Alcohol & Alcoholism, Suppl. 1, pp. 13–23

ALCOHOL AND ALDEHYDE DEHYDROGENASES: STRUCTURES OF THE HUMAN LIVER ENZYMES, FUNCTIONAL PROPERTIES AND EVOLUTIONARY ASPECTS

Hans Jörnvall, John Hempel, Hedvig von Bahr-Lindström, Jan-Olov Höög and Bert L. Vallee

Department of Chemistry I, Karolinska Institutet, S-104 01 Stockholm, Sweden, and Center for Biochemical and Biophysical Sciences and Medicine, Harvard University, Boston, MA 02115, USA

ABSTRACT

All three types of subunit of class I human alcohol dehydrogenase have been analyzed both at the protein and cDNA levels, and the structures of α, β_1, β_2, γ_1 and γ_2 subunits are known. The same applies to class II π subunits. Extensive protein data are also available for class III χ subunits.

In the class I human isozymes, amino acid exchanges occur at 35 positions in total, with 21-28 replacements between any pair of the $\alpha/\beta/\gamma$ chains. These values, compared with those from species differences between the corresponding human and horse enzymes, suggest that isozyme developments in the class I enzyme resulted from separate gene duplications after the divergence of the human and equine evolutionary lines. All subunits exhibit some unique properties, with slightly closer similarity between the human γ and horse enzyme subunits and somewhat greater deviations towards the human α subunit. Differences are large also in segments close to the active site zinc ligands and other functionally important positions. Species differences are distributed roughly equally between the two types of domain in the subunit, whereas isozyme differences are considerably more common in the catalytic than in the coenzyme-binding domain. These facts illustrate a functional divergence among the isozymes but otherwise similar changes during evolution. Polymorphic forms of β and γ subunits are characterized by single replacements at one and two positions, respectively, explaining known deviating properties.

Class II and class III subunits are considerably more divergent. Their homology with class I isozymes exhibits only 60-65% positional identity. Hence, they reflect further steps towards the development of new enzymes, with variations well above the horse/human species levels, in contrast to the class I forms. Again, functionally important residues are affected, and patterns resembling those previously established for the divergently related polyol dehydrogenases are encountered.

The two isozymes of human aldehyde dehydrogenase also exhibit considerable differences, with only 68% structural identity. The results show an early divergence into isozymes before the man/horse species radiation. Cys-302 is a functionally important residue and is located in one of the regions with conserved hydrophobic properties. Other regions with large differences in hydropathic properties may explain the absence of cross-hybridizing isozyme forms of human liver aldehyde dehydrogenase.

KEYWORDS

Alcohol dehydrogenase: aldehyde dehydrogenase: isozymes: enzyme evolution: gene duplications: structure-function relationships.

INTRODUCTION

Alcohol dehydrogenase

This enzyme from yeast was the first pyridine-nucleotide dehydrogenase crystallized (Negelein and Wulff, 1937). Structural data for the horse enzyme, with 374-residue subunits, were available early (cf. Brändén et al., 1975), and this form has served as a model for the enzyme from other species (Jörnvall et al., 1978; Brändén et al., 1984), and even for other enzymes which are distantly related (Jörnvall et al., 1981; Eklund et al., 1985). However, until recently comparatively little was known about the structure of human alcohol dehydrogenase, largely owing

to the presence of many different forms of the enzyme in man.

Three types of subunit were distinguished early (Smith et al., 1971). Introduction of an inhibitor-derived ligand for affinity chromatography (Lange and Vallee, 1976), allowed separation of further forms leading to the recognition of three classes, I, II, and III, of the enzyme (Vallee and Bazzone, 1983). Class I represents the initially discovered three types of subunit, α, β, and γ, with their allelic variants, including the type originally distinguished as "atypical" (von Wartburg et al., 1965) and now referred to as β_2 (Smith et al., 1971; Jörnvall et al., 1984a). Recent work correlating protein and cDNA studies have established the structures of α, β_1, β_2, γ_1, and γ_2 subunits (Jörnvall et al., 1984a; Hempel et al., 1985a; von Bahr-Lindström et al., 1986; Hedén et al., 1986; Höög et al., 1986a) in agreement with studies from other laboratories (Duester et al., 1984; Ikuta et al., 1985, 1986). Work on the cDNA for the enzyme from other mammalian species has also been reported (Edenberg et al., 1985). In addition, the genomic structure of a human class I enzyme subunit is known (Duester et al., 1986b), allowing further correlations with organizational patterns (Duester et al., 1986a), including exon/intron relationships for human alcohol dehydrogenase, in the same way as deduced previously for maize (Dennis et al., 1984, 1985; Brändén et al., 1984) and Drosophila (Benyajati et al., 1981) alcohol dehydrogenases. The three genes for the yeast enzyme have also been analyzed (cf. Young and Pilgrim, 1985).

Finally, two types of alcohol dehydrogenase in nature have been distinguished, as well as family assignments of several enzymes of both types (Jörnvall et al., 1981, 1984b). The types differ in subunit size and properties of catalysis. Thus, mammalian, yeast and most other characterized alcohol dehydrogenases are "long-chain" alcohol dehydrogenases with Zn at the active site, whereas the characterized insect enzyme (from Drosophila) is a "short-chain" alcohol dehydrogenase, lacking the typical Zn-liganding structures and apparently also the metal. Bacterial ribitol dehydrogenase and glucose dehydrogenase are distantly related to the short-chain alcohol dehydrogenase type, and mammalian sorbitol dehydrogenase to the long-chain type. The two types indicate the presence of ancient building unit(s), and reflect both convergent and divergent evolution (convergent in function on alcohol dehydrogenase activity, cf. Jörnvall, 1986, but divergent in structure from largely different organizations in the short-chain and long-chain enzymes).

