Molecular cloning of two human liver 3 α -hydroxysteroid/dihydrodiol dehydrogenase isoenzymes that are identical with chlordecone reductase and bile-acid binder

Yoshihiro DEYASHIKI,* Akihito OGASAWARA,* Toshihiro NAKAYAMA,* Masayuki NAKANISHI,* Yoshiyuki MIYABE,†
Kumiko SATO* and Akira HARA*†

*Biochemistry Laboratory, Gifu Pharmaceutical University, Mitahora-higashi, Gifu 502, Japan, and †Gifu Prefectural Hospital, Noishiki, Gifu 500, Japan

Human liver contains two dihydrodiol dehydrogenases, DD2 and DD4, associated with 3 α -hydroxysteroid dehydrogenase activity. We have raised polyclonal antibodies that cross-reacted with the two enzymes and isolated two 1.2 kb cDNA clones (C9 and C11) for the two enzymes from a human liver cDNA library using the antibodies. The clones of C9 and C11 contained coding sequences corresponding to 306 and 321 amino acid residues respectively, but lacked 5'-coding regions around the initiation codon. Sequence analyses of several peptides obtained by enzymic and chemical cleavages of the two purified enzymes verified that

the C9 and C11 clones encoded DD2 and DD4 respectively, and further indicated that the sequence of DD2 had at least additional 16 residues upward from the N-terminal sequence deduced from the cDNA. There was 82 % amino acid sequence identity between the two enzymes, indicating that the enzymes are genetic isoenzymes. A computer-based comparison of the cDNAs of the isoenzymes with the DNA sequence database revealed that the nucleotide and amino acid sequences of DD2 and DD4 are virtually identical with those of human bile-acid binder and human chlordecone reductase cDNAs respectively.

INTRODUCTION

3 α -Hydroxysteroid dehydrogenase (3 α -HSD) catalyses an NAD(P)⁺-dependent reversible oxidation of 3 α -hydroxy group of various steroids and functions in the metabolism of steroid hormones and bile acids. The enzyme activity is distributed in various mammalian tissues, and the cytosolic enzymes purified from liver [1–5], prostate [6,7], brain [8] and epididymis [9] are monomeric proteins with an M_r of about 35000. Most of the mammalian enzymes transfer hydrogen from the 4-*pro-R* position of the coenzyme to the substrates [5,10,11], and have been classified as the 'A-specific enzyme' (EC 1.1.1.213) to distinguish it from the 'B-specific enzyme' (EC 1.1.1.50). Since 3 α -HSDs of animal livers oxidize *trans*-dihydrodiols of aromatic hydrocarbons and reduce various exogenous carbonyl compounds [1–5], an additional role in xenobiotic metabolism has been suggested for the enzymes. Of the animal 3 α -HSDs, the molecular structure of the rat liver enzyme has been studied extensively [12–14]. The amino acid sequence of rat liver 3 α -HSD [12,13] has a significant sequencing identity with members of the aldo-keto reductase family, which includes prostaglandin F synthase (EC 1.1.1.188) [15], aldehyde reductase (EC 1.1.1.2) and aldose reductase (EC 1.1.1.21) of mammalian tissues [16,17].

Recently, Binstock et al. [18] have isolated a monomeric 3 α -HSD from human liver cytosol and reported that the sequences of several peptides from the enzyme are identical with regions of the deduced amino acid sequence of a cDNA for human liver chlordecone reductase (EC 1.1.1.225) [19]. However, the existence of several chlordecone reductase genes [19] and two proteins immunochemically related to the enzyme in human liver [20] have been reported. In addition, more than two forms of 3 α -HSD have been purified from human liver [21,22]. We previously purified two monomeric dihydrodiol dehydrogenases associated with 3 α -HSD activity from human liver [23] and suggested that

the two enzymes are A-specific 3 α -HSD isoenzymes on the basis of their differences in molecular mass, heat-stability and sensitivity to anti-inflammatory agents and bile acid [24]. In the present study, cDNA clones encoding the two human liver 3 α -HSDs were isolated to confirm that the two enzymes are genetic isoenzymes and to elucidate their structural relationship with rat 3 α -HSD, human liver chlordecone reductase and some other mammalian oxidoreductases.

EXPERIMENTAL

Materials

The λ gt11 human liver cDNA library (HL1001b #2104) was obtained from Clontech Laboratories; nitrocellulose membranes were from Schleicher and Schuell; poly(vinylidene difluoride) membranes were from Millipore; goat anti-rabbit IgG-peroxidase conjugate was from Bio-Rad Laboratories; and bacterial growth media were from Difco. Restriction and DNA-modifying enzymes were purchased from Nippon Gene (Tokyo, Japan) and Takara Shuzou Co. (Kyoto, Japan). The Taq Dye Primer Cycle Sequencing kit and gas-phase-protein-sequencer chemicals were obtained from Applied Biosystems Japan, and a plasmid, pBluescript II SK(+), for subcloning and sequencing, was from Stratagene. All other chemicals were of the highest grade that could be obtained commercially.

Immunochemical experiments

The two 3 α -HSDs, DD2 and DD4, were purified to homogeneity from human liver as previously described [24]. Antibodies against DD2 were raised in rabbits according to the method described by Sawada et al. [25]. The IgG fraction of the antisera was prepared by 0–40 % ammonium sulphate precipitation, and the specificity of the antibodies was analysed by immunoblot analysis [26].

Abbreviation used: HSD, hydroxysteroid dehydrogenase.

† To whom correspondence should be sent.

