

- flavoproteins (Bray, R. C., Engel, P. C. and Mayhew, S. G., eds.), pp. 695–698, Walter de Gruyter, Berlin
- 30 Gardlik, S. and Rajagopalan, K. V. (1990) *J. Biol. Chem.* **265**, 13047–13054
- 31 Gardlik, S. and Rajagopalan, K. V. (1991) *J. Biol. Chem.* **266**, 4889–4895
- 32 Gutteridge, S. and Bray, R. C. (1980) *Biochem. J.* **189**, 615–623
- 33 Bray, R. C. and George, G. N. (1985) *Biochem. Soc. Trans.* **13**, 560–567
- 34 Hille, R. and Sprecher, H. (1987) *J. Biol. Chem.* **262**, 10914–10917
- 35 Holm, R. H. (1990) *Coordination Chem. Rev.* **100**, 183–221
- 36 Greenwood, R. J., Wilson, G. L., Pilbrow, J. R. and Wedd, A. G. (1993) *J. Am. Chem. Soc.* **115**, 5385–5392
- 37 Enemark, J. H. and Young, C. G. (1993) *Adv. Inorg. Chem.* **40**, 1–88
- 37a Howes, B. D., Bray, R. C., Richards, R. L., Turner, N. A., Bennett, B. and Lowe, D. J. (1996) *Biochemistry*, in the press
- 38 Howes, B. D., Bennett, B., Bray, R. C., Richards, R. L. and Lowe, D. J. (1994) *J. Am. Chem. Soc.* **116**, 11624–11625
- 39 Gambarotta, S., Floriani, C., Chiesi-Villa, A. and Guastini, C. (1985) *J. Am. Chem. Soc.* **107**, 2985–2986
- 40 Hawkes, T. R., George, G. N. and Bray, R. C. (1984) *Biochem. J.* **218**, 961–968
- 41 Coucouvanis, D., Toupadakis, A., Lane, J. D., Koo, S. M., Kim, C. G. and Hadjikyriacou, A. (1991) *J. Am. Chem. Soc.* **113**, 5271–5282

Received 17 August 1995

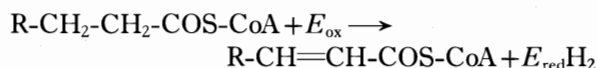
184

## Molecular evolution and substrate specificity of acyl-CoA dehydrogenases: chimaeric 'medium/long' chain-specific enzyme from medium-chain acyl-CoA dehydrogenase

A. Nandy, B. Kuchler and S. Ghisla

Fachbereich Biologie der Universität, Postfach 5560-M644, D-78434 Konstanz, Germany

Acyl-CoA dehydrogenases (ADHs) from mammalian mitochondria catalyse the  $\alpha,\beta$ -dehydrogenation of fatty chain acyl-CoA conjugates, the first and rate-limiting step in the  $\beta$ -oxidation cycle:



They are members of a superfamily to which belong also peroxisomal acyl-CoA oxidases (AOX) [1,2] and enzymes involved in the catabolism of branched-chain amino acids. Related enzymes from lower organisms, which are assumed to function by the same basic chemical mechanism, can also be appended to this family. The acyl-CoA dehydrogenases differ in their specificity for the configuration of the substrate acyl chain. The three-dimensional structure of medium-chain acyl-CoA dehydrogenase

(MCADH) [3–5], probably the best studied member of the family, and that of a bacterial butyryl-CoA dehydrogenase [6] have recently been resolved. That of the long-chain enzyme (LCADH) has been modelled on the structure of MCADH [7]. The sequences of approx. three dozen enzymes belonging to the superfamily have been reported and are available from public databases [8].

While the chemical mechanism of the  $\alpha,\beta$ -dehydrogenation is quite clear [9–11], we are only at the beginning of understanding the various modes by which the protein activates the substrate(s) and brings about catalysis [11,12]. However, this is only one main topic of current research with ADHs. A second point of interest deals with the molecular mechanisms that bring about chain length specificity of ADHs. This point will be addressed to some extent in this contribution.

### Functional groups of catalytic importance

Substrate  $\alpha,\beta$ -dehydrogenation is initiated by abstraction of the substrate  $\alpha$ -H as  $\text{H}^+$  by a base, which has turned out to be a glutamate- $\text{COO}^-$  in all enzymes studied to date. Most interestingly, the position of this Glu in the sequence is not

Abbreviations used: ADH, acyl-CoA dehydrogenase; AOX, acyl-CoA oxidase; MCADH, (human wild-type) medium chain ADH; LCADH, long-chain ADH; MLCADH: medium/long chain ADH (T255E/E376G MCADH); Glu/Glu-MCADH, Thr-255-Glu-MCADH; IVDH, isovaleryl-CoA dehydrogenase; SCADH, short-chain ADH; SBCADH, short, branched-chain acyl-CoA dehydrogenase; VLCADH, very long-chain ADH; GDH, glutaryl-CoA dehydrogenase.