This review will cover human liver alcohol dehydrogenase subunits of class I and will also touch upon the most recent progress towards the characterizations of class II subunits of the π type (Höög, et al., 1986b; and unpublished) and class III subunits of the χ type (unpublished).

Aldehyde dehydrogenase

These enzymes are larger proteins, with 500-residue subunits (Hempel et al., 1984b; 1985b). The enzyme occurs in two types, which differ in intracellular localization, and which were early purified from several species (Eckfeldt et al., 1976; Greenfield and Pietruszko, 1977). However, as for alcohol dehydrogenases, the structures of the human forms have become available only recently at the protein (Hempel et al., 1984b, 1985b) and cDNA levels (Hsu et al., 1985). The mitochondrial and cytosolic forms differ considerably in structure (Hempel et al., 1985b), explaining the lack of cross-hybridization. These properties suggest an emerging development of partially different enzymes (Jörnvall, 1980; Jörnvall et al., 1986). As with human alcohol dehydrogenase, an atypical form of the mitochondrial aldehyde dehydrogenase has been investigated and defined as a single amino acid replacement (Hsu et al., 1985; Yoshida et al., 1984; Hempel et al., 1984a). Functional correlations, utilizing thiol reagents (Hempel et al., 1982) and different reactive coenzyme analogs (von Bahr-Lindström et al., 1985) have further identified Cys-302 as an important residue close to the active site of the enzyme. Additional correlations are in progress. Most importantly, crystals of the mitochondrial enzyme have recently become available for analysis (unpublished), giving promise that detailed knowledge will soon be available for aldehyde dehydrogenase in the same way as that already available for alcohol dehydrogenase. This review will cover the structural properties of the two characterized forms of human liver aldehyde dehydrogenase.

MATERIALS AND METHODS

Details of enzyme purification, protein analysis, cDNA analysis and further correlations followed methods previously described. For studies from the authors laboratories, human liver alcohol dehydrogenase and human liver aldehyde dehydrogenase were purified utilizing affinity chromatography steps on CapGapp Sepharose (Lange and Vallee, 1976) and AMP Sepharose (Greenfield and Pietruszko, 1977), respectively. Pure enzymes were ^{14}C-carboxymethylated, and cleaved with CNBr and proteolytic enzymes; peptides were analyzed by reverse phase high performance liquid chromatography (von Bahr-Lindström et al., 1986; Hempel et al., 1984b) and characterized with different methods, utilizing degradations with a manual method (von Bahr-Lindström et al., 1982), with liquid phase sequencers, modified liquid phase sequencers (Carlquist et al., 1984) and gas phase sequencers. Partially known structures allowed screening of human liver cDNA libraries

utilizing oligonucleotide probes, cloning of hybridization-positive inserts, and subsequent sequence analysis with the dideoxy chain termination method (von Bahr-Lindström et al., 1986; Hedén et al., 1986; Höög et al., 1986a,b). Structural correlations with alcohol dehydrogenases utilized the originally known horse liver enzyme tertiary structure (Eklund et al., 1976), with subsequent refinements including details for the ternary complex with the coenzyme (Eklund et al., 1981), and correlations with other indirectly deduced structural variants, also derived from comparisons (Jörnvall et al., 1978; Eklund et al., 1985, 1986).

RESULTS AND DISCUSSION

I. ALCOHOL DEHYDROGENASE

Structures of human class I isozymes and cDNAs

All structures reported for the class I human isozymes are summarized in Fig. 1. Values for the differences among the three types of subunit are given in Table 1. These overall comparisons already reveal two types of imbalance in the distributions of differences for the subunits and their genes. Thus, amino acid substitutions in isozymes are considerably more frequent in the catalytic domain of the mature protein than in the coenzyme-binding domain (Table 1A, top row, and more clearly in Fig. 2). This fact is also visible when all the human/horse species differences are considered (Table 1A, middle row) and indicates functional differences that can be interpreted to account for metabolic roles and functional properties of the different isozymes (below). However, when species differences outside positions with isozyme differences are considered (Table 1A, bottom row), the two domains reveal fairly equal values. This shows that in the overall rate of evolutionary changes, disregarding the special effects from different functional roles accommodated by the isozyme developments, the two domains behave equally during

```
                    L       E       H       V   GT D     TM
  1 STAGKVIKCK AAVLWEVKKP FSIEDVEVAP PKAYEVRIKM VAVGICRTDD HVVSGNLVTP LPVILGHEAA GIVESVGEGV   80
                    L       E       H       A   RS E     NL

                 AI      I           VS  Q       S     RR      L I
 81 TTVKPGDKVI PLFTPQCGKC RVCKNPESNY CLKNDLGNPR GTLQDGTRRF TCRGKPIHHF LGTSTFSQYT VVDENAVAKI  161
                 FT      I           LG  R       R     SG      V V

                         N                       AI
161 DAASPLEKVC LIGCGFSTGY GSAVNVAKVT PGSTCAVFGL GGVGLSAVMG CKAAGAARII AVDINKDKFA KAKELGATEC  240
                         K                       VV

                                 R                       D     M             IL
241 INPQDYKKPI QEVLKEMTDG GVDFSFEVIG RLDTMMASLL CCHEACGTSV IVGVPPASQN LSINPMLLLT GRTWKGAVYG  320
                                 R/Q                     D     I             IF

          CV                    HV              H       I M
321 GFKSKEGIPK LVADFMAKKF SLDALITHVL PFEKINEGFD LLHSGKSIRT VLTF                            374
          SV                    N/V             R       V T
```

Fig. 1. Structures of all types of subunit (α, β, γ), including major allelic variations (β_1/β_2 and γ_1/γ_2), for human liver alcohol dehydrogenase.