Screening of cDNA library

The anti-DD2 IgG was used as the probe for cDNA cloning from the λ gt11 human liver cDNA library. Recombinant phages coding for DD2 antigen were identified by the plaque screening method [27] using *Escherichia coli* strain Y1090 as the host. Immunopositive phage plaques transferred from agar to nitrocellulose membranes were detected using the goat anti-rabbit IgG-peroxidase conjugate and 4-chloro-1-naphthol plus H_2O_2 .

DNA sequencing

The cDNA inserts from the positive clones were subcloned into the *Eco*RI sites of the pBluescript II plasmids. For the construction of other sequencing clones with 1.2 kb inserts, some restriction-endonuclease cleavage fragments of the cDNAs were ligated into the appropriate cloning sites of the plasmids. Sequencing reactions were performed with the Dye Primer Cycle Sequencing kit according to the dideoxy chain-termination method [28], and the reaction mixture was analysed using an automated model 373A DNA sequencing system (Applied Biosystems Japan).

Peptide sequencing

Homogeneous DD2 or DD4 (0.2 mg) was directly digested with 2 μ g of lysylendopeptidase (Wako Pure Chemical Industries) in 0.1 M ammonium bicarbonate, pH 8.1, for 8 h at 30 °C. For DD2 it was also cleaved with CNBr after the reductive pyridylethylation of the protein [17]. The peptides were separated by reversed-phase h.p.l.c. (Japan Spectroscopic Co.) on a Develosil 300 ODS-7 column (4.6 mm $\times$ 250 mm; Nomura Science, Seto, Japan) with a 0–60% (v/v) linear gradient of acetonitrile in water containing 0.1% trifluoroacetic acid. The effluent was pooled according to u.v. absorbance at 220 nm and freeze-dried; the pools were separately rechromatographed on the same column with a linear gradient of 0–60% acetonitrile in 10 mM ammonium formate, pH 6.3. Following freeze-drying, samples were sequenced by automated Edman degradation on a 473A protein sequencer (Applied Biosystems Japan). To determine N-terminal sequence of the protein, the sequence analysis was performed with the reduced and pyridylethylated intact proteins before and after digestion of an N-blocked amino acid with acylamino-acid-releasing enzyme (Takara Shuzou) at 37 °C for 6 h in 10 mM sodium phosphate, pH 7.2, containing 1 mM dithiothreitol.

Sequence comparison

DNA sequence data were stored and analysed by the computer program DNASIS (Hitachi Soft Engineering, Osaka, Japan). The amino acid sequences, obtained from peptide sequencing and deduced from the cDNA sequences, were compared against all sequences in the NBRF (PIR R34.0) and EMBL (SWISS-PROT R23.0) databases using the program.

RESULTS AND DISCUSSION

Cloning and cDNA sequence

The polyclonal antibodies against DD2 were not monospecific and reacted with both the antigen and DD4 (Figure 1). This suggests that DD2 and DD4 are immunochemically similar to

each other, despite their apparent structural and functional difference previously observed [24]. The cross-reactivity allowed us to screen cDNA clones encoding the two enzymes from a human liver cDNA library using the antibodies. A total of 200 000 plaques from the cDNA library were screened. In all, 18 immunopositive plaques were obtained, and six clones were isolated. Two had cDNA inserts of about 1.2 kb, and the other four had about 0.8 kb inserts. The two clones, C9 and C11, containing the largest *Eco*RI fragments were subjected to sequence determination on both strands, as shown in Figure 2. The nucleotide and deduced amino acid sequences of C9 and C11 are shown in Figures 3 and 4 respectively.

The cDNA sequence of C9 contained 1162 bp, of which 920 bp constituted an open reading frame specifying a 306-amino-acid polypeptide followed by 242 bp 3'-untranslated region containing a TAA stop codon, two polyadenylation signal-like sequences and a poly(A) track. There was no first ATG codon near the 5'-end of clone C9. The cDNA insert in C11 contained 1164 bp with a 965 bp open reading frame, specifying a 321-amino-acid polypeptide, and a 199 bp 3'-untranslated region containing a TAG stop codon, a polyadenylation hexamer and a poly(A) tail. An ATG codon was found at the 5'-end of this mRNA molecule, but the nucleotide context of this AUG (ACUUCAUGC) does not agree with the initiation consensus sequence (CCACCAUGG); specifically, it is very unusual to find a pyrimidine at the third base upstream from an initiator AUG [29]. The other clones, C8, C10, C16 and C18 were sequenced, but their inserts were *Eco*RI fragments containing bases 1–809 of the sequence of C9 or bases 1–854 of the sequence of C11 (Figure 2).

Amino acid sequences of two 3 α -HSDs

No amino acid was found by direct sequencing of the purified DD2 and DD4, suggesting that the N-termini of the enzymes are blocked. When the two proteins were treated with acylamino-acid-releasing enzyme and analysed with a sequencer, a sequence

Figure 1 Immunoblot analysis of the purified human liver dehydrogenases DD2 and DD4

The pyridylethylated enzymes (0.5 μ g of protein) were subjected to SDS/PAGE. The proteins were transferred by electroblotting to a poly(vinylidene difluoride) membrane, and the membrane was stained with Coomassie Brilliant Blue R-250 (a) and using the antibodies against DD2 as the primary antibody (b). Lanes 1 and 2 are DD4 and DD2 respectively. Arrowheads show the M_r markers.