Figure 1

Sequence alignments of active centre domains of acyl-CoA dehydrogenases and oxidases

The alignments were done following the criteria outlined in the text. The highlighted residues are those at either position 255 or 376, this numbering referring to the sequence of mature MCADH. 2MBCER, 2-methyl, branched-chain enoyl-CoA reductase; BDH, butyryl-CoA dehydrogenase; ADH, no chain specificity reported; *caiA* and *aidB*, products from *Escherichia coli* genes; AOX members of the *Candida* multigene families are not specified individually.

		Helix G										Helix J										Helix K									
MCADH	human	243	AGFKVA-MGAFDK	T	RPVVAAGAVGLAQRALDEATKYAL	ERKTFGKLLVEH	291	---	352	GGNGFNTPEY	VEKLMRDAKIYQ	IY	IE	GT	SQIQRLLIVAREHD	392															
MCADH	pig	243	AGFKIA-MGTFDK	T	RPPVAAGAVGLAQRALDEATKYAL	ERKTFGKLLAEH	291	---	352	GGNGFNTPEY	VEKLMRDAKIYQ	IY	IE	GT	AQIQRLIIAREHIG	392															
MCADH	rat	243	AGFKIA-MGAFDR	T	RPTVAAGAVGLAQRALDEATKYAL	DRKTFGKLLVEH	291	---	352	GGYGFNTPEY	VEKLMRDAKIYQ	IY	IE	GT	AQIQRLIIAREHIE	392															
MCADH	mouse	243	AGFKIA-MGAFDR	T	RPTVAAGAVGLAQRALDEATKYAL	DRKTFGKLLVEH	291	---	352	GGYGFNTPEY	VEKLMRDAKIYQ	IY	IE	GT	AQIQRLIIAREHIE	392															
2MBCER	<i>A. suum</i>	241	EGFVVA-MRTFDR	T	RPVAAGLGTGLAQRCLDEASKYAL	ERKTFGIPIANH	289	---	350	GGNGYNSPEY	VEKLMRDAKLFQ	IY	IE	GT	CQIQRVVVSRLNLS	390															
SCADH	human	235	MGFKIA-MQTLDM	G	RIGIASQALGIAQALDCAVNYAE	NRMAGFAPLTKL	283	---	344	GGMGYVTEMP	AERYHRDARITE	IY	IE	GT	SEIQRLLVIAGHLR	384															
SCADH	rat	235	MGFKIA-MQTLDM	G	RIGIASQALGIAQALDCAVNYAE	NRMAGFAPLTKL	283	---	344	GGMGYVTEMP	AERYHRDARITE	IY	IE	GT	SEIQRLLVIAGHLR	384															
SCADH	mouse	235	MGFKIA-MQTLDM	G	RIGIASQALGIAQALDCAVNYAE	NRMAGFAPLTKL	283	---	344	GSMGYVTEMP	AERYHRDARITE	IY	IE	GT	SEIQRLLVIAGHLR	384															
BDH	<i>M. elsdenii</i>	238	SGFKIA-MMTLDG	G	RIGVAAQALGIAEAALADAVEYK	NRQVFGKPLCKF	282	---	343	GGYGYSEEP	VARHMRDAKITQ	IY	IE	GT	NEVQLMVTGGALLR	383															
SBADH	human	248	HGYKYA-IGSLNE	G	RIGIAAQMGLGAQGCDFYITPYIK	ERIQFGKRLPDF	296	---	357	GGVGYTKDP	VEKYFRDAKITQ	IY	IE	GA	SNIQNLNTIAKHIDA	397															
LCADH	human	250	KGFYYL-MKELPQ	G	RLLIADVAISATEFMEFEETRNYVR	QRKAFGKTVAHN	298	---	359	GGWGYMMEYP	IAKAYVDARVQP	IY	IE	GT	NEIMKELIARQIVS	399															
LCADH	rat	250	KGFYYL-MKELPQ	G	RLLIADVAISATEFMEFEETRNYVR	QRKAFGKTVAHN	298	---	359	GGWGYMMEYP	IAKAYVDARVQP	IY	IE	GT	NEIMKELIARQIVS	399															
IVDH	human	242	KGVYVL-MSGLDL	G	RLVLAGGPLGLMQLVLDHTIPYLH	VREAFGQKIGHF	290	---	351	GGNGYINDFP	MGRFLRDAKLYE	IG	A	GT	SEVRRIVIGRAFNA	391															
IVDH	rat	242	KGVYVL-MSGLDL	G	RLVLAGGPLGLMQLVLDHTIPYLH	VREAFGQKIGHF	290	---	351	GGNGYINDFP	MGRFLRDAKLYE	IG	A	GT	SEVRRIVIGRAFNA	391															
VLCADH	human	288	SGFKVA-MHILNN	G	RFGMAAALAGTMRGIKAVDHT	NRTQFGDKIHNH	336	---	398	GGMGFMKEPG	VERVLRLRIFR	IF	GT	NDILRLFVALQGCC	438																
VLCADH	rat	288	SGFKVA-MHILNN	G	RFGMAAALAGTMRGIKAVDHT	NRTQFGDKIHNH	336	---	398	GGMGFMKEPG	VERVLRLRIFR	IF	GT	NDILRLFVALQGCC	438																