The continuous line is represented by the β_1 isozyme. At positions where any class I isozyme differs, residues in all isozymes are given, with the α alternative above, and the γ_1 alternative below the β alternative. At positions where allelic variants differ, the β_1/β_2 alternatives are both given (as R/H at position 47), as are also the γ_1/γ_2 alternatives (R/Q at 271 and I/V at 349). Data for π subunits of class II and χ subunits of class III are not yet published, but the zinc ligands (at positions 46, 67, 174 for the active site Zn, and 97, 100, 103, 111 for the second Zn) and some other important residues (including the coenzyme-interacting positions 199, 223, 228) known to be strictly conserved in all human liver alcohol dehydrogenase classes are underlined. Residues are printed in groups of 10, with start and end numbers given for each line. Data from Jörnvall et al., 1984a; Hempel et al., 1985a; Hedén et al., 1986; von Bahr-Lindström et al., 1986; and Höög et al., 1986a).

the class I divergence exactly as previously noticed from the differences between the yeast and horse alcohol dehydrogenases (Jörnvall, 1977). A second overall imbalance concerns the frequency of nucleotide exchanges between the coding and non-coding segments, as has been noted before (Höög, et al., 1986a) and is common in other systems, too. Thus, the 5' non-coding flanking region contains a higher frequency of nucleotide exchanges than remaining parts, including the 3' non-coding segment, and has exchanges about twice as frequent as those constituting the sum of the functional and silent mutations in the coding part (Table 1B). This can also be interpreted functionally and may reflect different control elements in the regulation of gene activities.

Functional properties

Zinc ligands. The three ligands to the active site zinc atom (Cys-46, His-67, and Cys-174) are conserved in class I, as well as in class II and class III alcohol dehydrogenases. Similarly, the four ligands to the other zinc atom (Cys-97, Cys-100, Cys-103 and Cys-111) are conserved in all three classes. This indicates overall common basic properties for all the enzymes, as detailed (Eklund et al., 1986) and recently reviewed for the class I isozymes (Jörnvall et al., 1986). Consequently, models of these human enzymes can be constructed based on the known conformation of

Table 1. Summary of overall differences within the class I proteins (A) and cDNAs (B).

In A, the values listed are those for positions with replacements in any of the $\alpha/\beta_1/\beta_2/\gamma_1/\gamma_2$ chains. Considering only pair-wise comparisons, all values become smaller and slightly variable depending on which two structures that are compared (the total of 35 then corresponds to 21-28, cf. Fig. 4). The same applies to the species variants where values given list positions with a difference between the horse E-type subunit and human class I subunit (the total of 64 corresponds in pair-wise comparisons to 40-44, cf. Fig. 4). In B, nucleotide positions with exchanges are given in per cent of positions compared in order to reveal frequency differences that are not directly visible in absolute numbers because of the large size differences between the segments listed. As in A, values given are those considering replacements in any of the positions that can be compared (total 1450) in the cDNAs corresponding to $\alpha/\beta_1/\gamma_2$, while pair-wise comparisons (Höög et al., 1986a) give slightly different values. Two types of imbalance are noticed. One concerns the distribution in A of domain differences between isozyme variants (top row) and species variants (bottom row) (sum of both differences in middle row); the other concerns the distribution in B between coding and non-coding segments. Data from Hedén et al., 1986, von Bahr-Lindström et al., 1986, Höög et al., 1986a, and Jörnvall et al., 1986.

A.

Property	Amino acid positions with differences		
	Total	Catalytic domain	Coenzyme-binding domain
Isozyme variants (human $\alpha/\beta/\gamma$)	35	28	7
Species variants (Horse - Human class I)			
all positions	64	42	22
excluding positions with human isozyme variants	29	14	15

B.

Property	Nucleotide positions with differences (%)					
	Total	Non-coding		Coding		
		5'	3'	Functional	Silent	Total
Class I cDNAs	9.1	18.7	13.0	3.5	4.1	7.6

the horse enzyme. A schematic representation in such a model of all amino acid replacements within the human class I isozymes is given in Fig. 2. Generally, superficial positions of the replacements, as well as replacements close to the active site are directly noticable in Fig. 2. Two of the most variable segments within the class I isozymes are the regions that affect the two zinc atoms. Thus, Cys-46 is within a non-conserved segment containing several functionally important replacements (below). The other region variable in amino acid sequence (Fig. 1) is the one encompassing positions 93-143, thus including the entire loop around the second zinc atom, and also contributing to the substrate cleft space relationships (especially for position 93; Eklund et al., 1986). Differences in catalytic details within the generally conserved nature of the class I isozymes are therefore to be expected. These conclusions about differences in enzymatic properties apply still more to the class II and class III alcohol dehydrogenases as indicated by the presence of still further structural differences.