Figure 2 Sequencing strategy for cDNA clones of C9 and C11

The restriction sites shown were used to generate deletion subclones in pBluescript II plasmid. The dark box (■) and solid line (—) indicate the open reading frame and 3'-untranslated region respectively, and arrows denote the length and direction of DNA sequencing.

of DSKYQCVKLNDGH was yielded from DD2, though no sequence was obtained in the case of DD4. Since the N-terminal sequence of DD2 did not match the deduced amino acid sequences of C9 and C11, the two enzymes were subjected to digestion by lysyl-endopeptidase or CNBr cleavage and sequences of the peptides purified by h.p.l.c. were determined to confirm that we have cloned the cDNAs for DD2 and DD4.

Two peptides (2A and 2D) were obtained from the CNBr cleavage of DD2 and three peptides (2B, 2C and 2E) were from the lysyl-endopeptidase digestion. The sequence data for the five peptides are shown in Table 1. Of the peptides, the sequences of 2C, 2D and 2E matched the regions in the deduced amino acid sequence of clone C9 (Figure 3), but those of the other two peptides [(2A) DSKYQCVKLNDGHF and (2B) LNDGHFMPVLGFGTYAPAEVPK] were not. However, the sequences of the two peptides were overlapped by the italicised six amino acids, and the peptide 2B contained the same sequence of 14 amino acids as that of the N-terminal portion deduced from the cDNA. In addition, peptide 2A contained the 13-amino-acid N-terminal sequence obtained by digestion with acylamino-acid-releasing enzyme of DD2. This indicates that DD2 has at least an additional sequence of 16 residues upward from the N-terminal amino acid predicted from the cDNA (Figure 5). The M_r of DD2 was calculated to be 36793 from the predicted overall amino acid sequence of 322 residues. This value is compatible with that reported for the purified DD2 (M_r 37000) [24]. Therefore the sequence of the enzyme in Figure 5 probably lacks only one acylated N-terminal residue of the intact protein.

On the other hand, all the sequences of the six peptides (4A–4F) obtained from the enzymic cleavage of DD4 (Table 2) were identical with regions in the amino acid sequence deduced from clone C11, especially one sequence of the peptide 4A

located within a region around the first ATG codon in the deduced amino acid sequence (Figure 4). The sequence with 321 amino acids predicted a protein with an M_r of 36847, which is slightly lower than that of the intact DD4 (M_r 39000) previously reported [24]. The result further indicated that the first ATG codon in the cDNA does not initiate translation of the DD4 protein, although the present sequence of DD4 also lacked several N-terminal residues.

Sequence comparison among human liver 3 α -HSDs and chlordecone reductase

DD2 and DD4 have been thought to be 3 α -HSD isoenzymes [24]. The overall amino acid sequences of the two enzymes show 82% identity (Figure 5). Therefore it is appropriate to conclude that the two enzymes are isoenzymes expressed from separate, but similar, genes. Brinstock et al. [18] have shown that sequences of several peptides obtained from a 3 α -HSD of human liver are identical with regions of the amino acid sequence deduced from a cDNA (CCDR12) for human liver chlordecone reductase, which has been thought to comprise 308 amino acids [19]. When the amino acid sequences of DD2 and DD4 were compared with that of the chlordecone reductase cDNA, the 308-amino-acid C-terminal sequence of DD4 was found to be completely identical with the deduced amino acid sequence of CCDR12. The nucleotide sequences of C11 and CCDR12 were also identical, except that the C11 lacks one nucleotide in the 5'-end and has the poly(A)_n compared with the CCDR12 sequence. Thus the isolated cDNA clone encoding DD4, C11, is virtually the same as CCDR12, but the present results demonstrated that the open reading frame of CCDR12 is longer than that previously reported [19]. Although DD4 is probably identical with 3 α -HSD purified

		CT	2
GTC CTG GGA TTT GGC ACC TAT GCG CCT GCA GAG GTT CCT AAA AGT AAA GCT TTA GAG GCC	62		
V L G F G T Y A P A E V P K S K A L E A	20		
2B			
ACC AAA TTG GCA ATT GAA GCT GGC TTC CGC CAT ATT GAT TCT GCT CAT TTA TAC AAT AAT	122		
T K L A I E A G F R H I D S A H L Y N N	40		
GAG GAG CAG GTT GGA CTG GCC ATC CGA AGC AAG ATT GCA GAT GGC AGT GTG AAG AGA GAA	182		
E E Q V G L A I R S K I A D G S V K R E	60		
GAC ATA TTC TAC ACT TCA AAG CTT TGG TGC AAT TCC CAT CGA CCA GAG TTG GTC CGA CCA	242		
D I F Y T S K L W C N S H R P E L V R P	80		
GCC TTG GAA AGG TCA CTG AAA AAT CTT CAA TTG GAT TAT GTT GAC CTC TAC CTT ATT CAT	302		
A L E R S L K N L Q L D Y V D L Y L I H	100		
2C			
TTT CCA GTG TCT GTA AAG CCA GGT GAG GAA GTG ATC CCA AAA GAT GAA AAT GGA AAA ATA	362		
F P V S V K P G E E V I P K D E N G K I	120		
CTA TTT GAC ACA GTG GAT CTC TGT GCC ACG TGG GAG GCC GTG GAG AAG TGT AAA GAT GCA	422		
L F D T V D L C A T W E A V E K C K D A	140		
GGA TTG GCC AAG TCC ATC GGG GTG TCC AAC TTC AAC CGC AGG CAG CTG GAG ATG ATC CTC	482		
G L A K S I G V S N F N R R Q L E M I L	160		
2D			
AAC AAG CCA GGG CTC AAG TAC AAG CCT GTC TGC AAC CAG GTG GAA TGT CAT CCT TAC TTC	542		
N K P G L K Y K P V C N Q V E C H P Y F	180		
AAC CAG AGA AAA CTG CTG GAT TTC TGC AAG TCA AAA GAC ATT GTT CTG GTT GCC TAT AGT	602		
N Q R K L L D F C K S K D I V L V A Y S	200		
GCT CTG GGA TCC CAC CGA GAA GAA CCA TGG GTG GAC CCG AAC TCC CCG GTG CTC TTG GAG	662		
A L G S H R E E P W V D P N S P V L L E	220		
GAC CCA GTC CTT TGT GCC TTG GCA AAA AAG CAC AAG CGA ACC CCA GCC CTG ATT GCC CTG	722		
D P V L C A L A K K H K R T P A L I A L	240		
CGC TAC CAG CTA CAG CGT GGG GTT GTG GTC CTG GCC AAG AGC TAC AAT GAG CAG CGC ATC	782		
R Y Q L Q R G V V V L A K S Y N E Q R I	260		
2E			
AGA CAG AAC GTG CAG GTG TTT GAA TTC CAG TTG ACT TCA GAG GAG ATG AAA GCC ATA GAT	842		
R Q N V Q V F E F Q L T S E E M K A I D	280		
GGC CTA AAC AGA AAT GTG CGA TAT TTG ACC CTT GAT ATT TTT GCT GGC CCC CCT AAT TAT	902		
G L N R N V R Y L T L D I F A G P P N Y	300		
CCA TTT TCT GAT GAA TAT TAA CATGGAGGGCATTGCATGAGGTCTGCCAGAAGGCCCTGCGTGTGGATGGTG	974		
P F S D E Y *	307		
ACACAGAGGATGGCTCTATGCTGTGACTGGACACATCGCCTCTGGTTAAATCTCTCTGCTTGGTGATTTCAGCAAGC	1053		
TACAGCAAAGCCCATTTGGCCAGAAAGGAAAGACAATTAATTTTGTGTTTTTCATTTTGAAGAAAATTAATGCTCTCTCCTA	1132		
AAGATTCTTCACCTAAAAA	1162		

