Human peroxisomal L-alanine:glyoxylate aminotransferase

Evolutionary loss of a mitochondrial targeting signal by point mutation of the initiation codon

Yoshikazu TAKADA,*† Naomasa KANEKO,* Hiroyasu ESUMI,* P. Edward PURDUE† and Christopher J. DANPURE†

*Biochemistry Division, National Cancer Center Research Institute, Tokyo, Japan, and †Biochemical Genetics Research Group, Clinical Research Centre, Watford Road, Harrow, Middx. HA1 3UJ, U.K.

The amino acid sequence of human hepatic peroxisomal L-alanine:glyoxylate aminotransferase 1 (AGT1) deduced from cDNA shows 78% sequence identity with that of rat mitochondrial AGT1, but lacks the *N*-terminal 22 amino acids (the putative mitochondrial targeting signal). In humans this signal appears to have been deleted during evolution by a point mutation of the initiation codon ATG to ATA. These data suggest that the targeting defect in primary hyperoxaluria type 1, in which AGT1 is diverted from the peroxisomes to the mitochondria, could be due to a point mutation that reintroduces all or part of the mitochondrial signal sequence.

INTRODUCTION

Hepatic L-alanine:glyoxylate aminotransferase 1 (AGT1), which is a homodimer with a subunit M_r of approx. 42000 [1–3], catalyses the irreversible transamination of glyoxylate to glycine with L-alanine as the amino-group donor [1,4,5]. The intracellular localization of AGT1 is species-dependent. It is found in peroxisomes in primates (including humans) and lagomorphs, mitochondria in carnivores, and in both organelles in rodents [2,3,6–8].

In the lethal autosomal recessive disease primary hyperoxaluria type 1 (PH1), hepatic peroxisomal AGT1 is deficient [9]. Although most PH1 patients have a complete, or nearly complete, absence of AGT1, a considerable proportion (18/45; C. J. Danpure, unpublished work) have significant residual enzyme activity (5–50% of normal levels). In at least nine of these patients the disease appears to be caused by a trafficking defect in which AGT1 is diverted from the peroxisomes to the mitochondria [10]. At least in humans, its peroxisomal localization would appear to be essential to detoxify glyoxylate efficiently [11]. In PH1 patients with no AGT1 activity or with AGT1 localized in the mitochondria instead of the peroxisomes, failure to transaminate glyoxylate within the peroxisome allows the accumulated glyoxylate to be oxidized to oxalate [9,10]. The very low solubility of calcium oxalate causes all the pathological sequelae of PH1 (e.g. severe urolithiasis) [12]. In an attempt to understand the mechanisms by which AGT1 is variably targeted to peroxisomes and/or mitochondria, the sequence of human peroxisomal AGT1 cDNA has been determined and compared with that of rat mitochondrial AGT1 cDNA [13].

EXPERIMENTAL

Cloning and sequencing of human peroxisomal AGT1 cDNA

A fragment of rat mitochondrial AGT1 cDNA (1164 bp, 66–1230 [13]) was prepared by the polymerase chain reaction (PCR), using an Amplitaq kit (Cetus) with 1.5 mM-MgCl₂ and 0.5 mM-dNTPs and synthetic oligonucleotides (5'-ATGGGC-

TCGCACCAGCTGCTG-3' and 5'-AGGACAATGTTGCAG-GGCCTC-3') based on the published sequence [13], as primers. The PCR product was cut out from a 1.5%-agarose gel and electroeluted. A human liver λ gt11 cDNA library (5×10^5 recombinant plaques; Clontech) was screened with the fragment as probe at low stringency by plaque hybridization in $6 \times$ SSC/10 mM-EDTA/5 $\times$ Denhardt's solution/0.5% SDS/denatured salmon sperm DNA (100 μ g/ml) at 42 °C, and washing three times in $1 \times$ SSC at 45 °C for 10 min. Ten positive clones were plaque-purified. Phage DNA was obtained by the plate-lysate method and digested with restriction endonuclease *Eco*R1 and tested for cross-hybridization of the insert by Southern blotting [14]. All clones cross-hybridized to each other. The longest insert (A6-18; 1.25 kb) was subcloned into pBluescript KS(+) and sequenced. It did not contain the entire coding region, so an additional clone (H1-2, 1.6 kb) was selected from a HepG2 cDNA λ gt11 library (obtained from Dr. M. Krangel, Dana-Farber Cancer Institute, Boston, MA, U.S.A.). The DNA sequences were determined in both directions with a series of overlapping deletion clones [15] by the dideoxynucleotide-chain-termination method [16] using adenosine 5'-[α -thio]triphosphate labelled at the α -thio position with ³⁵S.

