J Mol Evol (1992) 34:522–535

Journal of Molecular Evolution

Progressive Sequence Alignment and Molecular Evolution of the Zn-Containing Alcohol Dehydrogenase Family

Hong-Wei Sun and Bryce V. Plapp

Department of Biochemistry, The University of Iowa, Iowa City, IA 52242, USA

Summary. Sequences of 47 members of the Zn-containing alcohol dehydrogenase (ADH) family were aligned progressively, and an evolutionary tree with detailed branch order and branch lengths was produced. The alignment shows that only 9 amino acid residues (of 374 in the horse liver ADH sequence) are conserved in this family; these include eight Gly and one Val with structural roles. Three residues that bind the catalytic Zn and modulate its electrostatic environment are conserved in 45 members. Asp 223, which determines specificity for NAD, is found in all but the two NADP-dependent enzymes, which have Gly or Ala. Ser or Thr 48, which makes a hydrogen bond to the substrate, is present in 46 members. The four Cys ligands for the structural zinc are conserved except in ζ-crystallin, the sorbitol dehydrogenases, and two bacterial enzymes. Analysis of the evolutionary tree gives estimates of the times of divergence for different animal ADHs. The human class II (π) and class III (χ) ADHs probably diverged about 630 million years ago, and the newly identified human ADH6 appeared about 520 million years ago, implying that these classes of enzymes may exist or have existed in all vertebrates. The human class I ADH isoenzymes (α, β, and γ) diverged about 80 million years ago, suggesting that these isoenzymes may exist or have existed in all primates. Analysis of branch lengths shows that these plant ADHs are more conserved than the animal ones and that class III ADHs are more conserved than class I ADHs. The rate of acceptance of point mutations (PAM units) shows that selection pressure has existed for ADHs, implying that these enzymes play definite metabolic roles.

Offprint requests to: B.V. Plapp

Key words: Alcohol dehydrogenase — Molecular evolution — Sequence alignment — Phylogenetic tree — Structure and function of proteins

Introduction

Alcohol dehydrogenases (ADHs: EC 1.1.1.1) occur in a wide variety of organisms, including animals, plants, yeasts, and bacteria (Brändén et al. 1975). ADHs can be classified, according to the metal ions contained, into three groups: those with zinc, those without any metal ion, and those with iron (Jörnvall et al. 1987). These three groups are respectively represented by horse liver ADH, *Drosophila* ADH, and ADH2 from *Zymomonas mobilis*.

Many Zn-containing ADHs from different species have been characterized; they exist as dimers and tetramers, as represented by the horse and yeast enzymes, respectively. Human ADHs are dimers. They are grouped (Vallee and Bazzone 1983) into three classes: I, II (π), and III (χ). In addition, a fourth class, human ADH6, has been identified recently (Yasunami et al. 1991). There are three genes for human class I isoenzymes, and the protein subunits coded by these genes are named α, β, and γ (Smith et al. 1971). In addition to the Zn-containing ADHs, some other proteins are related to this family as judged by their sequence identities and possible functional and structural similarities. These include liver sorbitol dehydrogenases (Jörnvall et al. 1981; Eklund et al. 1985; Karlsson et al. 1991), *Escherichia coli* threonine dehydrogenase (Aronson and Somerville 1989), and ζ-crystallin from guinea pig lens (Borrás et al. 1989).

Of the Zn-containing ADHs, the enzyme from horse liver is the most extensively studied (Eklund and Brändén 1987). Its three-dimensional structure shows that the enzyme is a dimer of two identical subunits, each containing two zinc atoms and different domains for binding coenzyme and substrate. Other ADHs of this family are also thought to have these features of horse liver ADH (Eklund and Brändén 1987).

Comparisons using sequence alignments are important for studying structure–function and evolutionary relationships among members of the Zn-containing ADH family. Many sequences have been aligned to that of horse liver ADH E isoenzyme. These studies, however, have been limited by the number of sequences available. An evolutionary tree was presented based on the protein sequences of 17 members of this family (Jörnvall et al. 1987), but without detailed information on branch order. Another tree was prepared for 17 animal and plant ADHs based on their DNA sequences (Yokoyama et al. 1990). Both branch lengths and branch order were given, but the information revealed was limited due to the exclusion of many tetrameric members. We have studied 47 members of the Zn-containing ADH family using computer programs that align sequences progressively and produce a phylogenetic tree based on the alignment (Doolittle and Feng 1990; Feng and Doolittle 1990).

Materials and Methods

Sequences included in this study are 42 ADHs from 26 different species, 3 sorbitol dehydrogenases from 3 different species, a threonine dehydrogenase, and ζ crystallin. Letter codes used in this study, species names, common names, and references for these enzymes and the protein are listed in Table 1. The Genetics Computer Group sequence analysis software package (GCG, version 7.0, April 1991; Devereux et al. 1984) was used to search two protein sequence data banks, National Biomedial Research Foundation (release 29, June 1991) and SwissProt (release 19, August 1991). PAM stands for the accepted point mutations per 100 residues per 100 million years.

A set of C programs running on a VAX 6410/VMS 5.3 system, including FORMAT, SCORE, PREALIGN, TREE, PAPA3, BLEN, and MULPUB was used to generate the progressive alignment and evolutionary trees (Doolittle and Feng 1990; Feng and Doolittle 1990). The alignment uses the algorithm of Needleman and Wunsch (1970) and the minimum mutation matrix of Dayhoff et al. (1978) for the scoring. Gaps are introduced by comparing the most closely related pair of sequences and are retained by the "once a gap always a gap" rule. The evolutionary tree was based on the progressive alignment. The initial versions of the tree had some negative branch lengths, so modifications were made by switching nearby branches and regrouping some members within a cluster. This was facilitated by using the topology information of a tree generated by PAPA3. The final version of the tree had no negative branch lengths and a low percentage standard deviation.

Results

The Alignment

The alignment is given in Fig. 1. It is different from all previous alignments in that it was made progressively rather than pairwise. As a result, it has some gaps or insertions at positions that are different from those assigned previously. For instance, the only insertion for quail ADH appeared before 117 in this alignment instead of before 119 as shown in a previous alignment (Kaiser et al. 1990). For human ADH π, a Ser, a Lys, and two Asn are inserted before 115, 117, and 122 respectively, whereas all four residues were inserted as a single unit before 122 in an alignment made for 17 members of this family (Jörnvall et al. 1987). For barley ADH2, a six-residue gap was previously assigned to position 293–298 (Trick et al. 1988), whereas the current alignment introduced this gap into position 290–295. Perhaps the most striking difference observed is the assignment of a 21-residue deletion for yeast ADHs. This deletion was previously treated as a single gap and assigned to 119–139 (Jörnvall et al. 1987). The present alignment, however, treats this deletion as two gaps and assigns them to 112–127 and 135–139. We have observed that the alignment is dependent on the sequences included, as slightly different assignments for some gaps or insertions were obtained when the number of sequences changed. Nevertheless, the positions for gaps or insertions presented in the current alignment, if different from those assigned previously, should at least provide alternatives for consideration. The ultimate solutions to the position of gaps or insertions will require knowledge of the three-dimensional structures.

With the present alignment, nine residues, eight Gly and one Val, are conserved in all the sequences. These strictly conserved residues can be divided into two clusters. One has four Gly (at positions 66, 71, 77, and 86) and one Val (at 80) and is located in the substrate-binding domain; the other has four Gly (192, 201, 204, and 236) and is located in the coenzyme-binding domain. These 9 residues are among those 22 found strictly conserved when 17 members of this family were aligned (Jörnvall et al. 1987). It is interesting that Val 80, whose conservation was doubted when more sequences became available (Jörnvall et al. 1987), is strictly conserved among all the members of this family. The number of conserved residues increases to 12 when ζ-crystallin of guinea pig lens is excluded, with the three additional residues being Cys 46, His 67, and Glu 68, which are first or second sphere ligands to the catalytic Zn. (Asp 49, another second sphere ligand to Zn, is conserved among all the enzymes except human ADH6, which has Glu 49.) In addition to

