The W and L allelic forms of phenylalanine hydroxylase in the rat differ by a threonine to isoleucine substitution

Julian F. B. MERCER,*† Wendy McADAM,* Geoffrey W. CHAMBERS† and Ian D. WALKER†

*Birth Defects Research Institute, Royal Children's Hospital, Flemington Road, Parkville, Victoria 3052, Australia, and

†Research Centre for Cancer and Transplantation, Department of Pathology, University of Melbourne, Parkville, Victoria 3052, Australia

High performance liquid chromatography maps of tryptic and chymotryptic peptides from the W and L forms of rat phenylalanine hydroxylase differed by one peptide. Sequencing of the variant tryptic peptides showed a substitution of threonine in the W form by isoleucine in the L form and this same difference was confirmed in the chymotryptic peptides. This allelic substitution would result from a nucleotide change of ACA to ATA at amino acid position 371 of the full phenylalanine hydroxylase sequence. Altered sodium dodecyl sulphate binding is postulated to explain the change in mobility of the proteins observed on sodium dodecyl sulphate/polyacrylamide gels.

INTRODUCTION

Phenylalanine hydroxylase (EC 1.14.16.1) is responsible for the conversion of phenylalanine into tyrosine, and its deficiency in humans results in the disease phenylketonuria (for reviews, see Kaufman, 1977; Cotton, 1977; Goodwin, 1979). In the rat we have previously reported two forms of the enzyme which are distinguished by a small mobility difference on SDS/polyacrylamide-gel electrophoresis. These variants termed W (apparent M_r 50000) and L (apparent M_r 49000) are coded by allelic genes, which are differentially distributed in common inbred rat strains (Mercer *et al.*, 1984). In this study we have investigated the structural bases for this allelism. Our investigations have revealed only one amino acid difference between the two allelic variants, threonine in the W variant being substituted by isoleucine in the L variant.

MATERIALS AND METHODS

Materials

All chemicals used in this study were of at least AR quality. Trypsin (tosylphenylalanylchloromethane-treated) and chymotrypsin were obtained from Worthington and Boehringer Mannheim (Mannheim, Germany) respectively. Sequencing chemicals were obtained from Applied Biosystems Incorporated (Foster City, CA, U.S.A.).

Rat breeding

Sprague–Dawley rats of genotype WW and LL were identified by liver biopsy and analysis of phenylalanine hydroxylase type as described by Mercer *et al.* (1984). Breeding pairs were set up to provide sufficient progeny for isolation of the pure forms of the enzyme.

Phenylalanine hydroxylase purification

The enzyme was isolated by the two-step procedure of Shiman *et al.* (1979) without the use of DEAE-cellulose. Further purification was achieved by the phenyl-Sepharose

binding and elution procedure described by Shiman (1980). The protein purity was > 98% as estimated by SDS/polyacrylamide-gel electrophoresis.

Carboxamidomethylation and trypsin digestion

Purified, freeze-dried phenylalanine hydroxylase was redissolved in 1.0% (w/v) SDS/0.1M-Tris/HCl, pH 8.0, at a concentration of 1 mg/ml. Dithiothreitol was added to a final concentration of 10 mM and the reduction mixture was boiled for 5 min and then incubated for a further 1 h in a sealed tube at ambient temperature. Iodoacetamide was added to a final concentration of 22 mM and the mixture was incubated in the dark for a further 15 min. Ice-cold acetone (9 vol.) was then added followed immediately by an appropriate amount of either trypsin or chymotrypsin. The amount of proteinase added was 2% (w/w) of the amount of substrate present. The mixtures were incubated at –20°C for 1–3 h and the protein pellets were recovered by centrifugation (10000 g, 30 min). After careful decantation of the acetone, aqueous trimethylamine bicarbonate (1% w/v, pH 8.0) was added to dissolve the protein pellets at a final concentration of 1 mg/ml. A further 2% (w/w) of proteinase was added and digestions were conducted for 3 h at 37 °C and terminated by freezing.

Chymotrypsin digestion

Carboxamidomethylation was performed essentially as above except that the reduction was carried out in 6M-guanidinium chloride/2mM-EDTA/0.2M-Tris/HCl, pH 9.0. After the addition of iodoacetic acid, the reaction was stopped with excess β -mercaptoethanol. The material was then dialysed into 50 mM-NH₄HCO₃ (pH 8.0) and lyophilized ready for chymotryptic digestion. Chymotryptic digestion was performed at 37 °C for 6 h with a ratio of proteinase to substrate of 1:50. The reaction was stopped by boiling for 10 min.