Figure 3 Nucleotide and deduced amino acid sequences of the cDNA insert in clone C9 encoding DD2

The single-letter symbol for each amino acid is placed below the central nucleotide of its codon. The stop codon (TAA) is labelled with an asterisk. Broken underlining indicates putative polyadenylation signal-like sequences. Bold underlining below the amino acid sequence denotes regions confirmed by protein sequence analyses of peptides obtained by digestion with lysyl-endopeptidase (2B, 2C and 2E) and by CNBr cleavage (2D) of the purified DD2. *EcoRI* linkers located at either end are not shown.

by Brinstock et al. [18], it cannot be concluded that DD4 and chlordecone reductase are identical proteins, because of the following findings. First, two clones other than CCDR12 were isolated as the chlordecone reductase cDNAs from human liver cDNA libraries [19], but the authenticity of the cDNAs has not been confirmed by protein sequencing of the purified enzyme. Secondly, two proteins reacting with antibodies against chlordecone reductase were detected in immunoblot analysis of human liver cytosol [20]. Thirdly, chlordecone reductase has been reported not to reduce carbonyl compounds such as pyridine-4-aldehyde and 4-benzoylpyridine [20], whereas DD4 efficiently reduces them [24].

The amino acid sequence of DD2 exhibited the highest sequencing identity with that deduced from another cDNA clone, MDCR2, of those of the three cDNA clones for human liver chlordecone reductase [19]. MDCR2 was cloned from a cDNA library different from that for CCDR12 and is incomplete

in the open-reading-frame-predicted 304 amino acid residues. When the 304-amino-acid C-terminal sequence of DD2 was aligned without amino acid insertions with the deduced amino acid sequence of MDCR2, the identity between the sequences was 97%, with the following eight residue exchanges: amino acids at positions 19, 37, 46, 53, 86, 150, 169 and 171 in the DD2 sequence (Figure 5) are Pro, Val, His, Val, Ser, Met, His and Leu respectively in the deduced amino acid sequence from MDCR2. Since the residue difference between the two sequences is smaller than those between three subunit types (α , β and γ) of human liver class-I alcohol dehydrogenase [30], the translation product of MDCR2 might be another 3 α -HSD isoenzyme. The two clones, CCDR12 and MDCR2, have been cloned from liver cDNA libraries prepared from different individuals, whereas the present clones, C9 and C11, were isolated from the same cDNA library. In addition, three forms of monomeric 3 α -HSD have been purified from human livers by other researchers [21,22],