GDH	human	244	SSLGGP-FGCLNN	A	RYGIWGLVGAEEFLHTARQYAL	DRMQFGVPLARN	292	---	353	GGNGISDEYH	VIHAMNLEAVN	TY	GT	HDHALILGRAITG	393																
GDH	pig	245	SSLAVP-FGCLNN	A	RYGITSWGLVGAEEFLHTARQYAL	DRMQFGVPLARN	292	---	354	GGNGISDEYH	VIHAMNLEAVN	TY	GT	HDHALILGRAITG	394																
GDH	mouse	244	SSLGGP-FGCLNT	A	RYGITWGLVGAEEFLHTARQYAL	DRMQFGVPLARN	292	---	353	RGNGISDEYH	VIHAMNLEAVN	TY	GT	HDHALILGRAITG	393																
ADH	<i>M. leprae</i>	239	TGFKTA-LATLDH	T	RPTIGAQVGLAQGALDAAIVYTK	ERLQFGESISTF	287	---	349	GGAGYSDFP	VERFMRDAKITQ	IY	GT	NIQIQRVVMSRALLR	389																
ADH	<i>C. elegans</i>	177	RGFAYQ-MNQFND	G	RLVTVAVGLLPLQKCNETIEYAR	DRKLFKGLLDQ	224	---	286	GGAGYLDNDG	ISRFRDFRIFS	IG	CA	DEVMMQIIHKQTSK	326																
<i>caiA</i>	<i>E. coli</i>	228	NGFNRY-KEEFDH	G	RFLVALTYGTAMCAFEDAARYAN	QRQVGEAIGRF	276	---	337	GGVGIAGNHR	ISRFRDLRVD	VS	GS	DEMQLTLGRAVLK	377																
<i>aidB</i>	<i>E. coli</i>	286	EGIRLI-LKMGGM	T	RFDALGSHAMRRRAFSLAIYHAH	QRHVFGNPLIQ	334	---	401	GGIYCEESE	LPRLYREMPVNS	IW	IE	GS	GNIMCLDLVRLVNLK	441															
AOX	rat	269	LSNKLTYGTMVF	V	RSFLVGNAAQSLSKACTIAIRYSA	VRRQSEIKQSEP	317	---	397	GGHGYSHSSG	IPNIYVTFTPAC	TF	GE	NTVMMQLQARFLMK	437																
AOX	human	269	LSNKLTYGTMVF	V	RSFLVGEAARLSKACTIAIRYSA	VRRQSEMKGEP	317	---	397	GGHGYSHCSG	LPNIYVNFPSG	TF	GE	NTVMMQLQARFLMK	437																
AOX	<i>C. maltosa</i>	309	PLEQLS-YISLLE	G	RVMVLDSDYRMLARVSTIALRYAV	ARRQFSKGDQGP	357	---	448	GGHGYSAAYG	FGKTYGDWAVQC	TW	GD	NNVLGMSAGKTIK	488																
AOX	<i>C. maltosa</i>	287	PLEQLS-YISLLE	G	RVMVLDSDYRMLARVSTIALRYAV	ARRQFSKGDQGP	335	---	432	GGHGYSSYNG	FAKAFNDWVQC	TW	GD	NNVLSLVGKPIIK	472																
AOX	<i>C. tropicalis</i>	309	PLEQLS-YISLLE	G	RVMVLDSDYRMLARVSTIALRYAV	ARRQFSKKAQEP	357	---	448	GGHGYSSYNG	FGKTYDDWVQC	TW	GD	NNVLAMSAGKTIK	488																
AOX	<i>C. tropicalis</i>	287	PLEQLS-YISLLE	G	RVMVLDSDYRMLARVSTIALRYAV	ARRQFSKGDVDP	335	---	432	GGHGYSSYNG	FGKAYNDWVQC	TW	GD	NNVLAMSAGKPIVK	472																
AOX	<i>C. tropicalis</i>	287	PLEQLS-YISLLE	G	RVMVLDSDYRMLARVSTIALRYAV	ARRQFSKGDVDP	335	---	432	GGHGYSSYNG	FGKAYNDWVQC	TW	GD	NNVLAMSAGKPIVK	472																
AOX	<i>C. tropicalis</i>	272	PSEQLS-YISLIG	G	RVTMMVDSYRMTSRFRTIALRYAI	HRROFKKDDTD	308	---	412	GGHGYSSYNG	FGKAYSDWVQC	TW	GD	NNVLAMVAKPMVR	452																
AOX	<i>S. cerevisiae</i>	324	SLDQISGYSALLS	G	RVMVMVDSFRFGSKFATIAVRYAV	GRQQFAPRKGLS	373	---	464	GGHGYSSYNG	FGKAYSDWVQC	TW	GD	NNVLSLTSAKSILK	504																
AOX	<i>C. elegans</i>	281	QHFKVG-YTTMLY	M	RSEMIYHQAYLLAMAMASIRYSA	VRRQGEIKPGTQ	329	---	409	GGHGYSDASY	LPTLYTCSVGAC	TY	GE	NNVMLLQLSKYLK	449																
AOX	barley (fragment)							---	36	GGHGYLNSSG	LPFLFAYVVPAC	TY	GE	NNVMLLHVARFLMK	76																
consensus ADH (75%)			-G-----	R--A--G-----Y--		-R--FG-----			GG-G-----RD-----	I--GT--I-----																					
consensus AOX (75%)			---L--Y-----	R--M-----R-----TIA-RY--		-RRQ-----			GGHYS---G-----Y-----C T-E GD N-V-----K																						
invariant ADH+AOX				R		R			G G					G																	