Functionally important residues. Several of the residues to which special functional roles have been ascribed in alcohol dehydrogenase are among those which are exchanged within the class I isozyme variations (Fig. 1). In particular, the residue at position 47 interacts with the coenzyme pyrophosphate (Brändén et al., 1975; Eklund et al., 1981, 1986) but no less than three different residues can here replace each other (Fig. 1). The residue at this position affects coenzyme dissociation, and therefore the rate-limiting step in the enzymatic reaction, with consequences on the overall activity of the enzyme (Jörnvall et al., 1984a; Eklund et al., 1986). The His/Arg substitution at this position (Fig. 1) has been coupled directly with the presence of the atypical enzyme protein. The Gly/basic residue exchange between α and the other isozymes (Fig. 1) should increase coenzyme dissociation still further. However, other exchanges, in particular the one at position 48, may compensate the effect and maintain an efficient enzyme (Eklund et al., 1986). Thus, a position of central importance is the Thr/Ser exchange at position 48. Furthermore, this difference may restrict the space available in the substrate pocket close

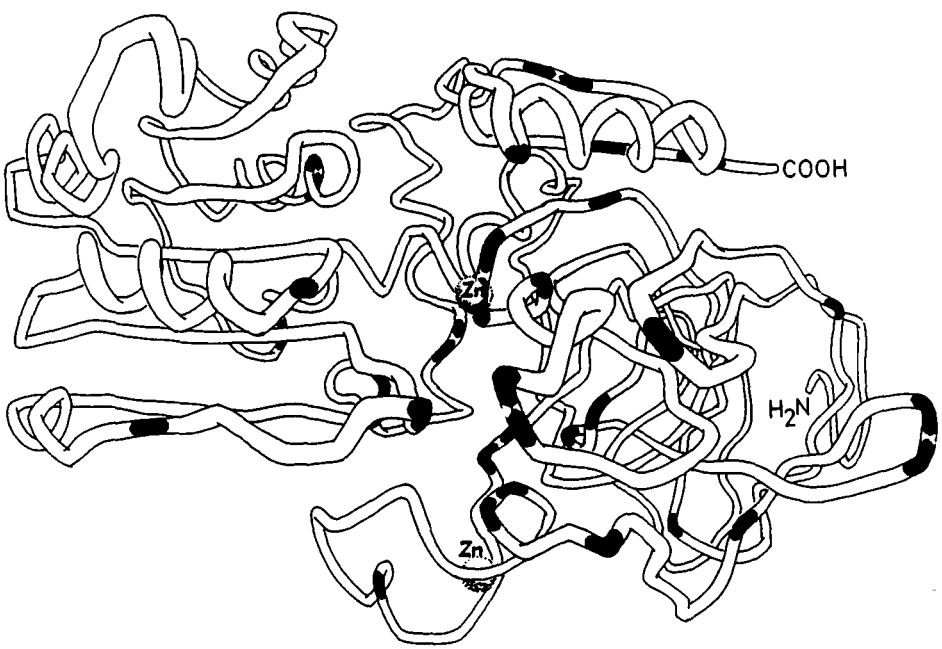


Fig. 2. Schematic representation of all positions with any residue differences in the three types of class I subunit of human alcohol dehydrogenase.

The conformation given (original drawing by Furugren) is that of the crystallographically analyzed horse E type enzyme subunit (cf. Brändén et al., 1979). The human class I subunits have related conformations (Eklund et al., 1986) and all positions with α/β/γ isozyme differences (cf. Fig. 1) are indicated in black. It is clearly visible that the catalytic domain (the right-hand half of the figure) contains most of the positions with isozyme differences (cf. Table I) and that many replacements are superficially located.

to the active site zinc atom. However, the conserved hydroxyl keeps a hydrogen bond with the NMN ribose of NAD intact.

The changes at position 47 and 48 have been interpreted to be those which are most significant in the explanations of the altered properties of the class I isozymes. In addition, an exchange at positions 271 (a γ_1/γ_2-difference) influences NAD-binding, accounting for the slower overall catalytic rate of γ_2. Similarly, an exchange at position 93, influencing the size of the substrate-binding pocket, and to a lesser extent possibly also those at positions 57 and 116, influencing the substrate cleft, are interpreted to explain the somewhat altered properties of the class I isozymes (Eklund et al., 1986). Yet, different properties in steroid binding and allosteric relationships for these isozymes (Mårdh et al., 1986) have not been accounted for in terms of structure. However, the large functional differences between the allelic variants β_1/β_2 and γ_1/γ_2 are explained entirely by the structural relationships characterized (Jörnvall et al., 1984a; Höög et al., 1986a; Eklund et al., 1986). For classes II and III differences are still more extensive.

Coenzyme-binding domain. The entire coenzyme-binding domain (positions 176-318) is extremely conserved. In the whole isozymes, 35 positions differ, but only eight of them effect this half of the molecule (Table 1). Only the exchange at position 271 has a minor influence on coenzyme-binding (above). The C-terminal part of the coenzyme-binding domain is the major area of subunit interactions in alcohol dehydrogenase (Brändén et al., 1975). The generally conserved properties in this domain of the class I isozymes therefore also explain why these three types of subunit all cross-hybridize into dimeric enzyme molecules, while larger differences towards the π and χ structures account for the absence of hybrids between class I and class II or class III subunits.

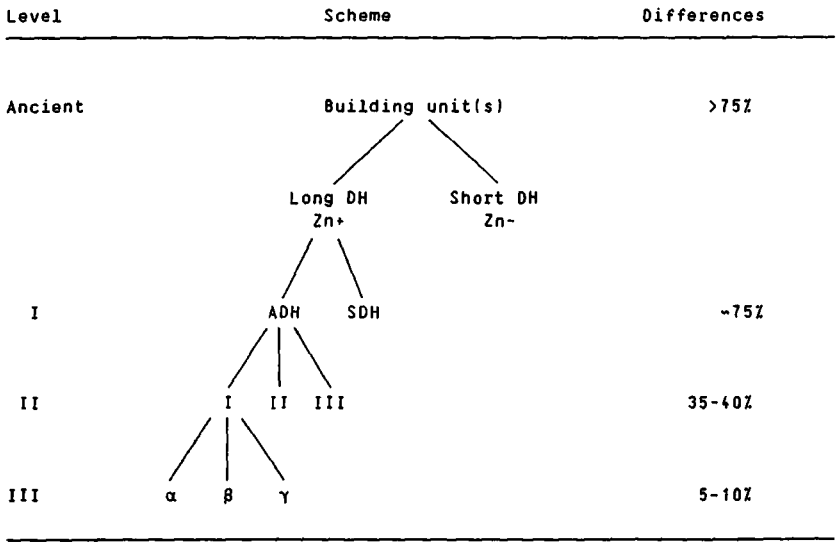


Fig. 3. Scheme showing structural relationships for zinc-containing dehydrogenases in human liver.