Table 1. List of enzymes/protein included in this study

Code	Species	Common name	References
Aeu	*Alcaligenes eutrophus*	*A. eutrophus* ADH	Jendrossek et al. 1988
Ani1	*Aspergillus nidulans*	*A. nidulans* ADH1	Gwynne et al. 1987
Ani3	*A. nidulans*	*A. nidulans* ADH3	McKnight et al. 1985
Ath	*Arabidopsis thaliana*	Mouse-ear cress ADH	Chang and Meyerowitz 1986
Cja	*Coturnix japonica*	Quail ADH	Kaiser et al. 1990
CpoZ	*Cavia porcellus*	Guinea pig lens crystallin ζ	Borrás et al. 1989
EcaE	*Equus caballus*	Horse ADHE	Jörnvall 1970
EcaS	*E. caballus*	Horse ADHS	Park and Plapp 1991
EcaX	*E. caballus*	Horse ADH χ	Kaiser et al. 1989
EcoT	*Escherichia coli*	*E. coli* threonine DH	Aronson and Somerville 1989
Fan	*Fragaria ananassa*	Strawberry ADH	Wolyn and Jelenkovic 1990
Gga	*Gallus gallus*	Chicken ADH	Estonius et al. 1990
HsaA	*Homo sapiens*	Human ADH α	Ikuta et al. 1986; von Bahr-Lindström et al. 1986
HsaB	*H. sapiens*	Human ADH β	Ikuta et al. 1985
HsaC	*H. sapiens*	Human ADH γ	Ikuta et al. 1986; Höög et al. 1986
HsaG	*H. sapiens*	Human sorbitol (glucitol) DH	Karlsson et al. 1991
HsaP	*H. sapiens*	Human ADH π	Höög et al. 1987
HsaX	*H. sapiens*	Human ADH χ	Kaiser et al. 1988
Hsa6	*H. sapiens*	Human ADH6	Yasunami et al. 1991
Hvu1	*Hordeum vulgare*	Barley ADH1	Good et al. 1988
Hvu2	*H. vulgare*	Barley ADH2	Trick et al. 1988
Hvu3	*H. vulgare*	Barley ADH3	Trick et al. 1988
Kla	*Kluyveromyces lactis*	*K. lactis* ADH	Saliola et al. 1990
MacA	*Macaca mulatta*	Rhesus monkey ADH α	Light et al. 1992
MmuA	*Mus musculus*	Mouse ADHA	Edenberg et al. 1985
MmuX	*M. musculus*	Mouse ADH χ	Edenberg et al. 1991
OarG	*Ovis aries*	Sheep sorbitol (glucitol) DH	Karlsson et al. 1991
Osa1	*Oryza sativa*	Rice ADH1	Xie and Wu 1989
Osa2	*O. sativa*	Rice ADH2	Xie and Wu 1990
Pam	*Pennisetum americanum*	Pearl millet ADH	Ha et al. 1989
PhaB	*Papio hamadrysa*	Baboon ADH β	Trezise et al. 1989
Psa	*Pisum sativum*	Garden pea ADH	Llewellyn et al. 1987
RnoA	*Rattus norvegicus*	Rat ADHA	Crabb and Edenberg 1987
RnoG	*R. norvegicus*	Rat sorbitol (glucitol) DH	Karlsson et al. 1991
RnoX	*R. norvegicus*	Rat ADH χ	Julià et al. 1988
Rpe	*Rana perezi*	Frog ADH	Cederlund et al. 1991
Sce1	*Saccharomyces cerevisiae*	Yeast ADH1	Bennetzen and Hall 1982
Sce2	*S. cerevisiae*	Yeast ADH2	Russell et al. 1983
Sce3	*S. cerevisiae*	Yeast ADH3	Young and Pilgrim 1985
Spo	*Schizosaccharomyces pombe*	*S. pombe* ADH	Russell and Hall 1983
Stu	*Solanum tuberosum*	Potato ADH	Matton and Brisson 1990
Tae	*Triticum aestivum*	Wheat ADH	Mitchell et al. 1989
Tbr	*Thermoanaerobium brockii*	*T. brockii* ADH	Peretz and Burstein 1989
Tre	*Trifolium repens*	White clover ADH	Ellison 1989
Zma1	*Zea mays*	Maize ADH1	Dennis et al. 1985
Zma2	*Z. mays*	Maize ADH2	Dennis et al. 1985
Zmo1	*Zymomonas mobilis*	*Z. mobilis* ADH1	Keshav et al. 1990

those strictly conserved residues, there are positions where only a few types of amino acid residues are found. These residues, together with those strictly conserved ones, are given in Table 2.

The number of conserved residues increases greatly when only animal and plant ADHs are considered, with the total number of conserved residues being 86. Even more residues are found conserved when plant and animal ADHs are treated separately. As can be seen in Fig. 1 (indicated by *a* and *p*), 116 (31%) residues are conserved among 18 animal ADHs and 213 (56%) residues are conserved among 14 plant ADHs. Apparently, these plant ADHs diverged less during evolution than the animal ones. In addition, the plant ADHs seem to have more aromatic residues conserved. Specifically, there are more Phe (14 out of 19 as compared to 6 out of 17 for animal ADHs), Trp (3 out of 3 as compared to 1 out of 3), and Tyr (3 out of 6 as compared to 1 out of 5) conserved among these plant ADHs.

The Evolutionary Tree

A phylogenetic tree of this family was produced on the basis of the progressive alignment; it is given in Fig. 2. The branch order for human class I ADHs in the current tree is different from that proposed previously, where different methods were employed and a limited number of sequences were included in the analysis (Ikuta et al. 1986; Trezise et al. 1989; Yokoyama et al. 1990). All three possible alternative ways of arrangement for the human class I ADHs were tested and the present order is the best as judged by the criteria for a better tree. In addition, the current order is supported by the fact that the model of the γ isoenzyme is most similar to the x-ray structure of horse ADH E (Eklund et al. 1987).

The 47 members are classified into two large groups. Group I includes all the dimeric ADHs of animals and plants, and group II has all the tetrameric ones including ζ crystallin. Group I is further divided into two major clusters, one for 18 animal ADHs and the other for 14 plant ADHs. Among the 18 animal ADHs, human class II enzyme diverged first, followed by class III enzymes and human ADH6, and then the class I ADH from frog diverged from the rest of class I enzymes of other species. The distances from points A, B, and C were given in Table 3 for animal and plant ADHs. They were calculated by adding up the corresponding branch lengths in the evolutionary tree.

In addition to the alignment and the evolutionary tree, the percent identities for all the possible pairs of these 47 enzymes/protein were calculated. The SCORE program produced the percent identities based on pairwise alignment. The TREE program, on the other hand, calculated the percent identities based on the progressive alignment. These two sets of percent identities are given in Table 4; the data show that these two sets differ only for distantly related pairs.

Discussion

The Alignment

Sequence alignments are generally valuable as it is unlikely that the three-dimensional protein structures (less than 500 currently) will be determined for most of the known protein sequences (over 20,000 currently). The alignment reveals functionally important residues and usually is the first step in building a reasonable model of the three-dimensional structure. Various sequences of ADHs have been aligned previously (e.g., Jörnvall et al. 1987; Eklund et al. 1990; Xie and Wu 1990). Most of these were produced by aligning the sequence with that of horse liver ADH E, and by applying subjective criteria when a gap or an insertion had to be assigned. In contrast, the alignment presented here, which was used to construct the evolutionary tree, was produced progressively based on objective criteria.

Requirements for an Enzymatically Functional Member

The alignment of 47 members of the Zn-containing ADH family yielded the current minimal requirements for a functional enzyme of this family. Only a few residues are strictly conserved among all 47 members, probably reflecting functional and structural diversities of the different ADHs. The fact that most of the strictly conserved residues are Gly indicates that they are located at crucial positions where a side chain would disrupt a structure that is required for a functional ADH. A stereoview of the positions of these strictly conserved residues in the three-dimensional structure of horse ADH E is illustrated in Fig. 3. It is interesting that the 9 strictly conserved residues are clustered, rather than distributed over the sequence, with the 5 in the catalytic domain being within a 21-residue fragment and the 4 in the coenzyme-binding domain being within a 45-residue fragment. This suggests that these strictly conserved residues are involved in forming compact cores for the two functional domains.

In addition to the nine strictly conserved residues, four more are conserved if ζ crystallin, an NADPH-quinone oxidoreductase (Rao et al. 1992), is not included. These four residues are all related to Zn binding. Cys 46 and His 67 are ligands to the catalytic Zn (Eklund and Brändén 1987). Asp 49 (except human ADH6, which has Glu 49) and Glu 68 are in the second sphere of Zn ligands and have been shown to affect the electrostatic environment near the catalytic Zn for yeast ADH1 (Ganzhorn and Plapp 1988). The third catalytic Zn-binding ligand (at position 174) may also be considered conserved for all but ζ crystallin if we assume Glu and Asp could serve as Zn-binding ligands for sorbitol dehydrogenases, threonine dehydrogenase, and ADHs from *Thermoanaerobium brockii* and *Alcaligenes eutrophus*. In other zinc enzymes, such as carboxypeptidase (Rees et al. 1981) and thermolysin (Holmes and Matthews 1981), Glu is a ligand. In addition, Asp has recently been identified as a Zn ligand in *E. coli* alkaline phosphatase (Kim and Wyckoff 1991). Thus, it seems that one major event during the evolution of ADHs is their acquisition of the capability to bind a Zn in their active sites.