conserved [1,4,13,14]. In the majority of cases a glutamate lies at a position that corresponds to position 376 of the mature sequence of MCADH, as shown by the sequence alignments of Figure 1. Helices J and K of MCADH flank its active site; Glu-376 is positioned on the loop that interconnects these helices, which, together with the aromatic Re plane of the flavin, sandwich the bound substrate [5]. In contrast, in LCADH and isovaleryl-CoA dehydrogenase, a Glu to which a corresponding H<sup>+</sup>-abstracting function has been assigned is positioned on a helix that corresponds to helix G of MCADH. Helix G also borders the active site cavity; it is positioned approximately parallel to helices J and K, and the Glu in question (at position 255 in the MCADH sequence) extends to occupy a similar position to that of Glu-376. Glu-255 carries out the same function in LCADH [7] as does Glu-376 in MCADH [15] since mutagenesis of both residues produces similar effects on activity. Interestingly, Glu-376 in MCADH can be replaced by various bases with the retention of activity. Thus, the Glu-376→Asp, Glu-376→His

and Glu-376→Cys variants show some degree of activity (V. Kieweg, P. Brass, A. Nandy and S. Ghisla, unpublished work). The mode of proton abstraction in ADHs seems to be unique in that the H<sup>+</sup>-abstracting Glu<sup>-</sup> apparently does not interact with other functional groups at the active centre [5]. In most biochemical systems, bases carrying out similar tasks are part of pairs or triads linked by hydrogen bridges [16]. The pK of Glu-376 at the active centre and in the complex with substrate is increased from the value of approx. 4.4 in the free state to above neutrality and to a value which is close to that of the bound substrate α-C-H. This increase has been proposed to be brought about by the hydrophobic environment at the active centre, and in particular by specific groups that flank Glu-376 [12]. Binding of acyl-CoA substrate brings the hydrophobic tail of the latter close to the Glu-COO<sup>-</sup>, and this further promotes the shift in pK<sub>a</sub> [5]. The increase in basicity of Glu-376 thus appears to result from the electrostatic effects brought about by the insertion of the negative charge in the hydrophobic environment just described.

An interaction of the utmost importance at the active centre of ADHs brings about the acidification of the substrate  $\alpha$ -C-H: the carbonyl of the acyl-thioester linkage is involved in two strong hydrogen bridges. One is with the backbone N-H of Glu-376 and the second with the 2'-OH group of the riboflavin side chain [5,17,18]. The importance of this interaction has been verified by replacing normal FAD in pig kidney MCADH with an analogue lacking the 2'-OH function (2'-deoxy-FAD). This leads to an artificial enzyme, 2'-deoxy-FAD-MCADH, which is essentially devoid of activity (approx.  $10^6$ -fold decrease in the rate of flavin reduction) [12]. Using substrate analogues in which the structure around the thioester carbonyl group was chemically modified, Colin Thorpe's group has reached similar conclusions [18].