Three levels of entire chain duplications are discerned. At level I, different enzyme lines were created, yielding the present-day different enzymes (ADH, alcohol dehydrogenase; SDH, sorbitol dehydrogenase; DH, dehydrogenase). At level III recent, multiple, and species-specific duplications explain intra-class isozymes of minor variability. At level II, inter-class alcohol dehydrogenase lines originated, yielding present day forms of an intermediate type, approaching enzyme evolution rather than typical isozyme variability. Values show extent of residue differences between present-day forms, with ranges to include all forms and both domains in separate pair-wise comparisons. Branches and levels are intended to illustrate principles, not exact details. Thus, the three branches at each of the duplicatory levels II and III are arbitrarily represented as similar and symmetrical, although branch topologies probably differ but cannot yet be further discerned, in the absence of additional data.

Evolution of alcohol dehydrogenase

The structures of the isozymes of human alcohol dehydrogenase, both those now available (Fig. 1) and those of π and χ subunits of classes II and III (Höög et al., 1986b; and unpublished) link previous observations on alcohol and polyol dehydrogenases by showing evolution of a large enzyme complex that encompasses several zinc metalloenzymes in human liver.

As shown in Fig. 3, three different levels of duplicatory events are clearly visible along the line to present-day Zn-containing alcohol and polyol dehydrogenases. In addition, there is an early stage, possibly encompassing both duplication(s) and gene rearrangement(s) resulting in still more ancestral parent molecules but this fourth stage(s) is too distant yet to trace in detail without more information on further structures. As shown in Fig. 3, it produced two descendent lines, leading to long-chain dehydrogenases with zinc at the active site and short-chain dehydrogenases without zinc. The first of the subsequent duplicatory levels on the descendant line of the long metalloenzyme subunit type is the sorbitol dehydrogenase/alcohol dehydrogenase duplication, which has been outlined earlier (Jörnvall et al., 1981). The resultant structures are clearly different (Jörnvall et al., 1984c), deviating at about 75% of all positions (Fig. 3), and forming separate enzymes locked at different positions in a metabolic pathway (Jeffery and Jörnvall, 1983).

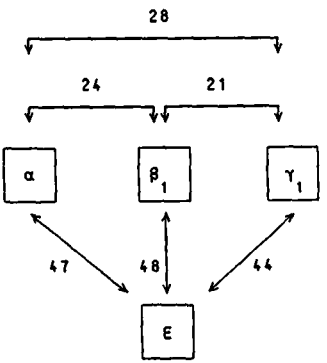
The next level of duplicatory events encompassed duplications leading to the various class descendents (Fig. 3). Here, the present-day structural identities are between those for the sorbitol/alcohol dehydrogenases of the first level and those of the class I isozymes (Fig. 3). The latter constitute the third level, which is recent, species-dependent, of repeated occurrence, and treated previously in a review (Jörnvall et al., 1986). As shown in Fig. 4, the third level of duplications has evolved considerably fewer residue differences (at 21-28 positions) than those between human and horse class I enzymes (44-48 positions). These values (Fig. 4), as well as the summaries in Table I, suggest that the third level of duplications to give the intra-class I isozymes in man can be timed to a period considerably after this species deviation. The second level of inter-class relations in the evolution of alcohol dehydrogenases has not been emphasized structurally before, but now forms a direct link between the early polyol/alcohol dehydrogenase separation and the recent isozyme separations. This level furthermore explains why the class II and class III enzymes, in contrast to the class I isozymes, appear common to several mammals (Oitlow et al., 1984; Beisswanger et al., 1985). In fact, intra-class variations at level II resemble evolution of new enzymes (as at level I), rather than of true, closely related isozymes (as at level III). Interestingly, similar different levels as now outlined for mammalian zinc-containing dehydrogenases also pertain to another complex enzyme group, that of cytosolic glutathione transferases (Persson et al., 1986). Both these completely different enzyme systems have multiple members in mammalian organs and show evolutionary stages in between recent isozyme evolution and distantly related enzymes; both also show further connections to still more separated other enzyme forms (cf. Jörnvall and Persson, 1986), demonstrating general patterns in protein evolution.

II. ALDEHYDE DEHYDROGENASE

Structures of human cytosolic and mitochondrial aldehyde dehydrogenases

The two structures are shown in Fig. 5. The enzymes differ considerably and show a structural identity at 68% of all positions. These enzymes are larger than the alcohol dehydrogenases and lack functional metal atoms. However, important cysteine residues have been distinguished, and one such reactive residue is at position 302, which is conserved in both enzyme forms (Fig. 5).

Fig. 4. Positions with residue differences in all pair-wise comparisons between the human α, β, and γ enzyme subunits and the horse E-type enzyme subunit.



In contrast to all the alcohol dehydrogenases and the cytosolic aldehyde dehydrogenase, the mitochondrial aldehyde dehydrogenase has a non-blocked N-terminus (the others being acetyl-blocked). Furthermore, this N-terminus has a "ragged end". Both the heterogeneity in the N-terminal starting position and the absence of a blocking group have been ascribed to remnants in the mature protein of the signals for mitochondrial transport of the mitochondrial form (Hempel et al., 1985b).