There are two other residues that seem to be functionally relevant. One of these is Asp 223, which is conserved in all but ζ crystallin and the ADH from *T. brockii* (which use NADP instead of NAD). Asp 223 has been suggested to be important in deter-

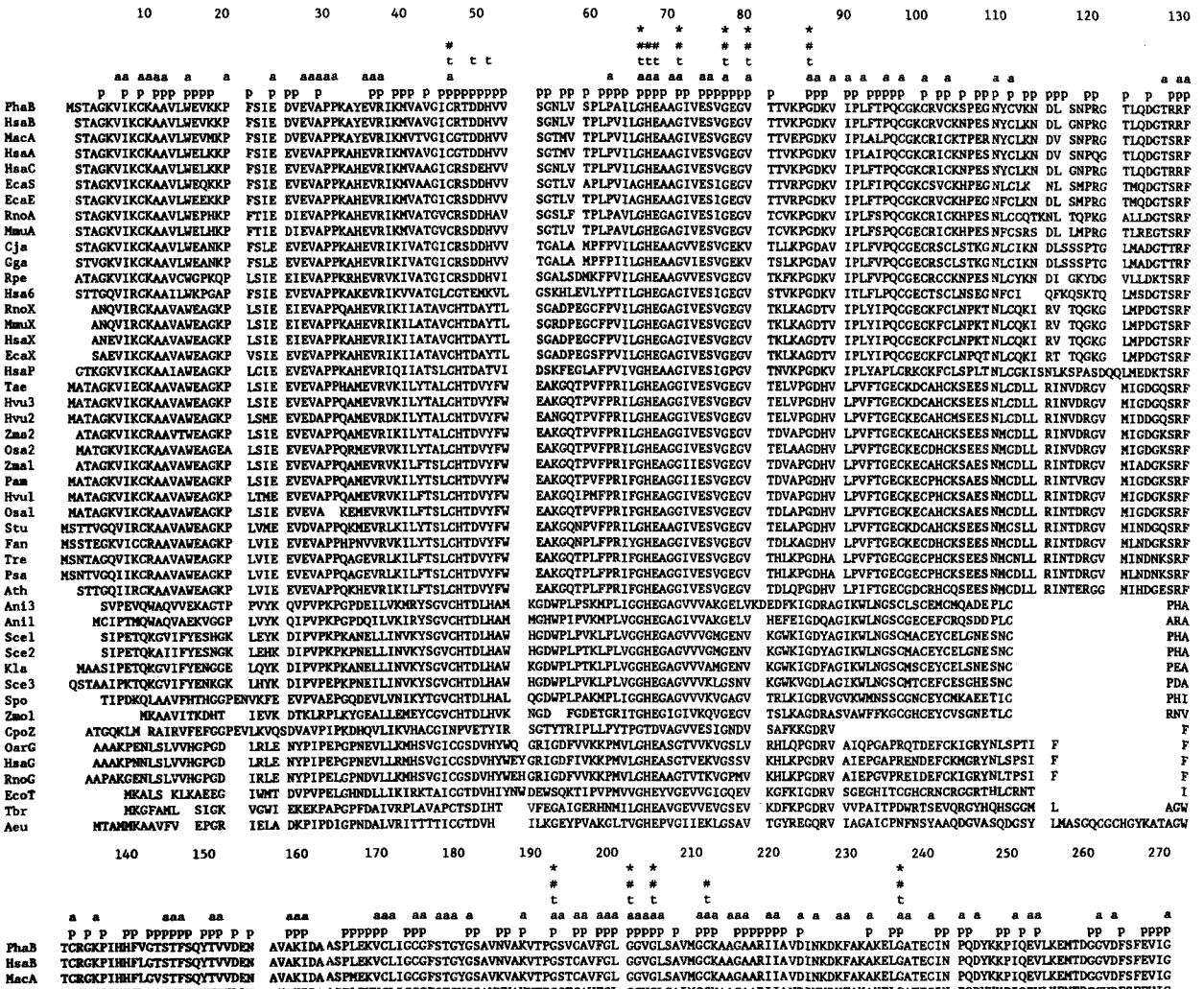

Fig. 1. Progressive sequence alignment of 47 members of the Zn-containing ADH family. *a*: Conserved in all animal ADHs (18); *p*: conserved in all plant ADHs (14); *t*: conserved in all tetrameric ADHs (14); #: conserved in all but ζ crystallin; *: conserved in all 47 members. For enzyme abbreviations, see Table 1.

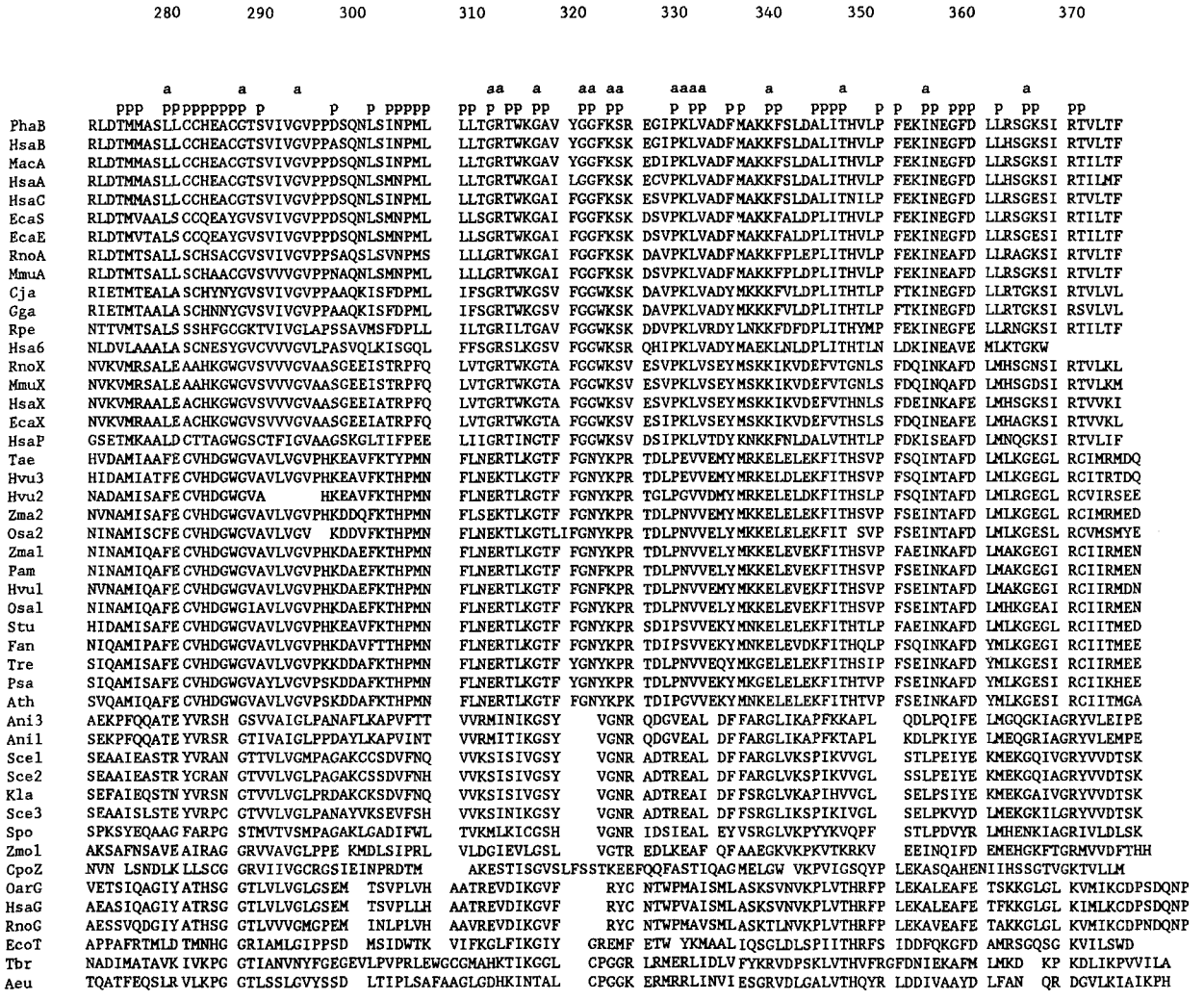

Fig. 1. Continued

mining the coenzyme specificity for ADHs in general (Ohlsson et al. 1974; Brändén et al. 1975; Eklund et al. 1984), and data supporting this role have been obtained for yeast ADH1 (Fan et al. 1991). The second functionally relevant residue is Ser 48 or Thr 48, which is conserved for all the members except ζ crystallin. Ser 48 is hydrogen-bonded to the hydroxyl group of the alcohol bound to the catalytic Zn and could function in a proton relay system to facilitate removal of the proton from the alcohol (Eklund et al. 1982). Substitution of Thr 48 with Ala in yeast ADH1 inactivates the enzyme (Plapp et al. 1991).

It appears, therefore, that the minimal requirements for alcohol dehydrogenases of this family include the following: several Gly residues at certain positions that are required to form a basic folded structure; residues that are necessary to bind the catalytic Zn and to modulate its electrostatic environment; an Asp that determines the specificity for NAD; and a Ser or Thr that facilitates proton removal from the substrate.

