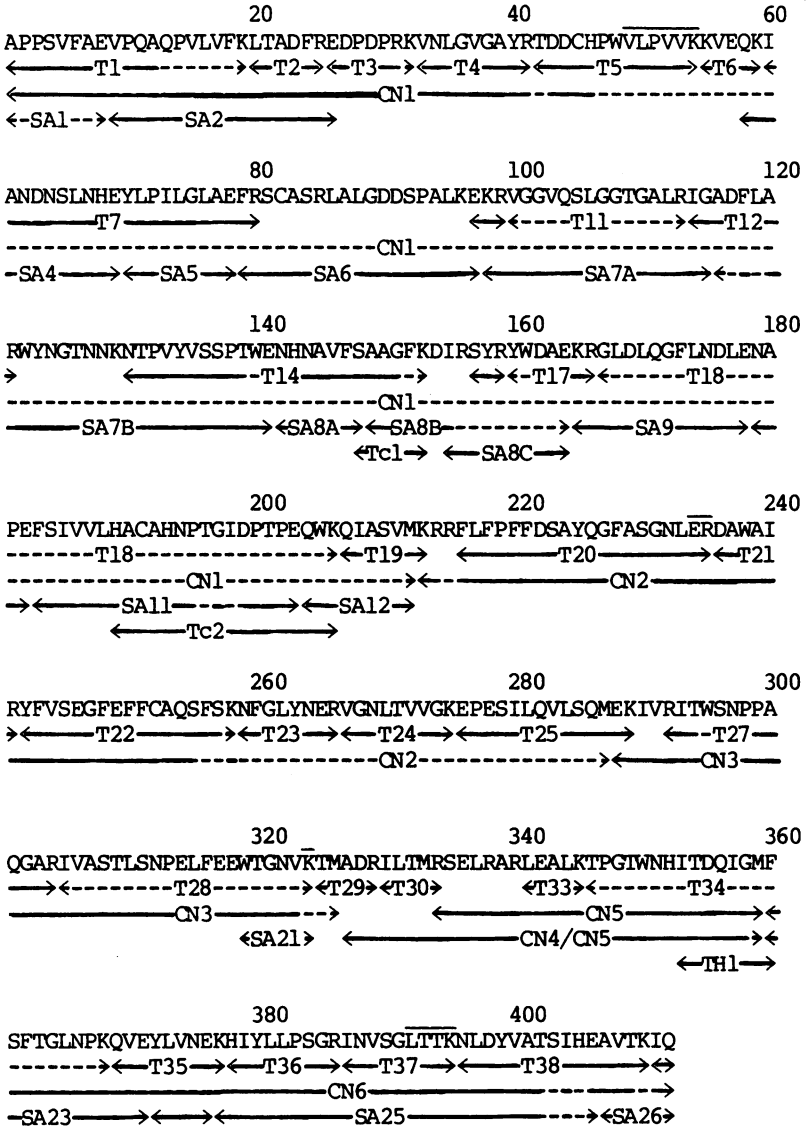

Fig. 1. Amino acid sequence of cytosolic aspartate aminotransferase from human liver

Underlinings show the peptides on which the sequence was based; dashed underlinings show peptides or parts of peptides for which only amino acid compositions were obtained. Overlinings show residues identified by digestion of peptides with carboxypeptidase Y. T, CN, SA and TH signify peptides produced by digestion with trypsin, CNBr, *S. aureus* V8 proteinase and thermolysin respectively. TC1 and TC2 signify peptides produced by chymotryptic-like cleavage in the digestion with trypsin. In numbering the sequence the first digit is placed over the residue specified by that number.

water/pyridine (15:3:12:10, by vol.) with detection by fluorescamine. Thirteen pooled fractions were taken.

Peptides in the pooled fractions were further purified by h.p.l.c. on C_{18} reverse-phase columns. A total of 34 peptides were isolated, two of which arose from chymotryptic-like cleavage (at leucine and phenylalanine). Purity of peptides was assessed by *N*-terminal analysis.