Northern-blot analysis of poly(A)⁺ mRNA from the livers of various PH1 patients and control subjects

Total RNA was extracted from liver stored in liquid N₂ by the guanidinium isothiocyanate/caesium chloride method [17]. Poly(A)⁺ mRNA was prepared by oligo(dT)-cellulose (BCL) column chromatography [18], and 0.5 μ g samples were fractionated by electrophoresis through 1.3%-agarose/2.2 M-formaldehyde gel. Marker tracks containing 28 S and 18 S rRNA and molecular-mass standards (BCL) were excised and stained with ethidium bromide. The remainder of the gel was capillary-blotted on to a nylon membrane (Hybond; Amersham International). The RNA was fixed to the membrane by u.v. irradiation (302 nm; 7000 μ W/cm²) for 60 s. The blot was pre-

Abbreviations used: AGT1, L-alanine:glyoxylate aminotransferase 1; PH1, primary hyperoxaluria type 1; PCR, polymerase chain reaction; SSC, saline/sodium citrate (1 $\times$ SSC is 0.15 M-NaCl/15 mM-sodium citrate); poly(A)⁺, polyadenylated; 1 $\times$ Denhardt's solution is 0.02% (w/v) BSA/0.02% (w/v) polyvinylpyrrolidone/0.02% (v/v) Ficoll.

† Present address and address for correspondence and reprint requests: Research Institute of Scripps Clinic, 10666 North Torrey Pines Road, La Jolla, CA 92037, U.S.A.

These sequence data will appear in the EMBL/GenBank/DDBJ Nucleotide Sequence Databases.

hybridized in $6 \times \text{SSC}/0.1\%$ SDS/50% deionized formamide/10% Denhardt's solution/sonicated and boiled salmon sperm DNA (100 $\mu\text{g}/\text{ml}$) at 40 °C for 12 h and then hybridized at 40 °C for 15 h with denatured probe (isolated cDNA insert of clone A6-18). After removal of the probe (0.1 $\times$ Denhardt's solution/5 mM-Tris/HCl/2 mM-EDTA, pH 8.0, 65 °C for 1 h) the same blot was reprobed with a 1.9 kb fragment of human