Conservation of Ligands for the Noncatalytic Zinc

It is interesting that the four residues responsible for binding the noncatalytic Zn, including cysteines 97, 100, 103, and 111, were conserved among all the members except the three sorbitol dehydrogenases and the ADHs from *A. eutrophus* and *T. brockii*. Sorbitol dehydrogenase contains one zinc atom per subunit (Jeffery et al. 1984). The two bacterial ADHs probably also lack the noncatalytic Zn. A structural role was proposed for the noncatalytic Zn of ADHs (Drum et al. 1969; Brändén et al. 1975). In contrast to the catalytic Zn ligands, which originate from different parts of the protein chain, the noncatalytic Zn ligands are close to each other, with the four cysteines being within a 14-residue fragment. This organization also supports a structural role for the noncatalytic Zn because it has been found for some other proteins that catalytic metal atoms have ligands from different parts of a protein, whereas structural metal atoms have ligands from the same

Table 2. Conserved and consensus residues among members of the Zn-coating ADH family

No.	Residue	Exceptions	No.	Residue	Exceptions
31	P	I(Aeu), L(RnoG, EcoT, Zmo1), Del(Osa1)	192	G	
35	E	V(Fan), D(RnoG, EcoT, Tbr, Aeu), Q(Ani1, CpoZ)	196	A	C(Spo), L(HsaG, OarG, RnoG, EcoT, CpoZ)
46	C	N(CpoZ)	197	V/I	
47	G/H/R	P(CpoZ)	199	G	A(Aeu)
48	S/T	V(CpoZ)	200	L/A	I(Tbr), Q(Aeu)
49	D	E(Hsa6, CpoZ)	201	G	
51	H/Y	Y(HsaP), K(Hsa6)	202	G/A/P	
62	P	N(Tbr), G(Asu, Zmo1)	203	V/L/I	
66	G		204	G	
67	H	T(CpoZ)	207	V/A	N(OarG), T(HsaG, RnoG)
68	E	D(CpoZ)	212	K/R	T(Sce3)
71	G		216	A/G	N(Zmo1), Del(CpoZ)
73	V	I(Gga, Zma1, Pam & Aeu)	218	R	K(Zmo1, CpoZ), T(Aeu), Q(HsaG, OarG, RnoG), N(EcoT)
77	G		219	I/V	T(Ani3)
80	V		220	I/V/L	
86	G		221	G/A	V(HsaG, OarG, RnoG), I(EcoT)
87	D	Q(Aeu)	222	V/I	T(OarG, HsaG, EcoT, CpoZ)
89	V/A		223	D	G(Tbr), A(CpoZ)
97	C	D(Tbr), N(Aeu), R(HsaG, OarG, RnoG), Del(CpoZ)	228	K/R	C(Tbr), Q(Zmo1), G(CpoZ)
100	C	T(Tbr), S(Aeu), D(HsaG, OarG, RnoG), Del(CpoZ)	236	G	
103	C	V(Tbr), A(Aeu), Del(CpoZ)	261	G	K(HsaG, OarG, RnoG)
111	C	S(Tbr, OarG, HsaG), D(Aeu), Del(CpoZ), T(RnoG)	287	G	S(Spo)
144	S/G		292	V	I(HspA, Ani1, Ani3), L(Aeu, EcoT) Del(Hvu2)
146	F	M(Zmol), Q(Aeu), L(HsaG, OarG, RnoG), Y(CpoZ)	293	G	S(Spo), N(Tbr), Del(Hvu2)
159	K/H/R	I(Spo), P(Aue)	294	V/L	Del(Hvu2), M(Sce1, Spo, RnoG), C(CpoZ), I(EcoT), Y(Tbr)
160	I/V/L		359	F/Y	V(Hsa6), H(CpoZ)
169	V/A	G(HsaG, OarG, RnoG, CpoZ)	365	G	N(Spo), Del(Tbr, Aeu), S(CpoZ)
174	C	D(Tbr, Aeu, EcoT), E(HsaG, OarG, RnoG), I(CpoZ)	369	R/K	Del(Hsa6), D(Aeu)
178	T	S(Osa2), V(HsaG, OarG, RnoG), N(EcoT)			

The numbers refer to the numbering system for horse liver ADH E. For enzyme abbreviations, see Table 1. Del: deletion

part of a protein (Matthews et al. 1974; Monaco et al. 1978; Vallee and Auld 1990; Kim and Wyckoff 1991). Whether the second Zn has a structural role remains to be explored, but the conservation of these residues indicates that they may be necessary at least for the eukaryotic ADHs. Thus, it seems that another major event during the evolution of ADHs was the acquisition of a capability for binding the noncatalytic Zn.

Other Conserved Amino Acids

His or Tyr is found at position 51 except for human class II ADH and ADH6, which have Thr and Lys at this position, respectively. Thus, there is His in all the class I and tetrameric enzymes and Tyr in all the rest, including all the plant ADHs and class III ADHs. Although both His 51 and Tyr 51 can form a hydrogen bond to the 2′-hydroxyl group of the nicotinamide ribose of NAD (Eklund et al. 1990), they must function differently. His 51 appears to act as a base catalyst for alcohol oxidation through the proton relay system (Eklund et al. 1982; Ehrig et al. 1991; Park 1991; Plapp et al. 1991). A role for Tyr 51, however, is more difficult to assign. The Tyr may bind the coenzyme but might not be required for base catalysis under the physiological conditions where the enzyme functions. It may be relevant that Tyr is also aligned at position 51 in ζ crystallin.

There are several other well-conserved positions. Asp is conserved at position 87 in all but the ADH of *A. eutrophus,* where it is Gln. Gly is conserved at position 199 except for the ADH from *A. eutrophus,* which has Ala. Lys or Arg is conserved at position 212 in all but ADH3 from *Saccharomyces cerevisiae,* which has Thr. Gly is found at position 261 for all but the sorbitol dehydrogenases, which have Lys. Gly is located at position 287 for all except for the ADH of *Schizosaccharomyces pombe,* which has Ser. Arg or Lys is conserved at position 369 except for *A. eutrophus* ADH, which has Asp. Although the structural or functional implications of these observations remain to be elucidated, they suggest directions for future studies.

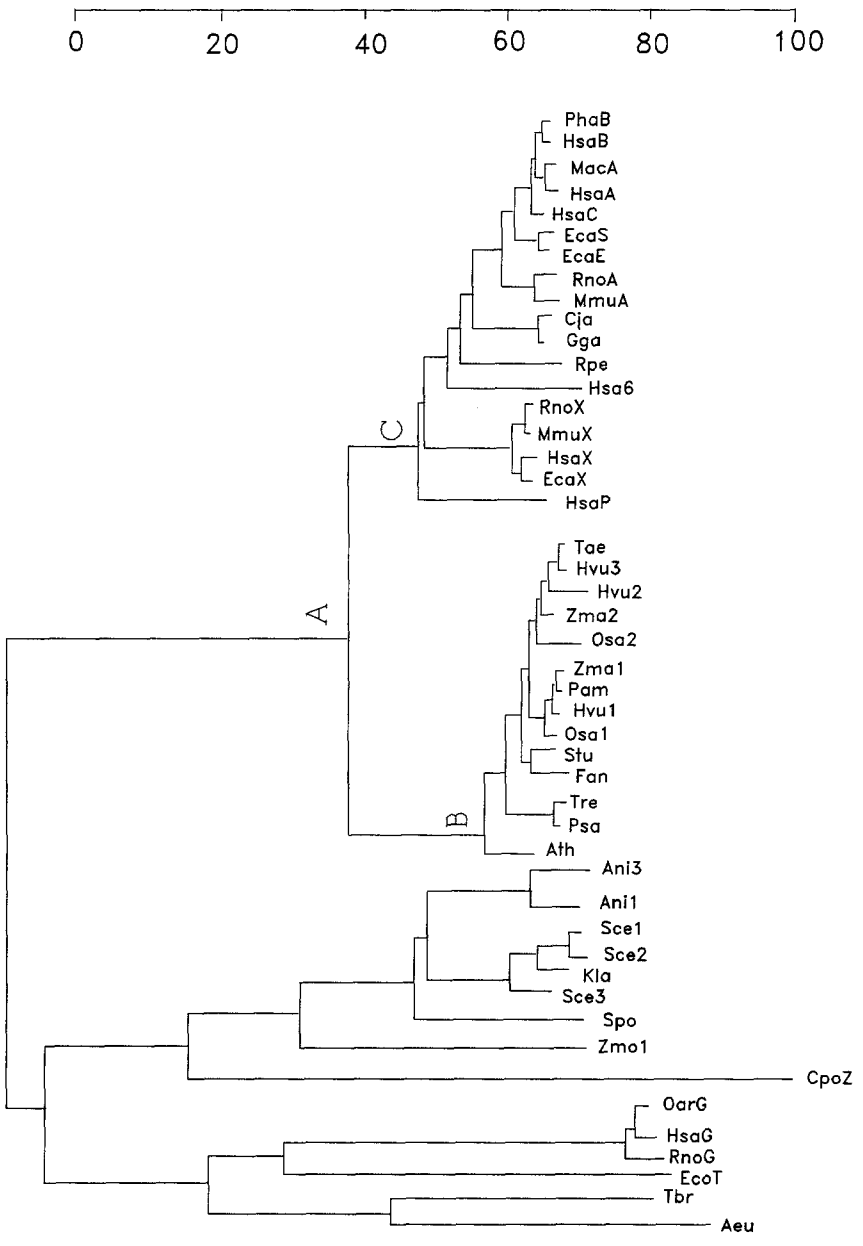

Fig. 2. Phylogenetic tree for 47 members of the Zn-containing ADH family. Percent standard deviation was 5.00 based on the progressive alignment. For enzyme abbreviations, see Table 1.

Assignment of Insertions or Deletions

Some potentially important residues were not aligned at the same positions as in previous reports. Leu was aligned at position 93 for yeast ADHs, instead of Trp. Whether Trp or Leu was assigned to the position was controlled by the insertion of one Gly. Trp would be at position 93 if a Gly were inserted between residues 95 and 96 instead of between 89 and 90. (The insertion produces a gap between these residues in the other sequences.) Residue 93 has been shown to determine the size of the active site of horse liver ADH (Eklund et al. 1982) and has been assumed to restrict the activities of yeast ADHs toward larger substrates due to the bulky side chain of Trp (Brändén et al. 1975). More importantly, the assignment of Trp to position 93 in yeast ADHs is supported by experimental data obtained from mutant enzymes of yeast ADH1 in which Trp was replaced by Ala or Phe (unpublished data from this laboratory; Creaser et al. 1990). Apparently, the assignment of Leu for Trp at position 93 produced a "better but less reliable" local alignment. Inserting a Gly between 89 and 90 gives a local alignment with better similarity because Ile is then aligned at position 90 where Ile is also found in all the animal ADHs. The alternative assignments for the positions of gaps or insertions should emphasize that the alternatives remain hypothetical until the three-dimensional structures are determined.