A topic that is of similar importance deals with the molecular basis for the lack of reactivity of these dehydrogenases with dioxygen compared to that of the acyl-CoA oxidases (AOXs). It is likely that in both subfamilies the chemistry of the dehydrogenation half-reaction is closely related, and in fact a Glu functioning in  $H^+$  abstraction is found at position 376 or 255 in the dehydrogenases, and always at position 376 in the AOXs. The alignment shown in Figure 1 indicates that substantial differences between the two families exist, which could well be related to the different function in the oxidative half-reaction. Notably, the active site environment of the oxidases appears to be more polar. The topic of differentiation of oxygen reactivity is currently under investigation in several laboratories, however a molecular interpretation has not yet emerged.

### Interchange of active centre functional groups between the 'Glu-376' and the 'Glu-255' families

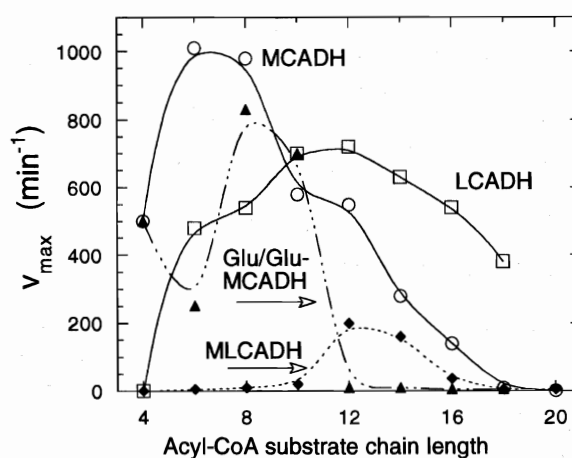
Pertinent to the questions mentioned in the introduction is a comparison at the molecular level of MCADH and LCADH which have the catalytic,  $H^+$ -abstracting Glus at two different positions (Figure 1) and differ substantially in their substrate chain length specificity. In order to study the effect of the Glu position on specificity and to verify the assumption that the position itself does not alter the chemical reactivity we have produced the following mutants of human MCADH: (a) Glu/Glu-MCADH (Thr-255Glu-MCADH). Here two Glu residues could theoretically compete for the chemical event of

$H^+$  abstraction. It is the phylogenetic progenitor to be assumed for enzymes of this family having Glu-255. (b) Glu-376Xaa. Several functional groups (Xaa) have been introduced in order to ascertain the chemical role of Glu-376 [15] (V. Kieweg, P. Bross, A. Nandy and S. Ghisla, unpublished work). Glu-376Gly mimics the amino acid sequence found in LCADH in the same domain and serves as a control. (c) Glu-376Gly/Thr-255Glu (MLCADH). Here the active site orientation of functional groups of LCADH is implemented in MCADH [14]. (d) Glu-376Thr/Thr-255Glu. In this configuration the positions of Glu and Thr have been interchanged. Of these mutants, Glu-376-Xaa-MCADHs have no activity when Xaa is not a base such as Asp, His or Cys. The properties of these mutants will be reported elsewhere (V. Kieweg, P. Bross, A. Nandy and S. Ghisla, unpublished work). The activity of mutant (d) is of the order of a few per cent of that of wtMCADH, suggesting that a Thr at position 376 is detrimental for activity in MCADH. Most interesting are the results observed with Glu/Glu-MCADH and MLCADH, which could be expressed and purified in good yield [14]. Figure 2 depicts their activity versus chain length profiles in comparison with those of wtMCADH and wtLCADH. The main point emerging from this comparison (Figure 2) is that both mutants are indeed highly active, the maximal values being well within the ranges reported for various ADHs. However the maxima of the activity profiles of the mutants are

Figure 2

Dependence of the activity of acyl-CoA dehydrogenases on the substrate chain length

The activities were determined using the ferricinium assay described by Lehman et al. [19].



108

substantially modified: they are much narrower, little activity being observed on the 'flanks' of the profiles. Furthermore, the activity maxima are shifted towards substrates with longer acyl chains. In particular, the maximum of MLCADH now coincides with that of LCADH, with which it shares the active site orientation of functional groups. From these results it can be concluded that the chemistry of dehydrogenation is not affected by the position of the Glu base, but that the latter has a strong influence on the chain length specificity. A molecular interpretation of these observations will be reported elsewhere (H. K. Lee, R. Paschke, M. Wang, A. Nandy, S. Ghisla and J. J. P. Kim, unpublished work).