Peptide chromatography

Tryptic peptide samples (0.1–1.0 ml) were separated by

Abbreviation used: PTH, phenylthiohydantoin.

† To whom correspondence and reprint requests should be addressed.

Fig. 1. Tryptic peptide maps of the W and L variants of phenylalanine hydroxylase

Digests of W (a), L (b) and a mixture (c) of both variants subjected to reverse phase h.p.l.c. as described in the Materials and methods section. The allele-specific peptides are designated TW and TL. Samples were applied at 0 min.

reverse-phase h.p.l.c. on a Novapak (C_{18}) column (Waters; 0.39 cm $\times$ 14 cm) eluted with a linear gradient of acetonitrile (0–50%) at a flow of 1 ml/min and a total run time of 60 min. Peptides were detected by their absorbance at 214 nm by using a Waters Model 441 fixed-wavelength detector.

Chymotryptic peptides were separated by a two-step procedure using a Brownlee Aquapore RP300 column (10 μ m; 4.6 $\times$ 25 cm). Initially the peptides were separated on the h.p.l.c. column equilibrated in 20 mM- NH_4HCO_3 with elution of peptides by a linear gradient of acetonitrile from 8% to 60% over a period of 40 min. Selected areas which differed between the two variants were then separated on the h.p.l.c. column equilibrated in 0.1% trifluoroacetic acid.

Peptide sequencing

Peptide samples were subjected to automated Edman degradation using an Applied Biosystems Model 470A

sequencer based on the original design of Hewick *et al.* (1981). The sequencing program used differed from that of Hewick *et al.* (1981) by the deletion of all steps involving vacuum applications and their replacement by extended argon deliveries to the cartridge and flask. The sequencing reagents used were all obtained from Applied Biosystems Inc. except for the 12.5% (w/v) trimethylamine, which was prepared from anhydrous trimethylamine purified by passage through a phthalic anhydride column.

Identification of PTH amino acids was accomplished by reverse phase h.p.l.c. on a Beckman PTH (C_{18}) column with an acetonitrile gradient elution system delivered by Waters chromatographic hardware. Detection of peaks was by absorbance at 254 nm. In cases where unambiguous identification of particular PTH amino acids was not possible, either trace enrichment or chromatography of remaining samples with an IBM Cyano column was conducted. The PTH derivative of carboxamidomethyl-cysteine eluted unambiguously from the Beckman PTH column as a major peak with a slightly longer retention time than that of PTH-Gly.

RESULTS

Tryptic peptide maps

Tryptic peptide maps of both variants of phenylalanine hydroxylase are shown in Fig. 1. In repeat experiments only one major tryptic peptide difference was detected reproducibly between the W and L forms of the enzyme (labelled respectively TW and TL in the peptide maps of the W and L variants, Figs. 1a and 1b). A mixed digest was then analysed to confirm that the chromatographic difference between the W and L variants was authentic (Fig. 1c). As anticipated, the two late eluting peptide peaks TW and TL were recovered from the mixture. The other major peptide peaks in the mixture were all present in similar amounts in the tryptic peptide maps of both the W and L variants of the enzyme, indicating that apart from the TW/TL difference the two variants were similar or identical.

Minor peptide differences between the W and L digests of Fig. 1 could not be reproduced in repeat experiments and could be due to small unintentional differences in digestion conditions. Alternatively, minor contaminants in either the W or L protein preparations may contribute peptide peaks exhibiting variation between different batches of substrate.

Sequence of allele-specific tryptic peptide

The partial N-terminal sequences of TW and TL were determined and the quantitative data from these experiments is presented in Figs. 2(a) and 2(b). The only sequence difference between the two peptides resided at their respective amino termini: Ile-1 in the L variant was substituted by Thr-1 in the W variant.

Although peptide TL was slightly contaminated with a minor impurity sequence (Fig. 2b), unambiguous residue assignments were made through to residue 18. The quantitative analysis of sequence data presented in Fig. 2 shows that most residues were recovered in the expected amounts. Threonine, serine and cysteine residues are always underestimated due to their destruction during the sequencing and conversion reactions and consequently deviate negatively from linearity in the semi-logarithmic plots of Fig. 2. The chromatographic

Fig. 2. Semi-logarithmic plots of PTH amino acid yields for peptides TW, TL and CL

Peptide sequencing was carried out as described in the Materials and methods section. The yields of individual PTH amino acids were calculated after subtraction of background from the previous cycle and normalization of peak height using PTH-norleucine as an internal standard. The data for peptides, TW, TL and CL are presented respectively in the upper (a), middle (b) and lower (c) panels; - indicates that no assignment was possible. Parentheses indicate PTH amino acids thought to be due to a minor contaminating peptide in the TL sample.

assignments of other residues present in lower than expected amounts (in particular Glu-5 and Glu-10 of peptide TW) were re-examined to confirm that they were the only major products of those cycles of Edman degradation.