The Evolutionary Tree

Another major result of this study is the evolutionary tree for 47 members of the Zn-containing ADH

Table 3. Distances for animal and plant ADHs since their divergence

	Animal			Plant	
	From A	From C		From A	From B
Pha	26.87	17.37	Tae	26.86	7.87
HsaB	26.58	17.08	Hvu3	26.84	7.85
MacA	28.28	18.78	Hvu2	31.87	12.88
HsaA	28.90	18.78	Zma2	26.24	7.25
HsaC	26.59	17.09	Osa2	30.47	11.48
EcaS	27.12	17.62	Zma1	26.40	7.41
EcaE	26.59	17.09	Pam	26.45	7.46
RnoA	27.35	17.85	Hvu1	27.27	8.28
MmuA	27.75	18.25	Osa1	27.71	8.72
Cja	27.01	17.51	Stu	26.06	7.07
Gga	26.31	16.81	Fan	28.01	9.02
Rpe	28.85	19.35	Tre	27.58	8.59
Hsa6	32.03	22.53	Psa	26.67	7.68
RnoX	23.07	13.57	Ath	24.16	5.17
MmuX	23.13	13.62			
HsaX	23.35	13.85			
EcaX	22.75	13.25			
HsaP	27.55	18.05			
Average	27 ± 2	17 ± 2		27 ± 2	8 ± 2

Distances were calculated by adding up the corresponding branch lengths shown in the evolutionary tree, which has a % standard deviation of 5.00. Branch points A, B, and C are as indicated in the tree. For abbreviations of enzymes, see Table 1

family. This tree was generated by the matrix-based method (Feng and Doolittle 1990), and was supported by an analysis using the nearest-neighbor procedure (Doolittle and Feng 1990). Like many other evolutionary trees for related sequences, this tree was produced assuming a common ancestor for all the members included. In general, there are two features that are important for an evolutionary tree: its topology or branch order, which shows how related members are grouped and have diverged from each other, and its distances or branch lengths, which should be proportional to the true evolutionary distances.

Topology of the Tree

The topology of the tree for this family shows several features. First, the two clusters and four subclusters have distinct structural or species differences. All the dimeric ADHs were clustered and further subclustered into two groups for animal ADHs and plant ones; all the tetrameric ones were clustered and further subclustered into two groups, with one including enzymes having two zinc atoms per subunit (represented by yeast ADHs), and the other including enzymes having one zinc atom per subunit (represented by the sorbitol dehydrogenases). It has been shown that evolutionary trees based on three-dimensional structures of proteins are almost identical to trees based on primary structures (Johnson et al. 1990). Thus, it is reasonable to assume that members of a cluster are more similar than nonmembers in their three-dimensional structures. In this regard, it is important to establish the three-dimensional structure for at least one member from each subcluster. The three-dimensional structures of horse ADH E and the human ADH β1 (Hurley et al. 1991) are the only ones known for this family.

Secondly, the topology of the tree shows that among all the animal ADHs included, the three classes of human ADHs diverged before speciation occurred for most animal ADHs. The divergence did not happen in humans, but instead in an an-

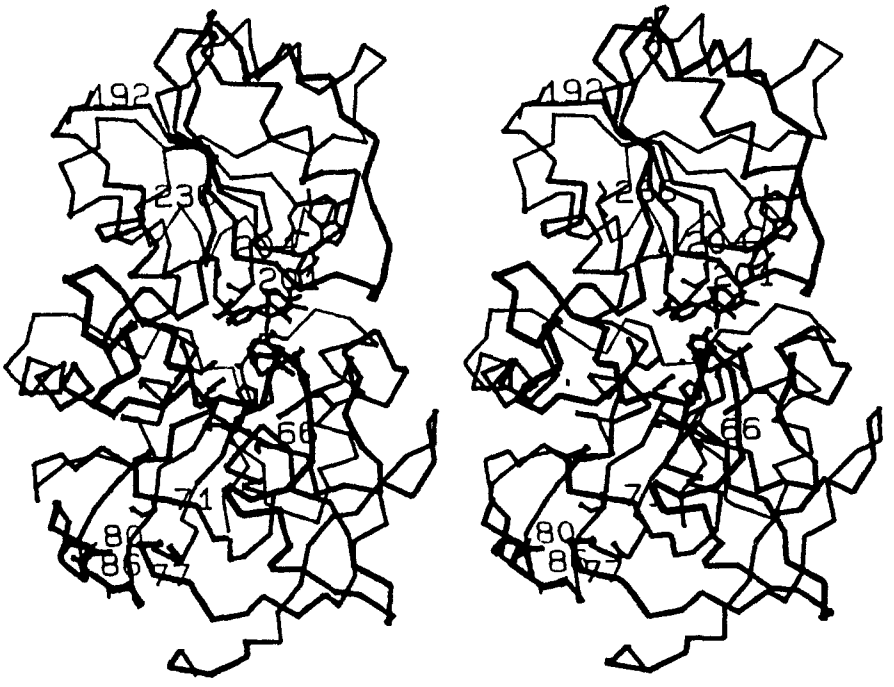

Fig. 3. Stereoview for locations of nine strictly conserved residues in the three-dimensional structure of horse alcohol dehydrogenase E isoenzyme. NAD^+ is bound in the cleft between the two domains of the subunit.

Table 4. Percent identities (based on aligned regions). Values above the diagonal were from the progressive alignment. Those below the diagonal were from the pairwise alignment