### Evolution of substrate specificity

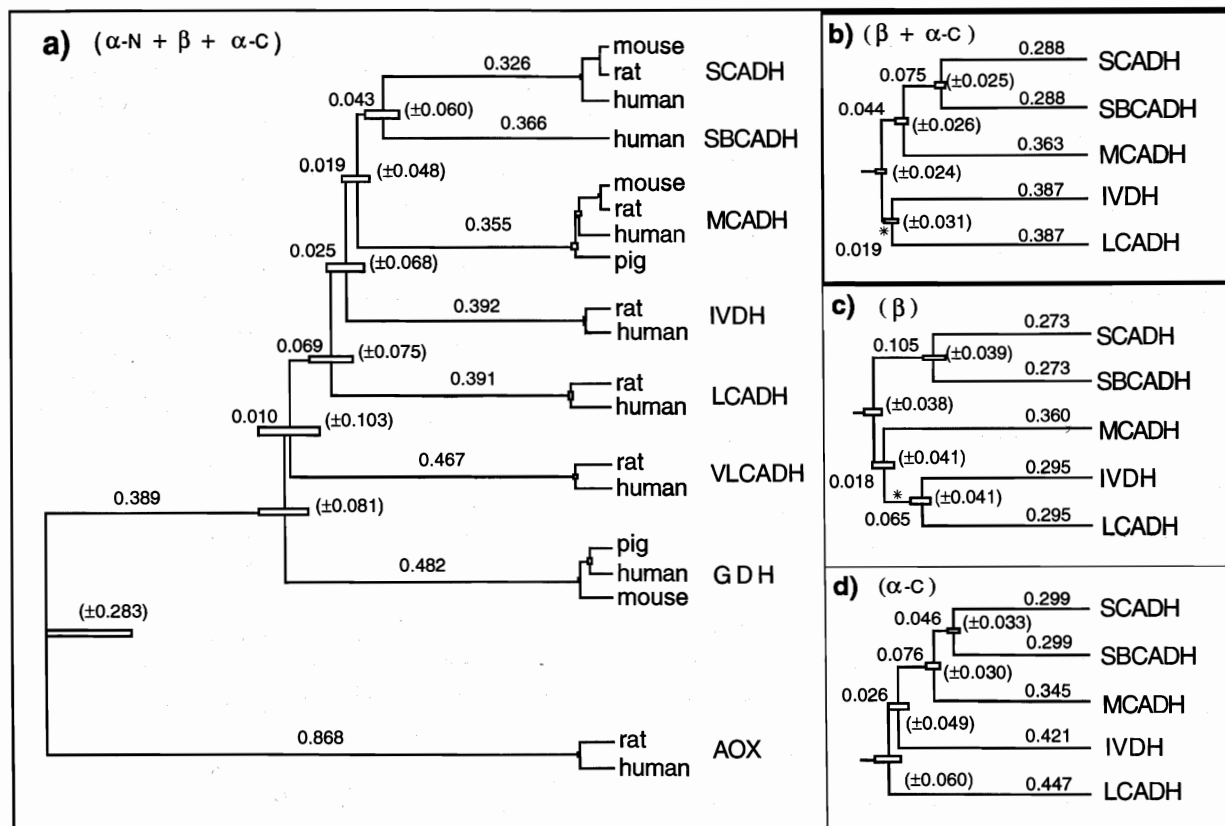
The three-dimensional folding of MCADH appears to be unique in that no enzymes other than ADHs have been found to share its topology [3]. On the other hand, the homologies within the family and its subclasses are substantial, as is evident from the motifs depicted in Figure 1. It is thus reasonable to assume that the family has emerged from a common progenitor [13]. The homologies and divergences of primary sequences of our ADHs were discussed in 1990 by Tanaka et al. [1], and a phylogenetic tree for the ADHs and AOX superfamily was presented in 1992 [2]. Since the last report, the number of sequences of ADHs/AOXs or of enzymes assumed to have similar catalytic properties has more than doubled (see Figure 1). In view of this new information and the forthcoming knowledge on the biochemistry of ADHs and AOXs, we have re-evaluated the sequence information using the program GeneWorks (version 2.4, IntelliGenetics, Mountain View, CA, U.S.A.) for the construction of UPGMA (unweighted pair group method with arithmetic mean) phylogenetic trees [19]. Initially a sequence alignment of the members of the ADH family with final manual adjustments was carried out using the sequences corresponding to the MCADH domains from helix A to helix K. The highly variable N- and C-termini, as well as the extra sequences of VLCADH, were not included in the evaluation. The AOX domains corresponding to the sequences from helix A to helix K were then added while the dehydrogenases alignment was kept constant with exception of the inclusion of gaps; the latter were kept consistent within the family. The dehydrogenases tree thus remains unaltered by the addition of the AOX sequences.