Functional and evolutionary interpretations

Aldehyde dehydrogenase clearly differs from alcohol dehydrogenases. Crystallographic data are not yet available, although crystals suitable for investigation have just been obtained (unpublished). Therefore, both conformational and evolutionary details for aldehyde dehydrogenase are still far less well known than for alcohol dehydrogenase.

However, one region with some large changes in hydrophobicity patterns (a segment at positions 320-395) has been suggested as one possible area for subunit interactions in aldehyde dehydrogenase (Hempel et al., 1985b), since the differences here appear large enough to explain the lack of cross-hybridization between the two types of subunit. Another area of special interest concerns the segment around position 302. Here, a reactive cysteine residue has been ascribed a position influenced by the inhibitor disulfiram (Hempel et al., 1982), and also a position close to the coenzyme-binding region (von 8ahr-Lindström et al., 1985). Furthermore, this segment is in a region of conserved hydrophobicity and might contribute some hydrophobic

```
  1 SSSGTPDLPV LLTDLKIQYT KIFINNEWHD SVSGKKFPVF NPATEEELCQ VEEGDKEDVD KAVKAARQAF QIGSPWRTMD  80
    (AAAA QAV A PNQQPEVFCN Q          A  R T  TV   S G VI    A               A   L     R
     ↑ ↑↑ ↑

 81 ASERGRLLYK LADLIERDRL LLATMESMNG GKLYSNAYLN DLAGCIKTLR YCAGWADKIQ GRTIPIDGNF FTYTRHEPIG 160
      H     NR           T Y  AL TLDN   P VIS  V   DMVL C    Y       YH  K      D   S       V

161 VCGQIIPWNF PLVMLIWKIG PALSCGNTVV VKPAEQTPLT ALHVASLIKE AGFPPGVVNI VPGYGPTAGA AISSHMDIDK 240
                  L QA  L      AT  V   M V           Y  N                F          A  E V

241 VAFTGSTEVG KLIKEAAGKS NLKRVTLELG GKSPCIVLAD ADLDNAVEFA HHGVFYHQGQ CCIAASRIFV EESIYDEFVR 320
            I  RV QV   S                    N IMS    M W   Q   FAL FN      C G  T   Q D       E

321 RSVERAKKYI LGNPLTPGVT QGPQIDKEQY DKILDLIESG KKEGAKLECG GGPWGNKGYF VQPTVFSNVT DEMRIAKEEI 400
      A   SRV V   FDSKTE      V ET F K   GY NT   Q     L     IAADR    I     GD Q  G T

401 FGPVQQIMKF KSLDDVIKRA NNTFYGLSAG VFTKDIDKAI TISSALQAGT VWVNCYGVVS AQCPFGGFKM SGNGRELGEY 480
      M  L    TIEE VG      ST   A A      L   N YL Q             D FG   S    Y     S

481 GFHEYTEVKT VTVKISQKNS                                                              500
     LQA           VP
```

Fig. 5. The primary structures of the two types of aldehyde dehydrogenase characterized from human liver.

The continuous line shows the cytosolic form. The residue alternatives in the mitochondrial enzyme are shown below at those positions only where the two enzymes differ. A residue within parentheses, and the arrows in the N-terminal region of the mitochondrial form indicate microheterogeneities in starting position (with alternative starts at the arrows), probably reflecting remnants of the signals for mitochondrial transport. Residues are printed in groups of 10, and start and end numbers are given for each line as in Fig. 1. The cysteine residue associated with special reactivity, protection by the inhibitor disulfiram, and accessibility to reactive coenzyme analogs, is underlined at position 302. Glu-487, harboring the mutated position in a largely non-functional mitochondrial enzyme variant of Oriental populations, is also underlined. Data from Hempel et al. (1982, 1984a,b, 1985b), Yoshida (1984), von 8ahr-Lindström et al. (1985), Hsu (1985).

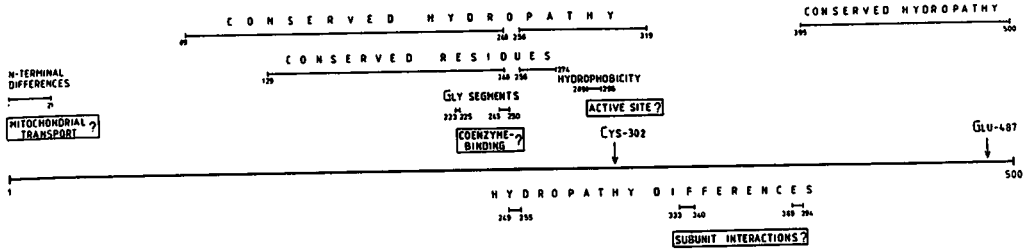


Fig. 6. Schematic representation of aldehyde dehydrogenase, and relative positions of special segments, residues or overall properties.

The interpretations (within boxes) are clearly tentative but are indicated in order to show the relationships that appear possible from presently known properties of overall hydropathy, residue differences between the mitochondrial and the cytosolic enzymes, distribution of glycine residues, and the positions of a reactive residue and a mutational difference.

properties to the environment of the active site cysteine residue(s) (Hempel et al., 1985b). Regarding coenzyme-binding in aldehyde dehydrogenase, a pattern with alternating glycine residues that is characteristic for coenzyme-binding segments of other dehydrogenases is difficult to define, but a few possible segments have been recognized (Hempel et al., 1985b). Finally, a functional mutant, explaining the presence of a largely inactive mitochondrial enzyme, has been defined in Oriental populations, explaining altered sensitivities to ethanol in affected populations; in this mutant form, Glu-487 is replaced by Lys in the C-terminal segment (Yoshida et al., 1984; Hempel et al., 1984a; Hsu et al., 1985). A summary of the overall properties discerned and their possible functional interpretations is shown in Fig. 6, as now discerned from the limited data available. It will be of interest to find out how well this model may represent true relationships that will be available in the future from crystallographic analyses.