	Cja	Gga	PhaB	HsaB	HsaC	MacA	HsaA	EcaS	EcaE	RnoA	MmuA	Rpe	Hsa6	RnoX	MmuX	HsaX	EcaX	HsaP	Tae	Zma1	Pam	Hvu1	Osa1	Zma2	Hvu3	Stu	Fan	Hvu2	Psa	Tre	Ath	Osa2	Tbr	Aeu	HsaG	OarG	RnoG	EcoT	Sce1	Sce2	Kla	Sce3	Ani3	Ani1	Spo	Zmo1	CpoZ
Cja		97	74	74	74	72	72	75	75	72	73	70	65	65	64	65	65	60	49	50	49	48	49	50	49	49	48	49	46	47	50	47	28	24	24	25	26	26	25	27	26	25	27	25	30	27	25
Gga	97		74	74	74	71	72	74	74	72	72	69	65	64	63	64	65	59	49	50	49	48	48	50	49	49	48	48	46	47	50	47	28	24	24	24	25	26	26	27	26	25	26	24	29	27	25
PhaB	74	74		97	93	94	92	86	87	81	83	67	64	62	61	62	62	61	50	50	50	49	49	50	50	50	49	49	47	48	50	47	28	24	25	26	24	25	26	27	26	26	26	26	27	29	25
HsaB	74	74	97		94	94	94	86	87	82	83	69	64	63	62	63	63	61	49	49	49	49	49	49	49	50	49	49	46	48	51	47	27	25	25	26	24	25	26	27	26	26	26	26	28	29	25
HsaC	74	74	93	94		92	93	87	89	82	84	69	63	63	62	63	63	60	49	49	49	48	48	49	49	49	48	48	45	47	50	47	27	24	24	26	24	24	26	26	26	25	26	25	27	28	25
MacA	72	71	94	94	92		94	86	86	80	82	67	63	61	60	61	61	60	49	49	49	48	48	49	49	49	48	47	45	47	50	46	27	25	25	26	24	26	26	27	26	26	26	26	27	29	24
HsaA	72	72	92	94	93	94		87	87	81	83	68	63	64	63	64	63	60	49	49	49	48	49	49	49	49	48	48	46	47	50	46	28	25	25	26	25	25	25	26	26	25	26	25	27	28	25
EcaS	75	74	86	86	87	86	87		97	81	83	68	63	63	62	63	63	61	50	50	50	49	48	50	49	49	48	48	45	46	50	46	26	22	24	25	23	24	24	24	24	24	25	25	25	27	25
EcaE	75	74	87	87	89	86	87	97		82	85	68	64	63	62	62	62	61	50	51	51	50	49	51	50	50	49	48	46	47	51	47	26	22	24	24	23	26	24	25	25	24	25	25	26	26	25
RnoA	72	71	81	82	82	80	81	82	82		89	68	62	64	63	63	63	59	49	49	49	49	48	48	49	49	47	48	46	47	50	46	27	24	25	25	25	25	26	27	27	26	27	27	26	30	25
MmuA	73	72	83	83	84	82	83	83	85	89		69	62	64	63	63	63	58	48	49	49	48	48	48	47	49	48	47	45	47	49	46	26	23	24	24	24	25	27	27	26	27	27	27	27	30	24
Rpe	70	69	67	69	69	67	68	68	68	68	69		59	61	61	62	61	56	49	49	49	48	48	47	49	47	48	48	46	47	48	47	28	22	26	26	26	25	23	23	23	23	25	24	28	26	25
Hsa6	65	65	65	65	63	64	64	64	65	63	62	59		59	58	59	60	58	46	47	47	46	45	47	46	47	45	45	43	44	47	45	24	20	24	25	25	22	25	25	25	24	25	25	25	24	23
RnoX	65	64	62	63	63	61	64	63	63	64	64	61	59		97	94	92	63	54	55	54	54	54	54	53	54	53	54	51	53	54	53	27	25	23	24	25	24	23	23	24	24	26	26	30	27	25
MmuX	64	63	61	62	62	60	63	62	62	63	63	61	58	97		93	91	62	54	54	53	53	54	53	53	53	54	51	52	53	53	26	24	23	24	25	23	23	24	24	26	26	30	28	26
HsaX	65	64	62	63	63	61	64	63	62	63	63	62	59	94	93		95	62	54	55	54	54	54	54	53	54	53	53	51	52	54	53	28	24	23	24	24	24	24	24	25	25	26	26	29	28	25
EcaX	65	65	62	63	63	61	63	63	62	63	63	61	60	92	91	95		63	54	54	54	53	54	54	53	53	53	54	50	52	54	53	28	24	23	24	25	23	24	24	24	24	27	26	30	28	25
HsaP	60	59	61	61	60	60	60	61	61	59	58	56	56	63	62	62	63		49	48	48	48	48	48	49	49	49	49	47	47	49	47	28	26	23	23	22	24	23	23	24	24	26	24	27	25	23
Tae	50	50	51	51	50	50	50	51	51	50	49	49	47	54	54	54	54	49		85	86	86	85	90	97	85	82	90	71	79	78	87	27	20	24	24	24	23	23	23	23	25	24	25	25	26	21
Zma1	50	50	51	51	50	50	50	51	52	49	50	49	47	55	54	55	54	49	84		97	94	94	87	85	86	85	81	83	83	81	83	27	24	22	23	22	24	21	21	22	22	24	24	24	23	23
Pam	49	50	51	51	50	50	50	51	52	49	50	49	47	55	54	55	54	48	86	97		96	94	88	86	85	84	82	83	83	81	84	28	23	22	23	22	23	22	21	23	23	24	24	26	23	23
Hvu1	49	49	50	49	49	48	49	50	51	49	49	48	46	54	53	54	54	48	86	94	96		94	87	85	84	83	81	82	83	81	83	27	23	23	23	23	24	22	21	22	23	24	24	25	23	23
Osa1	50	48	49	50	49	49	49	50	51	48	49	49	45	54	54	54	54	48	86	94	94	94		87	85	84	84	81	82	83	81	84	26	23	22	22	23	24	22	22	23	23	25	24	26	24	23
Zma2	50	50	51	51	51	50	50	51	52	49	49	47	47	55	54	54	54	49	90	87	88	87	87		90	83	82	85	81	82	79	89	28	22	22	23	22	24	23	23	24	25	24	24	26	24	24
Hvu3	50	50	51	51	50	50	50	51	51	49	48	49	47	54	53	53	54	49	97	85	86	85	85	90		84	81	90	78	79	78	88	28	22	24	24	24	24	24	23	24	25	24	24	26	24	22
Stu	49	49	51	51	50	50	49	50	51	50	50	48	48	55	54	54	54	49	85	86	85	84	85	83	85		85	82	83	83	84	80	27	21	23	24	24	23	22	22	24	24	24	25	25	26	23
Fan	49	49	50	50	49	49	49	50	51	47	49	48	46	53	53	53	53	49	82	85	84	83	84	82	81	85		79	82	83	81	80	28	23	23	23	24	25	22	22	23	24	25	25	25	23	24
Hvu2	49	49	50	51	49	48	49	49	50	49	49	49	46	55	55	54	55	49	90	81	82	81	81	85	90	82	79		76	78	75	84	28	21	22	22	22	22	21	21	23	24	26	25	26	25	21
Psa	47	47	48	48	47	47	47	47	48	46	47	46	44	51	51	51	50	48	78	83	83	82	83	81	78	83	82	76		96	84	79	26	21	22	22	22	25	23	22	24	24	24	24	26	27	22
Tre	47	47	50	49	48	48	48	48	49	48	48	47	44	53	53	53	52	48	80	83	84	83	84	82	79	83	83	78	96		84	80	26	21	22	22	22	25	23	23	24	25	24	25	26	27	23
Ath	50	50	51	52	51	51	51	51	52	51	50	48	48	54	53	54	53	49	78	81	81	81	82	79	78	84	81	75	84	84		77	26	22	22	22	22	24	22	22	24	24	24	24	26	26	24
Osa2	48	48	49	48	49	47	48	48	49	48	48	48	45	54	53	53	54	47	88	83	85	83	85	89	89	81	80	85	79	80	77		28	22	21	21	21	24	23	23	24	24	24	25	25	26	22
Tbr	31	30	31	32	31	29	30	28	30	30	29	29	26	30	28	31	30	32	30	28	28	28	26	29	30	28	29	28	27	27	28	29		36	26	26	24	24	16	16	16	18	20	18	21	18	19
Aeu	24	26	26	27	24	27	27	23	24	26	23	23	24	28	26	26	27	28	25	26	27	25	26	25	25	25	25	27	25	24	24	25	36		28	26	25	28	21	21	19	21	21	21	21	20	21
HsaG	24	25	27	27	27	26	27	26	26	26	26	28	23	25	26	26	26	25	27	24	25	26	24	26	27	25	26	25	23	23	25	26	28	28		90	82	27	22	23	22	25	24	22	24	22	24
OarG	25	25	27	27	27	27	27	26	27	25	26	29	24	25	26	26	26	24	26	24	25	25	23	24	25	26	26	23	24	24	25	25	27	28	88		83	27	21	22	22	24	23	22	23	21	23
RnoG	25	25	26	26	26	27	28	25	26	26	26	29	23	26	26	27	27	24	27	25	25	26	25	25	27	26	27	24	25	25	26	26	27	27	82	83		29	21	21	21	23	23	22	24	23	23
EcoT	27	27	27	27	26	28	27	27	28	26	26	27	25	27	26	29	25	25	24	26	26	25	25	25	25	25	27	23	25	25	26	24	26	31	28	27	30		26	26	25	25	24	25	25	24	24
Sce1	26	26	27	27	27	26	26	24	26	28	27	23	26	24	24	24	24	22	26	22	24	24	24	24	26	24	24	23	25	25	24	24	20	25	24	22	23	29		93	86	80	56	55	52	36	23
Sce2	27	27	27	27	27	26	26	24	26	27	27	23	26	24	23	24	24	22	26	22	23	23	24	24	26	23	24	23	25	25	24	24	21	25	24	23	22	29	93		84	79	56	55	52	37	23
Kla	26	26	27	26	28	26	26	24	26	27	28	24	25	25	25	26	25	24	27	24	26	26	27	26	27	25	26	24	27	26	28	26	20	24	24	23	23	28	86	84		81	57	57	50	37	24
Sce3	26	25	26	26	26	25	26	23	25	29	28	23	25	25	25	25	25	19	27	24	24	24	26	26	28	26	26	25	26	26	26	26	23	21	26	25	24	27	80	79	81		56	55	49	36	22
Ani3	27	27	27	27	26	26	26	24	24	27	27	26	29	26	27	25	26	27	26	25	26	26	27	26	26	27	28	26	28	28	27	26	24	25	25	24	25	27	56	56	57	56		78	54	41	20
Ani1	29	26	26	26	25	26	26	27	27	29	28	25	28	27	27	27	27	25	28	26	27	27	27	27	28	28	27	28	28	28	27	27	24	23	24	24	24	28	55	55	56	55	78		50	36	20
Spo	29	28	29	29	29	27	30	27	26	27	27	29	26	31	31	31	32	30	26	25	26	26	27	26	27	26	26	25	26	25	26	26	23	23	27	26	25	24	52	52	50	49	54	49		37	21
Zmo1	29	29	34	33	32	34	32	30	30	33	32	28	27	31	31	32	32	28	28	25	27	26	29	28	28	27	26	28	29	29	27	27	23	24	26	25	27	26	36	37	37	37	41	36	37		19
CpoZ	24	24	24	25	24	23	24	22	24	26	23	25	20	22	24	21	23	24	19	18	19	19	20	21	19	21	22	17	20	20	22	17	18	26	24	25	24	26	25	25	25	22	23	22	22	20	

cestor that was common to the animals (all vertebrates) whose ADHs have been included in this study. This implies that these three classes of enzymes may exist or have existed, not necessarily in active form in many, if not all, vertebrates. This implication is consistent with an earlier investigation of class III ADH from rat liver (Julià et al. 1988).

Thirdly, the topology of this tree shows that human class I ADHs (α, β, and γ) diverged before human, baboon, and monkey class I ADHs, supporting the proposal that duplications of the gene for primate class I ADH could have predated primate radiation (Trezise et al. 1989). This indicates that there should be isoenzymes of class I ADH in primates other than humans. This implication is supported by evidence for the existence of five major forms of baboon class I ADH (Holmes et al. 1986; Holmes and VandeBerg 1986). It is also consistent with the analysis of chromosomal DNA samples derived from various primates (Yasunami et al. 1990).

Finally, the newly identified human ADH6 (Yasunami et al. 1991) also diverged before speciation occurred for most animal ADHs. This enzyme may be the stomach μ or σ ADH (Chen and Yoshida 1991), and its position in the present tree suggests that other vertebrates may have similar enzymes.