All peptides were subjected to amino acid analysis, and the majority of them were sequenced, either partially or completely, by the dansyl–Edman method (experimental data are given in Supplementary Publication SUP 50158). The positions of the peptides isolated are shown in Fig. 1, with the peptides numbered in order from the *N*-terminus. Some peptides expected on the basis of the amino acid sequence (e.g. T-8 and T-9) were not isolated.

Some of the peptides were subjected to digestion with carboxypeptidase Y, and the amino acids liberated were identified as a function of time of hydrolysis. The results obtained are summarized in Fig. 1.

Peptides from digestion with CNBr

Carboxymethylated protein (20 mg) was digested with CNBr, and the products were partially purified by gel filtration through Sephadex G-50 (fine grade). Two peptides (CN-3 and CN-6) were obtained in a pure state in pools P4 and P3 respectively from gel filtration. Peptides CN-1 and CN-2 were co-eluted in pool P1, and were separated by gel filtration through Sephadex G-75 (superfine grade). Peptide CN-5 was purified from pool P5 by h.p.l.c. along with a peptide (Ala-327–Met-359) resulting from failure to cleave at Met-333.

CNBr-cleavage peptides were subjected to amino acid analysis and to partial or complete sequence analysis by means of the gas-phase method. Repetitive yields during sequencing were in the range 91–96%. The sequence data obtained are summarized in