Evolution-wise aldehyde dehydrogenase, exactly like alcohol dehydrogenases, exhibits isozyme formation. The two aldehyde dehydrogenases have diverged considerably. Functionally, they still retain similar activities, but structurally they deviate at 32% of all positions and do not form hybrid molecules. Furthermore, they have become segregated into different cellular compartments. Therefore, in the evolutionary scheme of Fig. 3, the position of the aldehyde dehydrogenases would most closely correspond to level II. Thus, aldehyde dehydrogenases, like the different classes of alcohol dehydrogenases (above) and the different classes of glutathione transferases (Persson et al., 1986) approach stages in the emergence of new enzyme functions rather than merely demonstrating recent isozyme evolution.

In a larger context, it has not yet been possible to discern details of ancestral connections between aldehyde dehydrogenases and other dehydrogenases. Thus, aldehyde dehydrogenase, although a dehydrogenase and therefore distantly related to other dehydrogenases (Rossmann et al., 1975), cannot yet be linked to the early building units of the scheme leading to the metalloenzyme dehydrogenases (Fig. 3). Obviously, further enzyme studies are required until all functional and evolutionary connections of aldehyde dehydrogenase are known. The accessibility of additional enzyme variants from other species (von Bahr-Lindström et al., 1984), reactive coenzyme analogs (von Bahr-Lindström et al., 1985), and crystals suitable for X-ray crystallography (unpublished) offer hopes that detailed knowledge on many aspects of aldehyde dehydrogenase may soon be available, exactly as has increasingly been the case for alcohol dehydrogenase. Thus, genetically, structurally, functionally and evolution-wise, both of the two first enzymes in alcohol metabolism in man are of increasing interest for general aspects in protein chemistry and enzyme structure.

ACKNOWLEDGEMENT

Support from the Swedish Medical Research Council (projects 03X-3532, 13X-7148), Knut and Alice Wallenberg Foundation, Magn. Bergvall Foundation, KabiGen AB, the Swedish Committee for Coordination and Planning in Research, and the Endowment for Research in Human Biology (Boston, MA) through funds from the Samuel Bronfman Foundation made available by Joseph E. Seagram and Sons, is gratefully acknowledged.

REFERENCES

Beisswanger, T.B., Holmquist, B. and Vallee, B.L. (1985) Proc. Natl. Acad. Sci. USA 82, 8369-8373.
Benyajati, C., Place, A.R., Powers, D.A. and Sofer, W. (1981). Proc. Natl. Acad. Sci. USA 78, 2717-2721.
Brändén, C.-I., Jörnvall, H., Eklund, H. and Furugren, B. (1975). In: The Enzymes: (ed. P.D. Boyer) 3rd ed., Vol. 11, Academic Press, New York. 103-190.
Brändén, C.-I., Eklund, H., Cambillau, C. and Pryor, A.J. (1984). EMBO J. 3, 1307-1310.
Carlquist, M., Kaiser, R., Tatemoto, K., Jörnvall, H. and Mutt, V. (1984). Eur. J. Biochem. 144, 243-247.
Dennis, E.S., Gerlach, W.L., Pryor, A.J., Bennetzen, J.L., Inglis, A., Llewellyn, D., Sachs, M.M., Ferl, R.J. and Peacock, W.J. (1984). Nucleic Acids Res. 12, 3983-4000.
Dennis, E.S., Sachs, M.M., Gerlach, W.L., Finnegan, E.J. and Peacock, W.J. (1985) Nucleic Acids Res. 13, 727-743.
Ditlow, C.C., Holmquist, B., Morelock, M.M. and Vallee, B.L. (1984) Biochemistry 23, 6363-6368.
Duester, G., Hatfield, G.W., Bühler, R., Hempel, J., Jörnvall, H. and Smith, M. (1984). Proc. Natl. Acad. Sci. USA 81, 4055-4059.
Duester, G., Jörnvall, H. and Hatfield, G.W. (1986a). Nucleic Acids Res. 14, 1931-1941.
Duester, G., Smith, M., Bilanchone, V. and Hatfield, G.W. (1986b). J. Biol. Chem. 261, 2027-2033.
Eckfeldt, J., Mube, L., Takio, K. and Yonetani, T. (1976). J. Biol. Chem. 251, 236-240.
Edenberg, H.J., Zhang, K., Fong, K., Bosron, R.F., and Li, T.-K. (1985) Proc. Natl. Acad. Sci. USA 82, 2262-2266.
Eklund, H., Nordström, B., Zeppezauer, E., Söderlund, G., Ohlsson, I., Boiwe, T., Söderberg, B.-O., Tapia, O., Brändén, C.-I. and Åkeson, Å. (1976). J. Mol. Biol. 102, 27-59.
Eklund, H., Samama, J.P., Wallén, L., Brändén, C.-I., Åkeson, Å. and Jones, T.A. (1981). J. Mol. Biol. 146, 561-587.
Eklund, H., Horjales, E., Jörnvall, H., Brändén, C.-I. and Jeffery, J. (1985). Biochemistry 24, 8005-8012.
Eklund, H., Horjales, E., Vallee, B. and Jörnvall, H. (1986). Eur. J. Biochem., submitted.
Greenfield, N.J. and Pietruszko, R. (1977). Biochim. Biophys. Acta 350, 121-128.
Hedén, L.-O., Höög, J.-O., Larsson, K., Lake, M., Lagerholm, E., Holmgren, A., Vallee, B.L., Jörnvall, H. and von Bahr-Lindström, H. (1986). FEBS Lett. 194, 327-332.
Hempel, J., Pietruszko, R., Fietzek, P. and Jörnvall, H. (1982). Biochemistry 21, 6834-6838.
Hempel, J., Kaiser, R. and Jörnvall, H. (1984a). FEBS Lett. 173, 367-373.
Hempel, J., von Bahr-Lindström, H. and Jörnvall, H. (1984b). Eur. J. Biochem. 141, 21-35.
Hempel, J., Holmquist, B., Fleetwood, L., Kaiser, R., Barros-Söderling, J., Bühler, R., Vallee, B.L. and Jörnvall, H. (1985a). Biochemistry 24, 5303-5307.
Hempel, J., Kaiser, R. and Jörnvall, H. (1985b). Eur. J. Biochem. 153, 13-28.
Höög, J.-O., Hedén, L.-O., Larsson, K., Jörnvall, H. and von Bahr-Lindström, H. (1986a). Eur. J. Biochem., in press.
Höög, J.-O., von Bahr-Lindström, H., Hedén, L.-O., Holmquist, B., Hempel, J., Larsson, K., Vallee, B.L. and Jörnvall, H. (1986b). Biochemistry, submitted.
Hsu, L.C., Tani, K., Fujiyoshi, T., Kurachi, K. and Yoshida, A. (1985). Proc. Natl. Acad. Sci. USA 82, 3771-3775.
Ikuta, T., Fujiyoshi, T., Kurachi, K. and Yoshida, A. (1985). Proc. Natl. Acad. Sci. USA 82, 2703-2707.
Ikuta, T., Szeto, S. and Yoshida, A. (1986). Proc. Natl. Acad. Sci. USA 83, 634-638.
Jeffery, J. and Jörnvall, H. (1983). Proc. Natl. Acad. Sci. USA 80, 901-905.
Jörnvall, H. (1977). Eur. J. Biochem. 72, 443-452.
Jörnvall, H. (1980). In: Dehydrogenases requiring nicotinamide coenzymes: (ed J. Jeffery) Birkhäuser Verlag, Basel. 126-148.
Jörnvall, H. (1986) Chemica Scripta 26, in press.
Jörnvall, H. and Persson, B. (1986). In: Thioredoxin and Glutaredoxin Systems: Structure and Function: (eds A. Holmgren, C.-I. Brändén, H. Jörnvall and B.-M. Sjöberg) Raven Press, New York. 111-120.
Jörnvall, H., Eklund, H. and Brändén, C.-I. (1978). J. Biol. Chem. 253, 8414-8419.
Jörnvall, H., Persson, M. and Jeffery, J. (1981). Proc. Natl. Acad. Sci. USA 78, 4226-4230.
Jörnvall, H., Hempel, J., Vallee, B.L., Bosron, W.F. and Li, T.-K. (1984a). Proc. Natl. Acad. Sci. USA 81, 3024-3028.