Branch Lengths of the Tree

If the topology reveals order of divergence, the analysis of branch lengths should then give estimates of the dates of divergence, provided that an evolutionary clock is available. Thus, if we assume that animals and plants diverged 1000 million years ago (Carroll 1988), it can be estimated that human ADH π, ADH χ, and ADH6 diverged about 640, 630, and 520 million years ago, respectively.[1] Therefore, the divergence of three classes of human ADHs occurred about 600 million years ago, which approximates the estimated time when vertebrates diverged from invertebrates (Carroll 1988). This estimated divergence time for the three classes of ADHs is about 400 million years earlier than the estimated times when mammals and birds, respectively, diverged from reptiles. Similarly, the divergence of human class I ADHs can be estimated to have occurred about 80 million years ago, which coincides with the time of divergence for primates and rodents (Carroll 1988). These analyses support the idea that three classes of ADHs may exist or may have existed in many vertebrates, and that isoenzymes of class I ADH may exist or have existed in all primates. More specifically, these analyses imply that the three classes of ADHs may exist or have existed in all mammals, birds, and reptiles. Supporting this implication is the fact that class I ADHs have been identified in mammals, birds, and amphibians, and that class III ADHs also exist in rat, mouse, horse, and human species (Julià et al. 1988; Kaiser et al. 1989; Edenberg et al. 1991).

The proposal that ADHs of three different classes may be ubiquitous in most or all vertebrates has also been made previously, based on the estimated divergence time for the ADH of frog, which was 430 million years ago (Cederlund et al. 1991). Frog ADH is the first class I ADH to have diverged (among those included) as shown in the present tree. Its divergence time can be estimated as being about 430 million years ago. The two estimations are identical; both indicate that the three classes diverged before the radiation of vertebrates. It should be emphasized that estimations of divergence times based on the branch lengths are approximate. This is because the number of sequences included is limited, and more importantly because the evolutionary rate for each ADH does not have to be constant. Nevertheless, it is encouraging to observe that the estimates made in this work are consistent with or supported by several earlier studies.

Further analysis of branch lengths reveals varying degrees of conservation for different groups of this family. The average distances since the beginning of their radiation are 17 and 8 for the animal ADHs and the plant ones, respectively. Thus, it appears that these plant ADHs are more conserved than the animal ones.

PAM Units

The values of PAM units are 4.4 [(165/378) × 10] for plant ADHs[2] and 6.9 [(259/375) × 10] for animal ones. These were calculated based on the progressive alignment and the assumption that animals and plants diverged 1000 million years ago (Carroll 1988). Normally, these calculated values should be corrected for multiple mutations at a site, but values less than 10 need not be. The PAM values for ADHs are higher than those for histone IV (0.09) and glyceraldehyde 3-phosphate dehydrogenase (2.2), comparable with those for insulin (3.5) and trypsinogen (5.1), and lower than those for lysozyme (10) and hemoglobin α and β chains (14) (Dayhoff 1976). Similar analysis can be applied to compare ADHs of class III with those of class I. The PAM value for four class III enzymes from four different species (human, horse, mouse, and rat) is 1.1 [(41/375) ×

[1] The average distance for 18 animal ADHs since the class II (π) enzyme diverged is 17.2: 17.2 × (1000/26.7) = 644

[2] The total number of conserved residues is 213 for the plant ADHs. The average sequence length of the plant ADHs is 378. [(378−213)/378] × 100 × $10^8/10^9$ = (165/378) × 10 = 4.4

10], whereas the PAM value for four class I enzymes from the same four species is 2.4 [(90/375) × 10]. This indicates that the class III ADHs are about twice as conserved as the class I enzymes. This observation is consistent with the results of a previous study, where it has been found that class III ADHs are less variable as compared to class I enzymes (Kaiser et al. 1989).

Human class I ADHs are the major enzymes responsible for ethanol oxidation in the liver, with class II ADH contributing less than 15% of the metabolism (Li et al. 1977). The human class III ADH is almost inactive on ethanol although it is as active as other human ADHs for alcohols with longer chains (Parés and Vallee 1981). Recently, the class III enzymes have been identified as glutathione-dependent formaldehyde dehydrogenases (Koivusalo and Uotila 1991).

Limited information is available about the physiological roles of the ADHs. Yeast ADH1 catalyzes the terminal step in glycolytic fermentation. Rice and maize ADHs are induced under anaerobic conditions where glycolysis is required (Ricard et al. 1986; Kadowaki et al. 1988; Xie and Wu 1989). Human ADH class I enzymes, which have broad substrate specificities, have been suggested to play a general role in the detoxification of various hydroxylated compounds (Kassam et al. 1989). Although the physiological roles of the ADHs remain to be fully established, this study clearly shows that selection pressures exist for plant and animal ADHs, implying that they have specific and important metabolic roles.

Acknowledgments. We thank Drs. Da Fei Feng and Russell Doolittle for providing the programs, and Zhongwen Chen for adapting these programs on the VAX system. This work was supported by USPHS grant AA06223.

References

Aronson BD, Somerville RL (1989) The primary structure of *E. coli* L-threonine dehydrogenase. J Biol Chem 264:5226–5232

Bennetzen JL, Hall BD (1982) The primary structure of the *Saccharomyces cerevisiae* gene for alcohol dehydrogenase I. J Biol Chem 257:3018–3025

Borrás T, Persson B, Jörnvall H (1989) Eye lens ζ crystallin relationships to the family of "long-chain" alcohol/polyol dehydrogenases. Protein trimming and conservation of stable parts. Biochemistry 28:6133–6139

Brändén CI, Jörnvall H, Eklund H, Furugren B (1975) Alcohol dehydrogenases. In: Boyer PD (ed) The enzymes, ed 3, vol XI. Academic Press, New York, pp 104–190

Carroll RL (1988) Vertebrate paleontology evolution. Freeman, New York

Cederlund E, Peralba JM, Parés X, Jörnvall H (1991) Amphibian alcohol dehydrogenase, the major frog liver enzyme. Relationships to other forms and assessment of an early gene duplication separating vertebrate class I and class III alcohol dehydrogenases. Biochemistry 30:2811–2816

Chang C, Meyerowitz EM (1986) Molecular cloning and DNA sequence of the *Arabidopsis thaliana* alcohol dehydrogenase gene. Proc Natl Acad Sci USA 83:1408–1412

Chen CS, Yoshida A (1991) Enzymatic properties of the protein encoded by newly cloned human alcohol dehydrogenase *ADH6* gene. Biochem Biophys Res Commun 181:743–747

Crabb DW, Edenberg HJ (1987) Complete amino acid sequence of rat liver alcohol dehydrogenase deduced from the cDNA sequence. Gene 48:287–290

Creaser EH, Murali C, Britt KA (1990) Protein engineering of yeast alcohol dehydrogenases; effects of amino acid changes at positions 93 and 48 of yeast ADH1. Protein Eng 3:523–526

Dayhoff MO (1976) Survey of new data and computer methods of analysis. In: Dayhoff MO (ed) Atlas of protein sequence and structure, vol 5, suppl 2. National Biomedical Research Foundation, Washington DC, pp 1–8

Dayhoff MO, Schwartz RM, Orcutt BC (1978) A model for evolutionary change. In: Dayhoff MO (ed) Atlas of protein sequence and structure, vol 5, suppl 3. National Biomedical Research Foundation, Washington DC, pp 345–358

Dennis ES, Sachs MM, Gerlach WL, Finnegan ET, Peacock WJ (1985) Molecular analysis of the alcohol dehydrogenase 2 (*Adh2*) gene of maize. Nucleic Acids Res 13:727–743

Devereux J, Haeberli P, Smithies O (1984) A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res 12:387–395

Doolittle RF, Feng DF (1990) Nearest neighbor procedure for relating progressively aligned amino acid sequences. Methods Enzymol 183:659–669

Drum DE, Li TK, Vallee BL (1969) Considerations in evaluating the zinc content of horse liver alcohol dehydrogenase preparations. Biochemistry 8:3783–3797

Edenberg HJ, Zhang K, Fong K, Bosron WF, Li TK (1985) Cloning and sequencing of cDNA encoding the complete mouse liver alcohol dehydrogenase. Proc Natl Acad Sci USA 82: 2262–2266

Edenberg HJ, Brown CJ, Carr LG, Ho WH, Hur MW (1991) Alcohol dehydrogenase gene expression and cloning of the mouse χ-like alcohol dehydrogenase. Adv Exp Med Biol 284: 253–262

Ehrig T, Hurley TD, Edenberg HJ, Bosron WF (1991) General base catalysis in a glutamine for histidine mutant at position 51 of human liver alcohol dehydrogenase. Biochemistry 30: 1062–1068

Eklund H, Brändén CI (1987) Alcohol dehydrogenase. In: Jurnak FA, McPherson A (eds) Biological macromolecules and assemblies, vol 3. Active sites of enzymes. John Wiley, New York, pp 73–142

Eklund H, Plapp BV, Samama JP, Brändén CI (1982) Binding of substrate in a ternary complex of horse liver alcohol dehydrogenase. J Biol Chem 257:14349–14358

Eklund H, Samama JP, Jones TA (1984) Crystallographic investigations of nicotinamide adenine dinucleotide binding to horse liver alcohol dehydrogenase. Biochemistry 23:5982–5996

Eklund H, Horjales E, Jörnvall H, Brändén CI, Jeffery J (1985) Molecular aspects of functional differences between alcohol and sorbitol dehydrogenases. Biochemistry 24:8005–8012