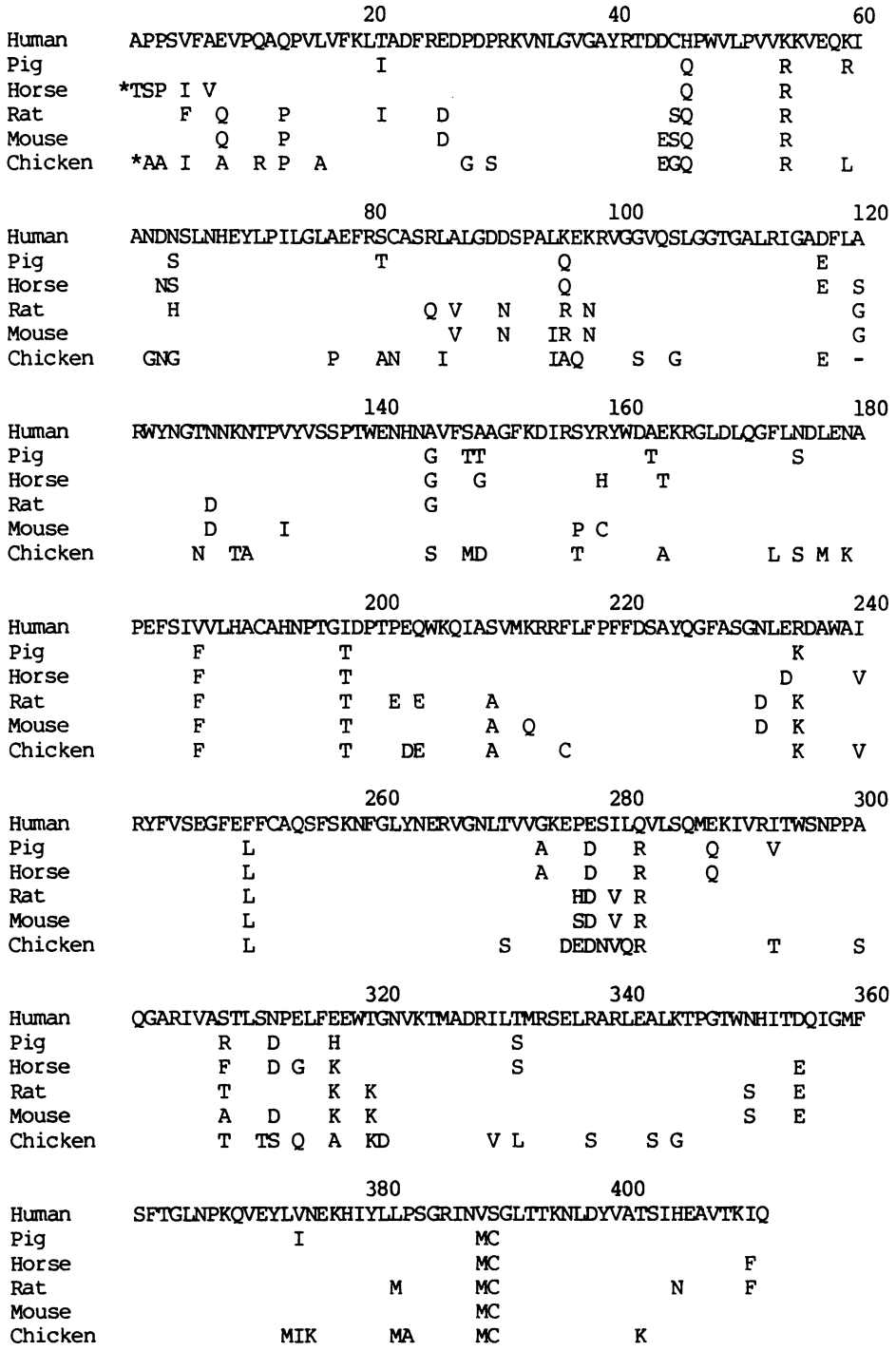

Fig. 2. Comparison of the amino acid sequences of cytosolic aspartate aminotransferases

For isoenzymes other than human, only residues that differ from those in the human form are shown. The symbol – at position 120 in the isoenzyme from chicken indicates a deletion. The symbol * at the *N*-terminus of the isoenzymes from horse and chicken indicates acetylated residues. The amino acid sequences of the isoenzymes from pig [3], horse [9] and chicken [4] were determined by direct protein sequencing; those of the isoenzymes from rat [11] and mouse [10] were obtained by cDNA sequencing.

Fig. 1, and detailed results are given in Supplementary Publication SUP 50158.

Sub-digestion of peptides

Peptides T-18 and T-28 were sub-digested with *S. aureus* V8 proteinase, as also were CNBr-cleavage peptides CN-1 and CN-6. Peptide T-34 was sub-digested with thermolysin. Product peptides were purified by h.p.l.c. Peptides were subjected to amino acid analysis and to sequence analysis by either the gas-phase method or the dansyl–Edman method (see Supplementary Publication SUP 50158).

Peptide SA-21A was subjected to digestion with carboxypeptidase Y. Lysine was identified as the *C*-terminal residue. This was important, since evidence for this residue (Lys-324) was otherwise indirect.

The sequence information obtained from sub-digests is sum-

Fig. 3. Comparison of the amino acid sequences of human cytosolic and mitochondrial aspartate aminotransferases

The symbol : indicates identical residues, whereas the symbol . indicates conservative substitutions. The symbol – indicates gaps introduced to maximize homology between the isoenzymes.

marized in Fig. 1. Not all the expected peptides were isolated (e.g. peptide SA-3 from digestion of peptide CN-1), but no attempt was made to isolate missing peptides if the region concerned had been sequenced from other digests.

Some unusual cleavages were observed, i.e. at Gly-115, Ser-148 and Asp-154. On the other hand, cleavage occurred only partially at Glu-203 and Glu-406 [resulting in peptides SA-11A (183–206) and SA-25B (377–417); not shown in Fig. 1].