The phylogenetic tree of all known sequences of *mammalian* ADHs and AOXs is shown in Figure 3(a). This tree resembles that of Tanaka and Indo [2], but our conclusions differ in some important aspects. The tree shows that there is a distinct difference between the evolution of the gene family as such and that of the mammalian species. Inclusion of sequences from lower organisms was found to lead to overlapping of gene family and species evolution and therefore to interfere with the derivation of a consistent gene family tree. An important point that is also consistent with the deductions of Tanaka and Indo's group [2] is that the subfamily of the AOXs is well separated from the remaining members (Figure 3a). On the other hand, the branching points within the dehydrogenase family are uncertain as deduced from the overlap of the error bars. This might be because of the relatively low overall homology within the gene family (15–40% identities). In order to improve the phylogenetic tree, we have analysed the homologies within specific domains based on the three-dimensional structure of MCADH. The three domains we have analysed separately are the  $\alpha$ -helical domain at the N-terminus ( $\alpha$ -N), the domain comprising the  $\beta$ -sheet ( $\beta$ ) and the C-terminal  $\alpha$ -helical domain ( $\alpha$ -C). This subdivision corresponds to that proposed by Kim and Wu [3]. In the  $\alpha$ -N domain the ADHs homologies are very low (18–32%), whereas in the  $\beta$ -sheet domain they are relatively very high (30–51%) and in the  $\alpha$ -C we found an intermediate degree of homology (21–47%), as exemplified by the sequences of the G, J and K helices shown in Figure 1. Figure 3(b) depicts the tree derived from the alignment of the  $\beta$  and  $\alpha$ -C domains of human ADHs only. The main point emerging from it is that two main subfamilies emerge, which comprise IVDH and LCADH, which share Glu-255 on the one branch and MCADH, SBCADH and SCADH on the other. It should be noted that the overlap of the error bars does not exclude ambiguity. A tree derived from alignment of the  $\beta$ -domain sequences (Figure 3c) indicates a clear-cut separation of LCADH and IVDH from the remaining members. From this tree the assignment of MCADH is ambiguous. In the alignment of Figure 3(d), when the  $\alpha$ -C domain was used, the relatedness of MCADH to the subfamily of SCADH and SBCADH becomes more apparent, while the assignment of the branching points of LCADH and IVDH is uncertain, as can be

Figure 3

## Phylogenetic trees of the acyl-CoA dehydrogenase and oxidase families

The trees and branching points were obtained as outlined in the text using the program GeneWorks v2.4. The length of the horizontal lines of the branches is proportional to the genetic distance of the connected sequences, and is given in distance scores [20]. The standard error of each branching point is represented by the horizontal bars and is given in brackets. When overlapping of error bars of neighbouring branching points occurs, determination of branching sequence is not possible. The trees in (a), (b), (c) and (d) were obtained using the domains  $\alpha$ -N +  $\beta$  +  $\alpha$ -C,  $\beta$  +  $\alpha$ -C,  $\beta$  and  $\alpha$ -C respectively. Note that the trees in (b), (c) and (d) were obtained using the human sequences only.



judged from the overlap of the error bars. The conclusions from the single trees of Figures 3(c) and 3(d), which are derived from  $\beta$  and  $\alpha$ -C domains, thus support those derived using the unseparated domains (Figure 3b). From trees analogous to those of Figures 3(b), 3(c) and 3(d) (not shown) a clear separation of GDH and VLCADH from the other members of the dehydrogenase family also results. We thus consider the tree represented by Figure 3(b) as the one that is most likely to best represent the phylogeny of the ADH family. The main difference from the proposal of Tanaka and Indo [2] resides in the branching pattern of LCADH and IVDH.

The common ancestor protein of LCADH and isovaleryl-CoA dehydrogenases (as well as of other members of the family containing Glu-255,

i.e. the branching point between the two sub-families) should have two glutamic acids, one each at positions 255 and 376 (see branching point denoted by an asterisk in Figures 3(b) and 3(c), and should be active. With respect to these two functions, this would thus correspond to the configuration of Glu/Glu-MCADH. Indeed, this mutant is as active as MCADH or LCADH, and the activity versus chain length maximum is shifted towards that of LCADH, as shown in Figure 2. Occurrence of this mutant might have been the first step in the development of LCADH and IVDH, subsequently complementary mutations would have modified the profile of the chain length specificity, bringing about a broader chain length spectrum in the case of LCADH and the specialized function of IVDH.

Thus, different patterns of mutagenesis could surely lead to similar effects as in the cases of VLCADH and GDH (Figure 3).