		AT	2
CCC AAA TAT CAG CGT GTA GAG CTA AAT GAT GGT CAC TTC ***	CCC GTA TTG GGA TTT GGC	62	
P K Y Q R V E L N D G H F M P V L G F G		20	
4A			
ACC TAT GCA CCT CCA GAG GTT CCG AGG AAC AGA GCT GTA GAG GTC ACC AAA TTA GCA ATA		122	
T Y A P P E V P R N R A V E V T K L A I		40	
GAA GCT GGC TTC CGC CAT ATT GAT TCT GCT TAT TTA TAC AAT AAT GAG GAG CAG GTT GGA		182	
E A G F R H I D S A Y L Y N N E E Q V G		60	
CTG GCC ATC CGA AGC AAG ATT GCA GAT GGC AGT GTG AAG AGA GAA GAC ATA TTC TAC ACT		242	
L A I R S K I A D G S V K R E D I F Y T		80	
4B			
TCA AAG CTT TGG TGC ACT TTC TTT CAA CCA CAG ATG GTC CAA CCA GCC TTG GAA AGC TCA		302	
S K L W C T F F Q P Q M V Q P A L E S S		100	
4C			
CTG AAA AAA CTT CAA CTG GAC TAT GTT GAC CTC TAT CTT CTT CAT TTC CCA ATG GCT CTC		362	
L K K L Q L D Y V D L Y L L H F P M A L		120	
4D			
AAG CCA GGT GAG ACG CCA CTA CCA AAA GAT GAA AAT GGA AAA GTA ATA TTC GAC ACA GTG		422	
K P G E T P L P K D E N G K V I F D T V		140	
GAT CTC TCT GCC ACA TGG GAG GTC ATG GAG AAG TGT AAG GAT GCA GGA TTG GCC AAG TCC		482	
D L S A T W E V M E K C K D A G L A K S		160	
ATC GGG GTG TCA AAC TTC AAC TGC AGG CAG CTG GAG ATG ATC CTC AAC AAG CCA GGA CTC		542	
I G V S N F N C R Q L E M I L N K P G L		180	
AAG TAC AAG CCT GTC TGC AAC CAG GTA GAA TGT CAT CCT TAC CTC AAC CAG AGC AAA CTG		602	
K Y K P V C N Q V E C H P Y L N Q S K L		200	
4E			
CTG GAT TTC TGC AAG TCA AAA GAC ATT GTT CTG GTT GCC CAC AGT GCT CTG GGA ACC CAA		662	
L D F C K S K D I V L V A H S A L G T Q		220	
CGA CAT AAA CTA TGG GTG GAC CCA AAC TCC CCA GTT CTT TTG GAG GAC CCA GTT CTT TGT		722	
R H K L W V D P N S P V L L E D P V L C		240	
GCC TTA GCA AAG AAA CAC AAA CGA ACC CCA GCC CTG ATT GCC CTG CGC TAC CAG CTG CAG		782	
A L A K K H K R T P A L I A L R Y Q L Q		260	
CGT GGG GTT GTG GTC CTG GCC AAG AGC TAC AAT GAG CAG CGG ATC AGA GAG AAC ATC CAG		842	
R G V V V L A K S Y N E Q R I R E N I Q		280	
4F			
GTT TTT GAA TTC CAG TTG ACA TCA GAG GAT ATG AAA GTT CTA GAT GGT CTA AAC AGA AAT		902	
V F E F Q L T S E D M K V L D G L N R N		300	
TAT CGA TAT GTT GTC ATG GAT TTT CTT ATG GAC CAT CCT GAT TAT CCA TTT TCA GAT GAA		962	
Y R Y V V M D F L M D H P D Y P F S D E		320	
TAT TAG CATAGAGGGTGTTCACGACATCTAGCAGAAGGCCCTGTGTGTGGATGGTGATGCAGAGGATGTCTCTATG		1039	
Y *		321	
CTGGTGACTGGACACACGGCCTCTGGTTAAATCCCTCCCTCCTGCTTGGCAACTTCAGCTAGCTAGATATATCCATGG		1118	
TCCAGAAAGCAACATAATAATTTTATCTTGAAGTAAAAA		1164	

Figure 4 Nucleotide and deduced amino acid sequences of the cDNA insert in clone C11 encoding DD4

The deduced amino acids, stop codon (TAG) and putative polyadenylation signal are shown as in Figure 3. The first ATG codon is dotted. Bold underlining denotes regions confirmed by protein sequencing of peptides obtained by digestion with lysyl-endopeptidase (4A–4F) of the purified DD4. The peptide sequences (4C and 4E) include some residues (▲) that were not identified.

and we previously isolated other monomeric dihydrodiol dehydrogenases associated with 3(20) α -HSD activity, in addition to DD2 and DD4, from a human liver [23]. These findings, together with the present results, suggest that the multiplicity of hepatic 3 α -HSD forms is due to the multiplicity and/or polymorphism of 3 α -HSD genes, although we cannot exclude the possibility that some forms of 3 α -HSD are derived from post-translational modification of the limited 3 α -HSD isoenzymes.

Sequence identity among human liver 3 α -HSDs and other proteins

A computer-based comparison of the sequence of DD2 with the entries in the NBRF and EMBL databases revealed that its nucleotide sequence was identical with that of the full-length cDNA encoding human liver bile-acid binder [31], which was recently published as the present paper was being prepared. The present amino acid sequence of DD2 (Figure 5) lacks only the

first Met of the deduced sequence from the bile-acid-binder cDNA. The N-terminal sequence analysis of DD2 suggested that the Met of the protein is probably acylated. Bile-acid binder purified from human liver is also a monomeric protein of M_r 36000 that exhibits dihydrodiol dehydrogenase activity [32]. Although the bile-acid binder was reported not to be co-purified with 3 α -HSD activity [22,32] and the recombinant protein is devoid of 3 α -HSD activity [31], bile acids such as lithocholic acid and chenodeoxycholic acid were used as the substrates to detect the 3 α -HSD activity in these studies, but 3 α -hydroxysteroids such as 3 α -hydroxy-5 β -androstan-17-one and 5 β -pregnane-3 α ,20 β -diol have not been tested as substrates for the assay of 3 α -HSD activity. DD2 oxidizes the limited C₁₉- and C₂₁-steroids, but not bile acids, which conversely act as competitive inhibitors for this enzyme [24]. These findings seem to support the identity of the 3 α -HSD isoenzyme, DD2, with bile-acid binder. However, it has been reported that at least five amino acid residues in the

Table 1 Automated Edman degradation of DD2 peptides generated by CNBr cleavage and lysyl-endopeptidase digestion

The peptides were sequenced by a protein sequencer, and the phenylthiohydantoin-amino acids released in each cycle were quantified against a standard mixture containing 100 pmol of each phenylthiohydantoin-amino acid. The average repetitive yields for sequencing of 2A, 2D, 2B, 2C and 2E were 96, 96, 89, 87 and 92% respectively.