Amino acid sequence of cytosolic aspartate aminotransferase from human liver

The amino acid sequence of the enzyme is given in Fig. 1 and was mainly based on the results described above. In a small number of cases direct evidence was not obtained for the position of a particular peptide (SA-10) or to establish the junction of two peptides (T-5/T-6, T-22/T-23, T-23/T-24, T-24/T-25, T-28/T-29). In these cases recourse was had to considerations of homology with other cytosolic aspartate aminotransferases (Fig. 2), which allowed the positions to be assigned unambiguously. For example peptides T-23, T-24 and T-25 share eight out of eight, eight out of nine and 11 out of 14 identical residues respectively with the corresponding region of the cytoplasmic isoenyzme from pig heart (Fig. 2). Given the clear amino acid analyses and sequence data for these peptides (see Supplementary Publication SUP 50158), this region of the structure may be considered to be firmly established.

Peptide CN-2 gave a clear *N*-terminal sequence except for the third residue, which could not be identified; the remainder of the *N*-terminal sequence was consistent with those of peptides T-21 and T-22. The amino acid composition was poor, so that residue 3 could not be obtained by difference, and attempts to improve these analyses were unsuccessful. Hence, although for reasons given above the rest of the sequence covered by peptide CN-2 is secure, residue 215 remains unidentified. It is given in Fig. 1 as arginine, since this amino acid is found at position 215 in all other cytosolic aspartate aminotransferases sequenced to date.

The evidence for one residue (Lys-258) was weak, depending only on the amino acid composition of peptide T-22. However, this is the active-site lysine residue present in all aspartate aminotransferases. In addition, Teranishi and co-workers [24] have sequenced the active-site peptide from the human heart cytosolic isoenzyme and find lysine at this position. Their sequence, covering residues 242–266, is completely in agreement with that presented here. They also sequenced residues 1–20 and 411–412, again with results in agreement with those given in Fig. 1.

All other residues were identified directly, in most cases from two or more peptides. Tryptophan residues were determined by automated gas-phase analysis of appropriate peptides. Amidation states of aspartic acid/asparagine and glutamic acid/glutamine residues were also directly assigned with the exception of Glu-57 and Glu-182, which were assigned on the basis of specificity of *S. aureus* V8 proteinase (glutamic acid occurs in these positions in all other aspartate aminotransferases).

Comparison of the sequence with those of other cytosolic aspartate aminotransferases

This is shown in Fig. 2. For the other sequences, only those residues differing from that found in the human isoenzyme are shown. The human isoenzyme contains unique residues at

positions 46, 54, 186, 198, 250, 278, 282, 389 and 390. The lack of cysteine at position 390 is of particular interest, as discussed above.

Comparison of the sequences of the cytosolic and mitochondrial isoenzymes

The computer program of Lipman & Pearson [19] was used to obtain the best alignment of the sequences of the two isoenzymes (Fig. 3); the sequence of the mitochondrial isoenzyme was taken from ref. [8] with one correction (Asp-234). In this comparison both identities and conservative substitutions are indicated. Conservative substitutions were taken to be those occurring with a relative frequency of 1.30 or greater as defined by McLachlan [25]. This is a more stringent definition than that employed by Lipman & Pearson [19], based on Dayhoff's PAM250 matrix [26].

Gaps required for optimal alignment occur at the same positions as for other cytosolic and mitochondrial isoenzyme pairs [9]. It is debatable whether it is justified to insert a gap at residues 250/251 in the cytosolic isoenzyme and 253/254 in the mitochondrial form; these gaps allow for matching of an adjacent cysteine residue, but cysteine residues are not particularly highly conserved in these isoenzymes.

Aligned as shown in Fig. 3, the isoenzymes have 48% identical residues. These are not uniformly distributed. The regions 1–31, 120–137 (including four deletions in the mitochondrial isoenzyme), 275–295 and 335–349 are particularly highly variable.