albumin cDNA as an internal loading control. The probes were labelled by the random-priming method [19] to a specific radioactivity of about 5×10^8 c.p.m./ μg . Filters were washed three times each for 15 min in $2 \times \text{SSC}/0.1\%$ SDS at 40 °C, once for 15 min in 0.1 $\times$ SSC/0.1% SDS and then exposed to preflashed Hyperfilm-MP (Amersham) at -70 °C with intensifying screens for 24 h.

```

CCGGAAGCCCATCCACCAATCCTCACCTCTCACCTCTGTGTCCGCCCTGCTGGGAAATAT 60
TCCAGGCTTTGGCCAAGGCCAGTGCAGCCCCAGGTTCCCGAGCGGCAGGTTGGGTGCGGA 120
CCATGGCCTCTCACAAAGCTGCTGGTGACCCCCCAAGGCCCTGCTCAAGCCCCCTCTCCA 180
      M A S H K L L V T P P K A L L K P L S
(22 aa) -- G Q P E S
TCCCAACCCAGCTCCTGCTGGGGCCTGGTCTTCAACCTGCCTCCTCGCATCATGGCAG 240
I P N Q L L L G P G P S N L P P R I M A
      K R R A V L
CCGGGGGGCTGCAGATGATCGGGTCCATGAGCAAGGATATGTACCAGATCATGGACGAGA 300
A G G L Q M I G S M S K D M Y O I M D E
      S R H Q E F
TCAAGGAAGGCATCCAGTACGTGTTCCAGACCAGGAACCCACTCACACTGGTCATCTCTG 360
I K E G I Q Y V F Q T R N P L T L V I S
      Q V
GCTCGGGACACTGTGCCCTGGAGGCCGCCCTGGTCAATGTGCTGGAGCCTGGGGACTCCT 420
G S G H C A L E A A L V N V L E P G D S
      M T F L
TCCTGGTTGGGGCAATGGCATTGGGGGCAGCGAGCCGTGGACATCGGGGAGCGCATAG 480
F L V G A N G I W G Q R A V D I G E R I
      T I P A E A
GAGCCCGAGTGCACCCGATGACCAAGGACCCTGGAGGCCACTACACTGCAGGAGGTGG 540
G A R V H P M T K D P G G H Y T L Q E V
      Q I K E
AGGAGGGCCTGGCCAGCACAAGCCAGTGTCTGTCTTAACCCACGGGGAGTCGTCCA 600
E E G L A Q H K P V L L F L T H G E S S
      V
CCGGCGTGTGCAGCCCCCTTGATGGCTTCGGGGAACCTTGCCACAGGTACAAGTGCCTGC 660
T G V L Q P L D G F G E L C H R Y K C L
      Q
TCCTGTGGATTTCGGTGGCATCCCTGGGCGGACCCCCCTTTACATGGACCGCAAGGCA 720
L L V D S V A S L G G T P L Y M D R Q G
      V I Q
TCGACATCCTGTACTCGGGCTCCAGAGGCGCCTGAACGCCCTCCAGGGACCTCGCTCA 780
I D I L Y S G S Q K A L N A P P G T S L
      V I
TCTCCTTCAGTGACAAGGCCAAAAAGATGTACTCCCGCAAGACGAAGCCCTTCTCCT 840
I S F S D K A K K K M Y S R K T K P F S
      N S V
TCTACCTGGACATCAAGTGGCTGGCCAACTTCTGGGGCTGTGACGACCCAGCCAGGATGT 900
F Y L D I K W L A N F W G C D D Q P R M
      T Y S K L E G K T V
ACCATCACACAATCCCGCTCATCAGCCTGTACAGCCTGAGAGAGAGCCTGGCCCTCATTG 960
Y H H T I P V I S L Y S L R E S L A L I
      I L C
CGGAACAGGGCCTGGAGAACAGCTGGCGCCAGCACCAGGCGCGGCGTATCTGCATG 1020
A E Q G L E N S W R Q H R E A A A Y L H
      S R T H
GGCGCCTGCAGGCACTGGGGCTGCAGTCTTTCGTGAAGGACCCGGCGCTCCGGCTTCCCA 1080
G R L Q A L G L Q L F V K D P A L R L P
      K C R E K F E I
CAGTCACCACTGTGGCTGTACCCGCTGGCTATGACTGGAGAGACATCGTCAGTACGTCA 1140
T V T T V A V P A G Y D W R D I V S Y V
      I T N M
TAGACCACTTCGACATTGAGATCATGGGTGGCCTTGGGCCCTCCACGGGAAGGTGCTGC 1200
I D H F D I E I M G G L G P S T G K V L
      L N S E D
GGATCGGCCTGCTGGGCTGCAATGCCACCCGCGAGAATGTGGACCGGTGACGGAGGCCCC 1260
R I G L L G C N A T R E N V D R V T E A
      P Y T A A
TGAGGGCGGCCCTGCAGCACTGCCCCAAGAAGAAGCTGTGACCTGCCACTGGCACACAG 1320
L R A A L Q H C P K K L U
      T E N U
CTGGCACTGGCACACCTGTCCCATGCCACCCTGAGGGATCAGGAGCAAACAGACCCT 1380
GCAAGGTCTCCAGGCCTGGGGACAGGAAAGCCACTGACCCAGCCCGGAGGCAGAACCA 1440
GGCAGCCTCCCTGGCCCCAGGCAGCCCTTTCCCTCCAGTGGCACTCCTGGAAACAGTC 1500
CACTTGGGCGCAAAACCCAGTGCCTTCCAAATGAGCTGCAGTCCCAGGCCATGAGCCTC 1560
CCGGGAATGTTTAAATAAAGGGCCTGGCCACCTCTCCTCACAAAAA

```

Fig. 1. Sequence of human peroxisomal AGT1 cDNA

The details of the cloning and sequencing procedures are described in the Experimental section. The AATAAA sequence near the 3'-end indicates the polyadenylation consensus signal. The residues substituted in the rat mitochondrial AGT1 sequence are shown underneath the human peroxisomal AGT1 amino acid sequence. The underlined runs of amino acids correspond to independently determined tryptic peptide sequences.