Cycle no.	Peptide from CNBr cleavage				Peptide from enzymic digestion					
	2A		2D		2B		2C		2E	
	Residue	Yield (pmol)	Residue	Yield (pmol)	Residue	Yield (pmol)	Residue	Yield (pmol)	Residue	Yield (pmol)
1	Asp	43	Ile	38	Leu	122	Asn	49	Ser	22
2	Ser	21	Leu	33	Asn	24	Leu	124	Tyr	41
3	Lys	115	Asn	10	Asp	102	Gln	96	Asn	19
4	Tyr	105	Lys	22	Gly	108	Leu	101	Glu	29
5	Gln	72	Pro	19	His	23	Asp	66	Gln	31
6	Cys	67	Gly	36	Phe	114	Tyr	104	Arg	11
7	Val	119	Leu	30	Met	88	Val	152	Ile	45
8	Lys	86	Lys	25	Pro	41	Asp	66	Arg	11
9	Leu	100	Tyr	28	Val	56	Leu	86	Gln	32
10	Asn	79	Lys	27	Leu	48	Tyr	105	Asn	23
11	Asp	51	Pro	12	Gly	53	Leu	89	Val	35
12	Gly	80	Val	19	Phe	51	Ile	133	Gln	25
13	His	89	Cys	10	Gly	51	His	39	Val	27
14	Phe	22	Asn	14	Thr	13	Phe	94	Phe	35
15			Gln	14	Tyr	25	Pro	30	Glu	8
16			Val	18	Ala	29	Val	70	Phe	25
17			Glu	9	Pro	22	Ser	5	Gln	19
18			Cys	5	Ala	28	Val	52	Leu	17
19					Glu	11	Lys	28	Thr	7
20					Val	22	Pro	13	Ser	3
21					Pro	12	Gly	43	Glu	8
22					Lys	7	Glu	10	Glu	8
23							Glu	11	Met	4
24							Val	11	Lys	3
25							Ile	9		
26							Pro	2		
27							Lys	1		

DD2 *DSKYQCVKLNDSGHFMPVLGFGTYAPAEVPSKALEATKLAIEAGFRHIDSAHLYNNEEQV*
 DD4 *.FKYQRVELNDGHPFVLGFGTYAPPEVPRNRAVEVTKLAIAGFRHIDSAHLYNNEEQV*

DD2 *GLAIRSKIADGSVKREDIFYTSKLWCNSHPELVPRPALERSLKNLQLDYVDLYLHFPVS*
 DD4 *GLAIRSKIADGSVKREDIFYTSKLWCFTFPQPMQPALESSLKLQLDYVDLYLHFPMA*

DD2 *VKPGEEVVPKDENGKILFDTVDLCATWEAVEKCKDAGLAKSIGVSNFNRRQLEMILNKP*
 DD4 *LKPGETPLPKDENGKILFDTVDLSATWEVMEKCKDAGLAKSIGVSNFNCRQLEMILNKP*

DD2 *LKYKPCVNCQVECHPYFNQKRLDFCKSKDIVLVAYSALGSHREEPWDPNPSVLLEDPVL*
 DD4 *LKYKPCVNCQVECHPYLNQSKLDFCKSKDIVLVASHALGTQRHKLWDPNPSVLLEDPVL*

DD2 *CALAKKHKRTPALIALRYQLQRGVVVLAKSYNEQRIQNVQVFEPQLTSEEMKAIDGLNR*
 DD4 *CALAKKHKRTPALIALRYQLQRGVVVLAKSYNEQRIQNVQVFEPQLTSEDMKVLGDGLNR*

DD2 *NVRYLTLDIFAGPPNYPFSDEY*
 DD4 *NYRVVMDPLMDHPDYPFSDEY*

Figure 5 Comparison of predicted sequences of DD2 and DD4

The sequence of DD2 includes the N-terminal portion that was determined only by sequencing of peptides 2A and 2B and is indicated in *italics*. Asterisks indicate identical amino acid residues between the two sequences. The putative coenzyme-binding and active-site sequences proposed for rat liver 3 α -HSD [12,14] are underlined below the sequence of DD4.

sequences of peptides from the purified human liver bile-acid binder do not match the sequence deduced from the full-length

cDNA for this protein [31], whereas the sequences (84 amino acids in total) of the four peptides (2B–2E) of DD2 were completely identical with the regions in the deduced sequence of clone C9. Further studies will be necessary to clarify the identity of DD2 and bile-acid binder in human liver, although the present clone C9 and the clone for the bile-acid binder are identical cDNAs.

In pairwise alignments, DD2 and DD4 revealed at least 48% amino acid sequence identity with the aldo-keto reductase family of proteins such as bovine lung prostaglandin F synthase [15], rat liver 3 α -HSD [12,13] and human aldehyde and aldose reductases [16], as reported with the sequence deduced from cDNAs for human liver chlordecone reductase [18,19] and bile-acid binder [31]. Of the aldo-keto reductase family of proteins, rat liver 3 α -HSD is similar to the two human 3 α -HSDs in dual-coenzyme specificity for NADP(H) and NAD(H), specificity for dihydrodiols, alicyclic alcohols and carbonyl compounds, and sensitivity to anti-inflammatory agents [1,2,23,24]. DD2 and DD4 exhibited relatively high sequence identities of about 68% with rat liver 3 α -HSD, and showed conservation of two structural domains that have been found in rat liver 3 α -HSD [12–14]. One is the 'fingerprint' structural motif in the coenzyme-binding sites of many dehydrogenases, GXGXXA (where X is any amino acid) [12,13,33,34], that is seen between residues 19 and 24 of the sequences of DD4 and DD2 (Figure 5). In recent crystallographic studies on the three-dimensional structure of aldose reductase, which has sequence similarity to rat 3 α -HSD, the coenzyme-