Quantitative comparisons of sequences of aspartate aminotransferases

Pair-wise comparisons were done, using the program of Lipman & Pearson [19], of the sequences of the vertebrate aspartate aminotransferases known to date, that is the cytosolic isoenzymes from human (this work), pig [3], horse [9], rat [11], mouse [10] and chicken [4], and the mitochondrial isoenzymes from human [8], pig [3], horse [9], rat [6], mouse [10] and chicken [5]. In each case the percentage sequence difference between pairs was calculated; when two sequences of different lengths were compared the percentage difference was calculated on the basis of the longer sequence with gaps counted as mismatches.

Also calculated were the average number of amino acid substitutions per site (K_{aa}) as defined by Kimura [20]. This is given by $K_{aa} = -\ln(1-p_d-0.2p_d^2)$, where p_d is the fraction of amino acid sequence differences. Calculated in this way K_{aa} values provide estimates of the total number of substitutions per site, including multiple substitutions at a single site.

The results are summarized in Table 3 in terms of the mean K_{aa} values for a particular set of comparisons together with the standard deviations and the number of comparisons.

The average K_{aa} value for mammalian mitochondrial isoenzymes (0.068 ± 0.015) is essential unchanged from that (0.069 ± 0.013) previously reported [9], as expected, since only one new sequence (mouse) has been included. The value for cytosolic isoenzymes differs considerably (0.092 ± 0.021 compared with 0.064). The value reported previously [9] was based on a single comparison (pig/horse) and hence was considered unreliable.

The essential point is that the two means are significantly different at $P < 0.01$. Consequently the results presented here constitute definitive evidence that the cytosolic aspartate aminotransferase has evolved faster (by a factor of about 1.3) than the mitochondrial isoenzyme in the mammalian lineage. This is consistent with the observation of specific conservation of certain residues ('m-residues') in the mitochondrial isoenzymes [5] and with the possibility that extra structural constraints are imposed on the structure of the mitochondrial isoenzyme concerned with its uptake from its site of synthesis in the cytosol into mitochondria.

Previous claims [13] that mammalian cytosolic aspartate aminotransferases have evolved faster than their mitochondrial counterparts were based on immunochemical comparisons. Although qualitatively correct, these conclusions were quantitatively grossly in error since the difference in rate of evolution was estimated to be a factor of 2. This underlines the danger of drawing conclusions about the degree of structural similarity between proteins from anything other than direct sequence determinations.

The absolute rates of evolution of the isoenzymes (k_{aa}) can be calculated from $k_{aa} = K_{aa}/2T$, where T is the time elapsed since divergence of mammals from their common ancestor [20]. Taking $T = 80 \times 10^6$ years, then $k_{aa} = 0.43 \times 10^{-9}$/year and 0.57×10^{-9}/year for mammalian mitochondrial and cytosolic isoenzymes respectively.

When all cytosolic and mitochondrial isoenzymes (including those from chicken) are compared (Table 3), the average K_{aa} value is 0.857 ± 0.033. Assuming that the rates of evolution in the two lines (cytosolic and mitochondrial) have not changed and taking the mean number of substitutions/site per year as 0.50×10^{-9}, then this puts the point of divergence of the isoenzymes from their common ancestor at 860×10^6 years ago. This value is somewhat different from that (10^9 years) previously reported [9] because of the greater number of sequences now available for comparison.

Table 3. Mean K_{aa} values for pairs of aspartate aminotransferases from various sources

The K_{aa} value gives an estimate of the average number of substitutions per site between any pair of isoenzymes.

Comparison	Mean $K_{aa} \pm$ S.D.	No. of pairs compared
Mammalian mitochondrial	0.068 ± 0.015	10
Mammalian cytosolic	0.092 ± 0.021	10
All cytosolic versus all mitochondrial	0.857 ± 0.033	36