RESULTS AND DISCUSSION

cDNA clones encoding human peroxisomal AGT1 were isolated by screening cDNA libraries from human liver and a hepatocarcinoma cell line using rat AGT1 cDNA as the probe. The nucleotide sequence (Fig. 1) from two overlapping clones (A6-18, H1-2) encoded a protein of 392 amino acids with a calculated M_r of 43046, which is in close agreement with the size of mature human peroxisomal AGT1. The identity of these clones was confirmed by comparison of the predicted amino acid sequence with the sequences of two internal peptides (DMYQIMDEIK and LPTVTTVAPAGY) generated by tryptic digestion of human AGT1. The sequence of human peroxisomal AGT1 shows about 78% amino acid sequence identity with rat mitochondrial AGT1. The putative pyridoxal phosphate-binding lysine residue at position 209 is conserved [13]. A comparison of the 5' sequence of rat with that of human AGT1 cDNA suggests that the N-terminal 22 amino acids of the nascent rat translation product are absent from the human protein (Fig. 2). The rat cDNA sequence encoding this 22-amino-acid extension, which is believed to be a mitochondrial-targeting signal, shows high sequence identity (74%) with the corresponding non-coding region in the human cDNA. The sequence GCTGGGAAA-TGTTCC in rat mitochondrial AGT1 cDNA around the first ATG corresponds to GCTGGGAAATATTC in human peroxisomal AGT1 cDNA, clearly indicating that the mitochondrial-targeting sequence in human AGT1 has been deleted during evolution by a single point mutation at the first ATG codon. The cDNA sequence in this region has been confirmed by analysis of the sequence of the corresponding portion of a human AGT1 genomic clone (P. E. Purdue, unpublished work). The loss of the mitochondrial targeting signal during evolution may partly explain the species differences in intracellular localization of AGT1 [2,3,6].

The identical rates of evolution of peroxisomal and mitochondrial AGT1 from different species suggests that they did not arise by gene duplication [20,21]. In rats, mature

peroxisomal and mitochondrial AGT1 appear to be identical with respect to physical, enzymic and immunological properties [2,22,23]. Southern-blotting analysis of rat genomic DNA suggests that there is only one gene for AGT1 [24], and preliminary data (P. E. Purdue, unpublished work) indicates that this is also the case in humans. It is likely, therefore, that the rat peroxisomal and mitochondrial AGT1 are encoded by the same gene. In this respect it is interesting to note that the two different rat AGT1 species, with and without a mitochondrial-targeting sequence, might arise from a common AGT1 mRNA by translation initiation from the first (or second) and the third ATG codon respectively (Fig. 2).

The molecular basis of targeting of AGT1 to peroxisomes is yet to be resolved. The C-terminal sequences of both human and rat AGT1 (-KKKL and -KNKL respectively) are similar to the C-terminal sequence of firefly luciferase (-KSKL), the last three amino acids of which have been proposed to constitute a peroxisomal-targeting sequence [25–28]. However, mutagenesis *in vitro* of the firefly luciferase C-terminal sequence to KKKL apparently abolishes peroxisomal-targeting activity in transfected monkey kidney-cell cultures [25]. Although the effect of mutating this sequence to KNKL is not known, this does raise the possibility that other sequences are responsible for targeting AGT1 to the peroxisomes.

We propose that point mutations which introduce in-frame ATG codons 5' to the human AGT1 open reading frame might be the possible mechanism by which AGT1 is diverted to the mitochondria in some PH1 patients. Northern-blot analysis shows that there are no major differences in the size of AGT1 mRNA in PH1 patients with the trafficking defect compared with controls (Fig. 3). This finding, together with the observations that analysis of a genomic clone reveals no additional ATGs within 1 kb 5' to the start of the cDNA, and that the 5' untranslated region of the cDNA and the first 180 bp of the coding sequence are within the same exon (P. E. Purdue, unpublished work), suggest that alternative transcription start sites or alternate pre-mRNA splicing at the 5' end are less likely,

Fig. 2. Alignment of the nucleotide sequences of the 5' non-coding region of human peroxisomal AGT1 cDNA and the corresponding coding region of rat mitochondrial AGT1 cDNA

The amino acid sequences of human and rat nascent AGT1 deduced from cDNA are shown boxed. The putative mitochondrial targeting sequence of rat mitochondrial AGT1 is underlined. The N-terminal amino acid sequence of mature mitochondrial AGT1 is GSHQLLV- [13,24]. The N-terminus of mature human peroxisomal AGT1 is blocked. The 5' non-coding sequence in human AGT1 cDNA was tentatively translated into amino acid sequence in the same frame as the coding region (shown in lower case). Asterisks mark the sites where point mutations could introduce in-frame ATG codons. The double asterisks indicate the site of point mutation of ATG to ATA, which destroyed the initiation codon. The colons (:) indicate sequence identity.