Chymotryptic peptide analyses and sequence of L-specific peptide

A further digestion of both phenylalanine hydroxylase variants with a different proteolytic enzyme was conducted to investigate whether other sites of allelic variation occurred. Fig. 3 shows the results of comparative chymotryptic peptide mapping of both variants. Most

Fig. 3. Chymotryptic maps of the W and L forms of phenylalanine hydroxylase

Digests of the W (a) and L (b) forms were subjected to reverse phase h.p.l.c. as described in the Materials and methods section. The peaks taken for rechromatography are indicated by the stippled box. The peaks labelled 'S' are due to the solvent fronts.

peptides seemed to be identical, with minor differences not reproducible between runs except for a group of poorly resolved peaks at the beginning of the gradient indicated by the stippled box which appeared to contain additional material in the digest of the L form. The peak indicated by 'S' in Fig. 3, which is larger in the W-form map, is due to the solvent front and is quite variable between runs. This group of peptides from both digests was re-chromatographed using trifluoroacetic acid as an ion-suppression agent, resulting in the resolution of a peptide (CL) which although prominent in the L variant sample (Fig. 4b) was qualitatively absent from the W variant sample (Fig. 4a). The sequence of peptide CL is shown in Fig. 2(c) and overlaps the tryptic peptide TL; in particular it contains isoleucine at position 5 rather than threonine.

Location of the allelic difference in the phenylalanine hydroxylase sequence

A predicted amino acid sequence of the C-terminal portion of rat phenylalanine hydroxylase has been derived from a partial cDNA clone (Robson *et al.*, 1984), and we have recently isolated and sequenced a full-length cDNA clone for this enzyme (Dahl & Mercer, 1986). The peptide sequence of TW matches exactly the published sequence of Robson *et al.* (1984) and that of our cDNA clone. This result confirms the sequence data and shows that the cDNA clones are derived from the W-form mRNA. The alignment of the peptides with the predicted amino acid sequence is shown in Fig. 5. The allelic difference occurs at amino acid 371 from the N-terminal alanine of the mature protein, and is located 81 amino acids from the C-terminus.

Fig. 4. Rechromatography of selected chymotryptic peptides

The selected peptides from Fig. 3 were subjected to reverse phase h.p.l.c. using 0.1% trifluoroacetic acid in place of 20 mM-NH₄HCO₃. Peptides from the W form are shown in (a) and those from the L form in (b); the L-specific peptide which was sequenced is indicated by CL. The arrows indicate sample application.

DISCUSSION

We have investigated the allelic difference between the two variants of rat phenylalanine hydroxylase. Comparative tryptic peptide mapping revealed one major peptide difference between the two variants (Fig. 1). Sequence analysis of the two allelic tryptic peptides (TW and TL) revealed an *N*-terminal threonine from the W form but isoleucine for the L form (Fig. 2). Otherwise the two peptide sequences were identical. Chymotryptic peptide mapping experiments resulted in the isolation of a peptide (CL) which although present in the L variant was completely absent from the W variant. No W variant homologue of this peptide was isolated. The sequence of this peptide (Fig. 2) overlapped and extended the sequence of peptide TL; in particular it clearly defined the same allele-specific sequence as peptide TL.

It is likely at least in the tryptic digests of Fig. 1 that most or all of the phenylalanine hydroxylase molecule is represented by the 37 or so major peptide peaks present in both W and L forms. The cDNA sequence of phenylalanine hydroxylase (Dahl & Mercer, 1986) predicts 42 sites of tryptic cleavage, some of which should produce single amino acids (arginine or lysine). Clearly most or all of the expected tryptic peptides are present in the chromatograms of Fig. 1 and it is most unlikely that a further major allelic peptide would have been missed. The fact that the peptide CL, the only chymotryptic peptide isolated from either W or L digests which was allele-specific, exactly overlapped peptide TL strongly argues for a single site of allelic variation. It seems

Fig. 5. Location of the allelic difference in the predicted amino acid sequence of phenylalanine hydroxylase

The amino acid sequence shown is part of the predicted amino acid sequence from a cloned cDNA (Robson *et al.*, 1984). The amino acids are numbered from the *N*-terminal alanine predicted from a full-length rat cDNA (Dahl & Mercer, 1986). The allelic variable position is shown by the box and the L-specific chymotryptic peptide (CL) and the allele-specific tryptic peptides (TW, TL) are shown by the solid lines.