REFERENCES

1. Barra, D., Bossa, F., Doonan, S., Fahmy, H. M. A., Hughes, G. J., Kakoz, K. Y., Martini, F. & Petruzelli, R. (1977) FEBS Lett. **83**, 241–244
2. Kagamiyama, H., Sakakibara, R., Wada, H., Tanase, S. & Morino, Y. (1977) J. Biochem. (Tokyo) **82**, 291–294
3. Barra, D., Bossa, F., Doonan, S., Fahmy, H. M. A., Hughes, G. J., Martini, F., Petruzzelli, R. & Wittman-Liebold, B. (1980) Eur. J. Biochem. **108**, 405–414
4. Shlyapnikov, S. V., Myasnikov, A. N., Severin, E. S., Myagkova, M. A., Torchinsky, Yu. M. & Braunstein, A. E. (1979) FEBS Lett. **106**, 385–388
5. Graf-Hausner, U., Wilson, K. J. & Christen, P. (1983) J. Biol. Chem. **258**, 8813–8826
6. Huynh, Q. K., Sakakibara, R., Watanabe, T. & Wada, H. (1980) Biochem. Biophys. Res. Commun. **97**, 474–479
7. Kondo, K., Wakabayashi, S., Yagi, T. & Kagamiyama, H. (1984) Biochem. Biophys. Res. Commun. **122**, 62–67
8. Martini, F., Angelaccio, S., Barra, D., Pascarella, S., Maras, B., Doonan, S. & Bossa, F. (1985) Biochim. Biophys. Acta **832**, 46–51
9. Doonan, S., Martini, F., Angelaccio, S., Pascarella, S., Barra, D. & Bossa, F. (1986) J. Mol. Evol. **23**, 328–335
10. Obaru, K., Nomiyama, H., Shimada, K., Nagashima, F. & Morino, Y. (1986) J. Biol. Chem. **261**, 16976–16983

11. Horio, Y., Tanaka, T., Taketoshi, M., Nagashima, F., Tanase, S., Morino, Y. & Wada, H. (1988) J. Biochem. (Tokyo) **103**, 797–804
12. Pol, S., Bousquet-Lemercier, B., Pave-Preux, M., Pawlak, A., Nalpas, B., Berthelot, P., Hanoune, J. & Barouki, R. (1988) Biochem. Biophys. Res. Commun. **157**, 1309–1315
13. Sonderegger, P. & Christen, P. (1978) Nature (London) **275**, 157–159
14. Doonan, S., Barra, D. & Bossa, F. (1984) Int. J. Biochem. **16**, 1193–1199
15. Porter, P. B., Barra, D., Bossa, F., Cantalupo, G., Doonan, S., Martini, F., Sheehan, D. & Wilkinson, S. M. (1981) Comp. Biochem. Physiol. B **69**, 737–746
16. Leung, F. Y. & Henderson, A. R. (1981) Clin. Chem. **27**, 232–238
17. Ellman, G. L. (1959) Arch. Biochem. Biophys. **82**, 70–77
18. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951) J. Biol. Chem. **193**, 265–275
19. Lipman, D. J. & Pearson, W. R. (1985) Science **227**, 1435–1441
20. Kimura, M. (1983) The Neutral Theory of Molecular Evolution, pp. 55–96, Cambridge University Press, Cambridge
21. Petrilli, P., Pucci, P., Garzillo, A. M., Sannia, F. & Morino, G. (1981) Mol. Cell. Biochem. **35**, 121–128
22. Martini, F., Angelaccio, S., Barra, D., Doonan, S. & Bossa, F. (1984) Biochim. Biophys. Acta **789**, 51–56
23. Critz, W. J. & Martinez-Carrion, M. (1977) Biochemistry **16**, 1554–1558
24. Teranishi, H., Kagamiyama, H., Teranishi, K., Wada, H., Yamano, T. & Morino, Y. (1978) J. Biol. Chem. **253**, 8842–8847
25. McLachlan, A. D. (1971) J. Mol. Biol. **61**, 409–424
26. Dayhoff, M. (1978) Atlas of Protein Sequence and Structure, vol. 5, suppl. 3, National Biomedical Research Foundation, Bethesda

Received 15 November 1989/12 March 1990; accepted 3 April 1990