Fig. 3. Northern blot of poly(A)⁺ mRNA isolated from the livers of various primary hyperoxaluria type 1 patients and a control subject

The details of the procedures are described in the Experimental section. The samples were from a control subject (C) and four PH1 patients (P1–4). Abbreviations: AGT%, residual AGT1 enzyme activity as a percentage of the mean of the normal range [29]; CRM, amount of immunoreactive AGT1 protein as determined by immunoblotting (+++, normal; –, undetectable); Subcell, subcellular localization of immunoreactive AGT1 protein as determined by Protein A–gold immunoelectron microscopy.

although they cannot be ruled out. There are five possible sites in the 5' non-coding region of the human AGT1 cDNA that could form in-frame ATG codons by appropriate single point mutations (Fig. 2). One of these mutations would reintroduce the ATG codon and the complete mitochondrial targeting sequence lost during evolution. Three of the others would lead to the synthesis of partial mitochondrial signal sequences, and one would produce a protein with a complete mitochondrial targeting sequence possessing an *N*-terminal extension. It is probable that the efficiencies with which each of these five mutant AGT1 species can be targeted to mitochondria will be different; and in this respect it is interesting to note that the residual levels of AGT1 in the nine definite (18 possible) PH1-trafficking-defect variants so far identified range from 5 to 50% of normal (C. J. Danpure, unpublished work).

We thank the following: Mrs. B. Tracey and Dr. A. Aitkin for the h.p.l.c. separation and amino acid sequencing of the tryptic peptides respectively, Miss J. Allsop, Mrs. P. R. Jennings and Mr. M. Lumb for

technical help, and Dr. Y. Minatogawa and Dr. S. Tone (Wakayama Medical College) for helpful discussion.

REFERENCES

- Noguchi, T., Okuno, E., Takada, Y., Minatogawa, Y., Okai, K. & Kido, R. (1978) *Biochem. J.* **169**, 113–122
- Takada, Y. & Noguchi, T. (1982) *Comp. Biochem. Physiol.* **72B**, 597–604
- Noguchi, T. & Takada, Y. (1979) *Arch. Biochem. Biophys.* **196**, 645–647
- Takada, Y. & Noguchi, T. (1987) *Methods Enzymol.* **142**, 273–279
- Thompson, J. S. & Richardson, K. E. (1967) *J. Biol. Chem.* **242**, 3614–3619
- Okuno, E., Minatogawa, Y., Nakanishi, J., Nakamura, M., Kamoda, N., Makino, M. & Kido, R. (1979) *Biochem. J.* **182**, 877–879
- Cooper, P. J., Danpure, C. J., Wise, P. J. & Guttridge, K. M. (1988) *J. Histochem. Cytochem.* **36**, 1285–1294
- Yokota, S., Oda, T. & Ichiyama, A. (1987) *Histochemistry* **87**, 601–606
- Danpure, C. J. & Jennings, P. R. (1986) *FEBS Lett* **201**, 20–24
- Danpure, C. J., Cooper, P. J., Wise, P. J. & Jennings, P. R. (1989) *J. Cell Biol.* **108**, 1345–1352.
- Takada, Y., Mori, T. & Noguchi, T. (1984) *Arch. Biochem. Biophys.* **229**, 1–6
- Williams, H. E. & Wandzilak, T. R. (1989) *J. Urol.* **141**, 742–747
- Oda, T., Miyajima, H., Suzuki, Y. & Ichiyama, A. (1987) *Eur. J. Biochem.* **168**, 537–542
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular Cloning: a Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
- Henikoff, S. (1984) *Gene* **28**, 351–359
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463–5467
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299
- Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. U.S.A.* **69**, 1408
- Feinberg, A. P. & Vogelstein, B. (1984) *Anal. Biochem.* **137**, 266–267 (addendum)
- Takada, Y. & Noguchi, T. (1982) *Biochem. Biophys. Res. Commun.* **108**, 152–157
- Wilson, A. C., Carlson, S. S. & White, T. J. (1977) *Annu. Rev. Biochem.* **46**, 573–639
- Noguchi, T., Takada, Y. & Kido, R. (1977) *Biochem. J.* **161**, 609–614
- Noguchi, T. & Takada, Y. (1978) *Biochem. J.* **175**, 765–768
- Ichiyama, A., Oda, T., Miyajima, H., Funai, T., Itoh, T. & Nishiyama, K. (1988) *Seikagaku* **60**, 1253–1266
- Gould, S. J., Keller, G. A., Wilkinson, J. & Subramani, S. (1989) *J. Cell Biol.* **108**, 1657–1664
- Gould, S. J., Keller, G. A. & Subramani, S. (1987) *J. Cell Biol.* **105**, 2923–2931
- Gould, S. J., Keller, G. A. & Subramani, S. (1988) *J. Cell Biol.* **107**, 897–905
- Miyazawa, S., Osumi, T., Hashimoto, T., Ohno, K., Miura, S. & Fujiki, Y. (1989) *Mol. Cell. Biol.* **9**, 83–91
- Danpure, C. J. & Jennings, P. R. (1988) *Clin. Sci.* **75**, 315–322

Received 19 December 1989/19 March 1990; accepted 23 March 1990