Table 2 Automated Edman degradation of DD4 peptides generated by lysyl-endopeptidase digestion

The peptides were sequenced, and the phenylthiohydantoin-amino acids released in each cycle were quantified as described in Table 1. Peptides 4C and 4E contained unidentifiable phenylthiohydantoin-amino acids that are indicated as '—'. The average repetitive yields for sequencing of 4A, 4B, 4C, 4D, 4E and 4F were 86, 87, 92, 88, 86 and 100% respectively.

Peptide ...	4A		4B		4C		4D		4E		4F	
Cycle no.	Residue	Yield (pmol)	Residue	Yield (pmol)	Residue	Yield (pmol)	Residue	Yield (pmol)	Residue	Yield (pmol)	Residue	Yield (pmol)
1	Tyr	123	Arg	7	Leu	21	Lys	64	Tyr	58	Ser	10
2	Gln	87	Glu	29	Trp	36	Leu	116	Lys	59	Tyr	57
3	Arg	21	Asp	38	—	—	Gln	67	Pro	30	Asn	51
4	Val	114	Ile	65	Thr	3	Leu	139	Val	36	Glu	28
5	Glu	45	Phe	71	Phe	5	Asp	79	—	—	Gln	26
6	Leu	79	Tyr	49	Phe	7	Tyr	64	Asn	8	Arg	11
7	Asn	12	Thr	21	Gln	4	Val	87	Gln	9	Ile	35
8	Asp	16	Ser	5	Pro	3	Asp	29	Val	23	Arg	10
9	Gly	26	Lys	24	Gln	3	Leu	60	Glu	10	Glu	10
10	His	17	—	—	Met	2	Tyr	42	—	—	Asn	11
11	Phe	38	—	—	Val	3	Leu	61	His	10	Ile	21
12	Met	32	—	—	Gln	3	Leu	59	Pro	14	Gln	8
13	Pro	31	—	—	Pro	2	His	17	Tyr	13	Val	13
14	Val	19	—	—	Ala	2	Phe	53	Leu	13	Phe	10
15	Leu	20	—	—	Leu	3	Pro	36	Asn	6	Glu	5
16	Gly	27	—	—	Glu	2	Met	28	Gln	6	Phe	8
17	Phe	22	—	—	—	—	Ala	37	Ser	4	Gln	6
18	Gly	28	—	—	—	—	Leu	39	Lys	6	Leu	8
19	Thr	8	—	—	Leu	4	Lys	29	—	—	Thr	6
20	Tyr	34	—	—	Lys	2	Pro	26	—	—	Ser	4
21	Ala	14	—	—	—	—	Gly	34	—	—	Glu	6
22	Pro	7	—	—	—	—	Glu	11	—	—	Asp	6
23	Pro	9	—	—	—	—	Thr	6	—	—	Met	4
24	Glu	4	—	—	—	—	Pro	6	—	—	Lys	3
25	Val	9	—	—	—	—	Leu	7	—	—	—	—
26	—	—	—	—	—	—	Pro	4	—	—	—	—
27	—	—	—	—	—	—	Lys	1	—	—	—	—

binding site exhibits a single domain folded in an eight-stranded parallel β/α barrel, where 17 residues involved in the NADPH binding have been shown to exist in the middle and C-terminal regions of the enzyme sequence as well as in the N-terminal region [35,36]. Most of the critical NADPH-binding residues are conserved in the sequences of DD2 and DD4, which suggests that the residues other than the N-terminal residues in the 'fingerprint' motif are also responsible for the coenzyme binding of the human liver 3 α -HSDs. Another structural domain common to the human and rat 3 α -HSDs is a sequence of SIGVSNFNCR (residues between 161 and 170 in Figure 5). The sequence has been identified as an active-site peptide of rat liver 3 α -HSD, and the Cys in the sequence has been suggested to be critical to the catalytic function of rat liver 3 α -HSD [14]. This Cys is conserved in DD4, whereas it is replaced by Arg in DD2. The broad specificity of DD4 for androgens, pregnanes and bile acids is similar to that of rat liver 3 α -HSD, but DD2 catalyses the oxidoreduction of some androgens and pregnanes [24]. The substitution of Cys with Arg might contribute to the difference in steroid specificity between DD2 and DD4. Alternatively, the substitution might be related to the difference in the sensitivity to bile acids between the two human liver enzymes, of which DD2 is potentially inhibited [24]. Since the carboxy group at C-24 on free bile-acid molecules, competitive inhibitors with respect to substrate, has been suggested to be important for the binding to DD2 [24], the electrostatic interaction between the anionic carboxy group of bile acid and the basic amino acid residue might be involved in the binding of inhibitory bile acid to the active site of the enzyme. To clarify the two possibilities, our

efforts are currently focused on the structure-function relationship of DD2 by affinity-labelling using bile acid derivatives and by site-directed mutagenesis of the important Arg as the target.

This work was supported by the Tokyo Biochemical Research Foundation and by the Kato Memorial Bioscience Foundation.