Jörnvall, H., von Bahr-Lindström, H., Jany, K.-D., Ulmer, W. and Fröschle, M. (1984b). FEBS Lett. 165, 190-196.
Jörnvall, H., von Bahr-Lindström, H. and Jeffery, J. (1984c). Eur. J. Biochem. 140, 17-23.
Jörnvall, H., Hempel, J. and Vallee, B.L. (1986). Enzyme, in press.
Lange, L.G. and Vallee, B.L. (1976). Biochemistry 15, 4681-4686.
Mårdh, G., Falchuk, K.H., Auld, D.S. and Vallee, B.L. (1986) Proc. Natl. Acad. Sci. USA 83, 2836-2840.
Negelein, E. and Wulff, H.-J. (1937). Biochem. Z. 289, 436-437.
Persson, B., Ålin, P., Jörnvall, H. and Mannervik, B. (1986). In preparation.
Rossmann, M.G., Liljas, A., Brändén, C.-I. and Banaszak, L.J. (1975). In: The Enzymes (ed P.D. Boyer) 3rd ed., Vol. 11, Academic Press, New York. 61-102.
Smith, M., Hopkinson, D.A. and Harris, H. (1971). Ann. Hum. Genet. 34, 251-271.
Vallee, B.L. and Bazzone, T.J. (1983). In: Isozymes: Current topics in biological and medical research: (eds M.C. Rattazzi, J.G. Scandalios and G.S. Witt) Vol. 8, Liss, New York. 219-244.
von Bahr-Lindström, H., Hempel, J. and Jörnvall, H. (1982). J. Prot. Chem. 1, 257-262.
von Bahr-Lindström, H., Hempel, J. and Jörnvall, H. (1984) Eur. J. Biochem. 141, 37-42.
von Bahr-Lindström, H., Jeck, R., Woenckhaus, C., Sohn, S., Hempel, J. and Jörnvall, H. (1985) Biochemistry 24, 5847-5851.
von Bahr-Lindström, H., Höög, J.-O., Hedén, L.-O., Kaiser, R., Fleetwood, L., Larsson, K., Lake, M., Holmquist, B., Holmgren, A., Hempel, J., Vallee, B.L. and Jörnvall, H. (1986). Biochemistry 25, 2465-2470.
von Wartburg, J.-P., Papenberg, J. and Aebi, H. (1965). Can. J. Biochem. 43, 889-898.
Yoshida, A., Huang, I.Y. and Ikawa, M. (1984). Proc. Natl. Acad. Sci. USA 81, 258-261.
Young, E.T. and Pilgrim, D. (1985). Mol. Cell. Biol. 5, 3024-3034.