### Conclusions

From the data discussed in part in this contribution, it is evident that the Glu- base that initiates  $\alpha,\beta$ -dehydrogenation can be located at different positions at the active centre of catalytically active acyl-CoA dehydrogenases. The position of this Glu, however, is an important factor in determining substrate specificity. From alignments of ADHs and AOXs sequences, a phylogenetic tree can be derived, the ancestor of which has a Glu residue at position 376. Evolution towards the LCADH and IVDH subfamily suggests the occurrence of an (extinct) ancestor which has Glu residues at both positions 376 and 255.

This work was supported by a grant from the Deutsche Forschungsgemeinschaft (Gh 2/4-7, Gh 2/6-1) to S.G. and by the EU through the Human Capital and Mobility Programme (FLAPS network contract no. CHRX-CT93-0166).

- 1 Tanaka, K., Matsubara, Y., Indo, Y., Naito, E., Kraus, J. and Ozasa, H. (1990) in *Progress in Clinical and Biological Research*, vol. 321, *Fatty Acid Oxidation, Clinical, Biochemical, and Molecular Aspects* (Tanaka, K. and Coates, P. M., eds), pp. 577-598, Alan R. Liss, New York
- 2 Tanaka, K. and Indo, Y. (1992) in *Progress in Clinical and Biological Research*, vol. 275, *New Developments in Fatty Acid Oxidation* (Coates, P. M. and Tanaka, K., eds), pp. 95-110, John Wiley, New York
- 3 Kim, J. J. and Wu, J. (1988) *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6677-6681
- 4 Kim, J. J., Wang, M., Djordjevic, S. and Paschke, R. (1992) in *Progress in Clinical and Biological Research*, vol. 375, *New Developments in Fatty Acid Oxidation* (Coates, P. M. and Tanaka, K., eds), pp. 111-126, John Wiley, New York
- 5 Kim, J. J. P., Wang, M. and Paschke, R. (1993) *Proc. Natl. Acad. Sci. U.S.A.* **90**, 7523-7527
- 6 Djordjevic, S., Stankovich, M. T. and Kim, J. J. (1995) *Biochemistry* **34**, 2163-2171
- 7 Djordjevic, S., Dong, Y., Paschke, R., Frerman, F. E., Strauss, A. W. and Kim, J. J. (1994) *Biochemistry* **33**, 4258-4264
- 8 MIPS (1995) PIR Data bank of the Martinsrieder Institut für Proteinsequenzen (MIPS), Martinsried, Germany
- 9 Ghisla, S., Thorpe, C. and Massey, V. (1984) *Biochemistry* **23**, 3154-3161
- 10 Ghisla, S. and Massey, V. (1989) *Eur. J. Biochem.* **181**, 1-17
- 11 Thorpe, C. and Kim, J. J. (1995) *FASEB J.* **9**, 718-725
- 12 Ghisla, S., Engst, S., Vock, P., Kieweg, V., Bross, P., Nandy, A., Rasched, I. and Strauss, A. W. (1992) in *Progress in Clinical and Biological Research*, vol. 375, *New Developments in Fatty Acid Oxidation* (Coates, P. M. and Tanaka, K., eds), pp. 283-292, John Wiley, New York
- 13 Matsubara, Y., Indo, Y., Naito, E., Ozasa, H., Glassberg, R., Vockley, J., Ikeda, Y., Kraus, J. and Tanaka, K. (1989) *J. Biol. Chem.* **264**, 16321-16331
- 14 Nandy, A., Bross, P., Kräutle, F.-G., Rasched, I. and Ghisla, S. (1994) in *Flavins and Flavoproteins: Proceedings of the Eleventh International Symposium, Nagoya (Japan)* (Yagi, K., ed.), pp. 327-330, Walter de Gruyter, Berlin, New York
- 15 Bross, P., Engst, S., Strauss, A. W., Kelly, D. P., Rasched, I. and Ghisla, S. (1990) *J. Biol. Chem.* **265**, 7116-7119
- 16 Voet, D. and Voet, J. G. (1990) in *Biochemistry*, pp. 355-382, John Wiley, New York
- 17 Ghisla, S., Melde, K., Zeller, H. D. and Boschert, W. (1990) in *Progress in Clinical and Biological Research*, vol. 321, *Fatty Acid Oxidation, Clinical, Biochemical, and Molecular Aspects* (Tanaka, K. and Coates, P. M., eds.), pp. 185-192, Alan R. Liss, New York
- 18 Triebel, R. C., Wang, R., Anderson, V. E. and Thorpe, C. (1995) *Biochemistry* **34**, 8597-8605
- 19 Lehman, T. C., Hale, D. E., Bhala, A. and Thorpe, C. (1990) *Analytical Biochemistry* **186**, 280-284
- 20 Nei, M. (1987) in *Molecular Evolutionary Genetics*, pp. 293-298 Columbia University Press, New York

Received 23 August 1995