unlikely that other allelic sites within the W and L forms would have been missed by both enzymes under conditions where the same allelic difference was revealed in both digests. We cannot, however, formally rule out the existence of any other allelic difference between the W and L forms of phenylalanine hydroxylase. Nevertheless a single Ile $\leftrightarrow$ Thr interchange could account for several other recent findings. For example, the change of threonine to isoleucine does not alter the charge of the polypeptide chain and this is in agreement with the identity of isoelectric points of the different polypeptides of the W and L forms (Smith *et al.*, 1985; J. F. B. Mercer, unpublished work). The location of the allelic difference 81 amino acids from the C-terminus of the protein is in good agreement with the V8 proteinase digestion data which indicated that the difference between the W and L forms was within about 70 amino acids from one end of the polypeptide chain (Mercer *et al.*, 1984). The change from threonine to the more hydrophobic isoleucine may result in additional SDS binding which would explain the higher mobility of the L form on SDS/polyacrylamide gels. A similar change in mobility on SDS/polyacrylamide gels was observed following a substitution of leucine by glutamine in α -crystallin chains (De Jong *et al.*, 1978).

The C-terminal region of phenylalanine hydroxylase is not particularly well conserved between the rat and human enzymes, whilst other regions are completely conserved (Kwok *et al.*, 1985; Dahl & Mercer, 1986). A number of amino acid changes in this region are tolerated without abolition of enzyme activity. Thus the change of threonine to isoleucine would not be expected to result in any difference in activity and indeed we have been unable to show any differences in function or stability between the W and L forms.

Recently we have found that monkeys have two apparent M_r forms of phenylalanine hydroxylase that are encoded by different mRNA species (Smith *et al.*, 1985). In addition we have found that a number of mouse species have a form of phenylalanine hydroxylase which comigrates with the W form of the rat, but that *Mus spretus* contains a form with the same apparent M_r as the L form (J. F. B. Mercer, unpublished work). Thus phenylalanine hydroxylase forms which differ in apparent M_r are found in a number of species. Our data in this paper suggest that the molecular basis for the difference in the rat is a single amino acid substitution that apparently alters SDS binding. If analysis of phenyl-

alanine hydroxylase sequences from these other species shows the same substitution as the basis of the apparent M_r differences, this would imply that the origin of the W and L forms predates the mammalian radiation. Alternatively, a number of different amino acid substitutions which affect mobility on SDS/polyacrylamide-gel electrophoresis may have occurred in the poorly conserved portions of the phenylalanine hydroxylase molecule.

We thank Dr. R. G. H. Cotton and Professor David Danks for their advice and criticism of the manuscript and Andrew Grimes for his expert technical assistance. This work was supported in part by the National Health and Medical Research Council of Australia.

REFERENCES

- Cotton, R. G. H. (1977) *Int. J. Biochem.* **8**, 333–341
Dahl, H.-H. M. & Mercer, J. F. B. (1986) *J. Biol. Chem.*, in the press
- De Jong, W. W., Zweers, A. & Cohen, L. H. (1978) *Biochem. Biophys. Res. Commun.* **82**, 532–539
Goodwin, B. L. (1979 in *Aromatic Amino Acid Hydroxylases and Mental Disease* (Youdim, M. B. H., ed.), pp. 5–76, John Wiley and Sons, New York
Hewick, R. M., Hunkapillar, M. W., Hood, L. E. & Dreyer, W. J. (1981) *J. Biol. Chem.* **256**, 7990–7997
Kaufman, S. (1977) *Adv. Neurochem.* **12**, 1–132
Kwok, S. C. M., Ledley, F. D., DiLella, A. G., Robson, K. J. H. & Woo, S. L. C. (1985) *Biochemistry* **24**, 556–561
Mercer, J. F. B., Grimes, A., Jennings, I. & Cotton, R. G. H. (1984) *Biochem. J.* **219**, 891–898
Robson, K. J. H., Beattie, W., James, R. J., Cotton, R. G. H., Morgan, F. J. & Woo, S. L. C. (1984) *Biochemistry* **23**, 5671–5675
Shiman, R. (1980) *J. Biol. Chem.* **255**, 10029–10032
Shiman, R., Gray, D. W. & Pater, A. (1979) *J. Biol. Chem.* **254**, 11300–11306
Smith, S. C., McAdam, W., Cotton, R. G. H. & Mercer, J. F. B. (1985) *Biochem. J.* **231**, 197–199

Received 10 December 1985/18 February 1986; accepted 2 April 1986