REFERENCES

- 1 Wörner, W. and Oesch, F. (1984) FEBS Lett. **170**, 263–267
- 2 Penning, T. M., Mukharji, I., Barrows, S. and Talalay, P. (1984) Biochem. J. **222**, 601–611
- 3 Ohmura, M., Hara, A., Nakagawa, M. and Sawada, H. (1990) Biochem. J. **266**, 583–589
- 4 Hara, A., Inoue, Y., Nakagawa, M., Nagane, F. and Sawada, H. (1988) J. Biochem. (Tokyo) **103**, 1027–1034
- 5 Hara, A., Nakagawa, M., Taniguchi, H. and Sawada, H. (1989) J. Biochem. (Tokyo) **106**, 900–903
- 6 Taurog, J. D., Moore, R. J. and Wilson, J. D. (1975) Biochemistry **14**, 810–817
- 7 Amet, Y., Simon, B., Quemener, E., Mangin, P., Floch, H. H. and Abalain, J. H. (1992) J. Steroid Biochem. Mol. Biol. **41**, 689–692
- 8 Penning, T. M., Sharp, R. B. and Krieger, N. R. (1985) J. Biol. Chem. **260**, 15266–15272
- 9 Hastings, C. D. and Hansson, V. (1979) Int. J. Androl. **2**, 263–274
- 10 Berseus, O. and Björkhem, I. (1967) Eur. J. Biochem. **2**, 503–507
- 11 Sawada, H., Hara, A., Ohmura, M., Nakayama, T. and Deyashiki, Y. (1991) J. Biochem. (Tokyo) **109**, 770–775
- 12 Pawlowski, J. E., Huizinga, M. and Penning, T. M. (1991) J. Biol. Chem. **266**, 8820–8825

- 13 Stolz, A., Rahim-Kiani, M., Ameis, D., Chan, E., Ronk, M. and Shively, J. E. (1991) *J. Biol. Chem.* **266**, 15253–15257
- 14 Penning, T. M., Abrams, W. and Pawlowski, J. E. (1991) *J. Biol. Chem.* **266**, 8826–8834
- 15 Watanabe, K., Fujii, Y., Nakayama, K., Ohkubo, H., Kuramitsu, S., Kagamiyama, H., Nakanishi, S. and Hayaishi, O. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 11–15
- 16 Bohren, K. M., Bullock, B., Wermuth, B. and Gabbay, K. H. (1989) *J. Biol. Chem.* **264**, 9547–9551
- 17 Schade, S. Z., Early, S. L., Williams, T. R., Kezdy, F. J., Heinrikson, R. L., Grimshaw, C. E. and Doughty, C. C. (1990) *J. Biol. Chem.* **265**, 3628–3635
- 18 Brinstock, J. M., Iyer, R. B., Hamby, C. V., Fried, V. A., Schwart, I. S., Weinstein, B. I. and Southren, A. L. (1992) *Biochem. Biophys. Res. Commun.* **187**, 760–766
- 19 Winters, C. J., Molowa, D. T. and Guzelian, P. S. (1990) *Biochemistry* **29**, 1080–1087
- 20 Molowa, D. T., Shayne, A. G. and Guzelian, P. S. (1986) *J. Biol. Chem.* **261**, 12624–12627
- 21 Kudo, K., Amuro, Y., Hada, T. and Higashino, K. (1990) *Biochim. Biophys. Acta* **108**, 250–254
- 22 Takikawa, H., Fujimoto, M., Nishikawa, K. and Yamanaka, M. (1992) *Hepatology* **16**, 365–371
- 23 Hara, A., Taniguchi, H., Nakayama, T. and Sawada, H. (1990) *J. Biochem. (Tokyo)* **108**, 250–254
- 24 Deyashiki, Y., Taniguchi, H., Amano, T., Nakayama, T., Hara, A. and Sawada, H. (1992) *Biochem. J.* **282**, 741–746
- 25 Sawada, H., Hara, A., Hayashibara, M., Nakayama, T. and Usui, S. (1984) *Biochim. Biophys. Acta* **799**, 322–325
- 26 Towbin, H., Staehelin, T. and Gordon, J. (1979) *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4350–4354
- 27 Hunyu, T. V., Young, R. A. and Davis, R. W. (1985) in *DNA Cloning: A Practical Approach* (Glover, D. M., ed.), vol. 1, pp. 49–78, IRL Press, Oxford
- 28 Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) in *Molecular Cloning: A Laboratory Manual*, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- 29 Kozak, M. (1989) *J. Cell Biol.* **108**, 229–241
- 30 Jörnvall, H., von Bahr-Lindström, H. and Höög, J.-O. (1989) in *Human Metabolism of Alcohol* (Crow, K. E. and Batt, R. D., eds.), vol. 2, pp. 43–64, CRC Press, Boca Raton, FL
- 31 Stolz, A. A., Hammond, L., Takikawa, H., Ronk, M. and Schively, J. E. (1993) *J. Biol. Chem.* **268**, 10448–10457
- 32 Takikawa, H., Stolz, A., Sugiyama, Y., Yoshida, H., Yamanaka, M. and Kaplowitz, N. (1990) *J. Biol. Chem.* **265**, 2132–2136
- 33 Wierenga, R. K., De Maeyer, M. C. H. and Hol, W. G. J. (1985) *Biochemistry* **24**, 1346–1357
- 34 Scrutton, N. S., Berry, A. and Perham, A. (1990) *Nature (London)* **343**, 38–43
- 35 Rondeau, J. M., Tête-Favier, F., Podjarny, A., Reymann, J. M., Barth, P., Biellmann, J.-F. and Moras, D. (1992) *Nature (London)* **355**, 469–472
- 36 Wilson, D. K., Bohren, K. M., Gabbay, K. H. and Quiocho, F. A. (1992) *Science* **275**, 81–84