Eklund H, Horjales E, Vallee BL, Jörnvall H (1987) Computer-graphics interpretations of residue exchanges between the α, β and γ subunits of human liver alcohol dehydrogenase class I isozymes. Eur J Biochem 167:185–193

Eklund H, Müller-Wille P, Horjales E, Futer O, Holmquist B, Vallee BL, Höög JO, Kaiser R, Jörnvall H (1990) Comparison of three classes of human liver alcohol dehydrogenase. Eur J Biochem 193:303–310

Ellison NW (1989) EMBL Data Library, S06200
Estonius M, Karlsson C, Fox EA, Höög JO, Holmquist B, Vallee BL, Davidson WS, Jörnvall H (1990) Avian alcohol dehydrogenase: the chicken liver enzyme. Eur J Biochem 194: 593–602
Fan F, Lorenzen JA, Plapp BV (1991) An aspartate residue in yeast alcohol dehydrogenase I determines the specificity for coenzyme. Biochemistry 30:6397–6401
Feng DF, Doolittle RF (1990) Progressive alignment and phylogenetic tree construction of protein sequences. Methods Enzymol 183:375–389
Ganzhorn AJ, Plapp BV (1988) Carboxyl groups near the active site zinc contribute to catalysis in yeast alcohol dehydrogenase. J Biol Chem 263:5446–5454
Good AG, Pelchen LE, Crosby WL (1988) Nucleotide sequence of a complete barley alcohol dehydrogenase I cDNA. Nucleic Acids Res 16:7182
Gwynne DI, Buxton FP, Sibley S, Davies RW, Lockington RA, Scazzocchio C, Sealy-Lewis HM (1987) Comparison of the cis-acting control regions of two coordinately controlled genes involved in ethanol utilization in *Aspergillus nidulans*. Gene 51:205–216
Ha BD, Buffardd D, Breda C, Esnault R (1989) EMBL Data Library, S06693
Holmes MA, Matthews BW (1981) Binding of hydroxamic acid inhibitors to crystalline thermolysin suggests a pentacoordinate zinc intermediate in catalysis. Biochemistry 20:6912–6920
Holmes RS, Courtney YR, VandeBerg JL (1986) Alcohol dehydrogenase in baboons: tissue distribution, catalytic properties, and variant phenotypes in liver, kidney, stomach, and testis. Alcoholism 10:623–630
Holmes RS, VandeBerg JL (1986) Ocular NAD-dependent alcohol dehydrogenase and aldehyde dehydrogenase in the baboon. Exp Eye Res 43:383–396
Höög JO, Hedén LO, Larsson K, Jörnvall H, von Bahr-Lindström H (1986) The γ_1 and γ_2 subunits of human liver alcohol deydrogenase. cDNA structures, two amino acid replacements, and compatibility with changes in the enzymatic properties. Eur J Biochem 159:215–218
Höög JO, von Bahr-Lindström H, Hedén LO, Holmquist B, Larsson K, Hempel J, Vallee BL, Jörnvall H (1987) Structure of the class II enzyme of human liver alcohol dehydrogenase: combined cDNA and protein sequence determination of the π subunit. Biochemistry 26:1926–1932
Hurley TD, Bosron WF, Hamilton JA, Amzel LM (1991) Structure of human $\beta_1\beta_1$ alcohol dehydrogenase: catalytic effects of non-active-site substitutions. Proc Natl Acad Sci USA 88: 8149–8153
Ikuta T, Fujiyoshi T, Kurachi K, Yoshida A (1985) Molecular cloning of a full-length cDNA for human alcohol dehydrogenase. Proc Natl Acad Sci USA 82:2703–2707
Ikuta T, Szeto S, Yoshida A (1986) Three human alcohol dehydrogenase subunits: cDNA structure and molecular evolutionary divergence. Proc Natl Acad Sci USA 83:634–638
Jeffery J, Chester J, Mills C, Sadler PJ, Jörnvall H (1984) Sorbitol dehydrogenase is a zinc enzyme. EMBO J 3:357–360
Jendrossek D, Steinbüchel A, Schlegel HG (1988) Alcohol dehydrogenase gene from *Alcaligenes eutrophus*: subcloning, heterologous expression in *E. coli*, sequencing, and location of Tn5 insertions. J Bacteriol 170:5248–5256
Johnson MS, Sutcliffe MJ, Blundell TL (1990) Molecular anatomy: phyletic relationships derived from three-dimensional structures of proteins. J Mol Evol 30:43–59
Jörnvall H (1970) Horse liver alcohol dehydrogenase. Eur J Biochem 16:25–40
Jörnvall H, Persson M, Jeffery J (1981) Alcohol and polyol dehydrogenases are both divided into two protein types, and structural properties cross-relate the different enzyme activities within each type. Proc Natl Acad Sci USA 78:4226–4230
Jörnvall H, Persson B, Jeffery J (1987) Characteristics of alcohol/polyol dehydrogenases. Eur J Biochem 167:195–201
Julià P, Parés X, Jörnvall H (1988) Rat liver alcohol dehydrogenase of class III. Eur J Biochem 172:73–83
Kadowaki KI, Matsuoka M, Murai N, Harada K (1988) Induction of two alcohol dehydrogenase polypeptides in rice roots during anaerobiosis. Plant Science 54:29–36
Kaiser R, Holmquist B, Hempel J, Vallee BL, Jörnvall H (1988) Class III human liver alcohol dehydrogenase: a novel structural type equidistantly related to the class I and class II enzymes. Biochemistry 27:1132–1140
Kaiser R, Holmquist B, Vallee BL, Jörnvall H (1989) Characteristics of mammalian class III alcohol dehydrogenases, an enzyme less variable than the traditional liver enzyme of class I. Biochemistry 28:8432–8438
Kaiser R, Nussrallah B, Dam R, Wagner FW, Jörnvall H (1990) Avian alcohol dehydrogenase. Characterization of the quail enzyme, functional interpretations, and relationships to the different classes of mammalian alcohol dehydrogenase. Biochemistry 29:8365–8371
Karlsson C, Jörnvall H, Höög JO (1991) Sorbitol dehydrogenase: cDNA coding for the rat enzyme. Eur J Biochem 198: 761–765
Kassam JP, Tang BK, Kadar D, Kalow W (1989) In vitro studies of human liver ADH variants using a variety of substrates. Drug Metab Dispos 17:567–573
Keshav KF, Yomano LP, An H, Ingram LO (1990) Cloning of the *Zymomonas mobilis* structural gene encoding alcohol dehydrogenase I (*adh*A): sequence comparison and expression in *E. coli*. J Bacteriol 172:2491–2497
Kim EE, Wyckoff HW (1991) Reaction mechanism of alkaline phosphatase based on crystal structures. Two-metal ion catalysis. J Mol Biol 218:449–464
Koivusalo M, Uotila L (1991) Glutathione-dependent formaldehyde dehydrogenase: evidence for the identity with class III alcohol dehydrogenase. Adv Exp Med Biol 284:241–251
Li TK, Bosron WF, Dafeldecker WP, Lange LG, Vallee BL (1977) Isolation of π-alcohol dehydrogenase of human liver: is it a determinant of alcoholism? Proc Natl Acad Sci USA 74:4378–4381
Light DR, Dennis MS, Forsythe IJ, Liu CC, Green DW, Kratzer DA, Plapp BV (1992) GenBank, M81807
Llewellyn DJ, Finnegan EJ, Ellis JG, Dennis ES, Peacock WJ (1987) Structure and expression of an alcohol dehydrogenase 1 gene from *Pisum sativum*. J Mol Biol 195:115–123
Matthews BW, Weaver LH, Kester WR (1974) The conformation of thermolysin. J Biol Chem 249:8030–8044
Matton DP, Brisson N (1990) Nucleotide sequence of two potato alcohol dehydrogenase cDNAs. Nucleic Acids Res 18: 3070
McKnight GL, Kato H, Upshall A, Parker MD, Saari G, O'Hara PJ (1985) Identification and molecular analysis of a third *Aspergillus nidulans* alcohol dehydrogenase gene. EMBO J 4: 2093–2099
Mitchell LE, Dennis ES, Peacock WJ (1989) Molecular analysis of an alcohol dehydrogenase gene from chromosome 1 of wheat. Genome 32:349–358
Monaco HL, Crawford JL, Lipscomb WN (1978) Three-dimensional structures of aspartate carbamoyltransferase from *E. coli* and of its complex with cytidine triphosphate. Proc Natl Acad Sci USA 75:5276–5280
Needleman SB, Wunsch CD (1970) A general method applicable to the search for similarities in the amino acid sequence of two proteins. J Mol Biol 48:443–453
Ohlsson I, Nordström B, Brändén CI (1974) Structural and functional similarities within the coenzyme binding domains of dehydrogenases. J Mol Biol 89:339–354
Parés X, Vallee BL (1981) New human liver alcohol dehydrogenase forms with unique kinetic characteristics. Biochem Biophys Res Commun 98:122–130

Park DH (1991) PhD Thesis, Structure and function of isoenzymes of horse liver alcohol dehydrogenase. The University of Iowa, Iowa City

Park DH, Plapp BV (1991) Isoenzymes of horse liver alcohol dehydrogenase active on ethanol and steroids. J Biol Chem 266:13296–13302

Peretz M, Burstein Y (1989) Amino acid sequence of alcohol dehydrogenase from the thermophilic bacteria *Thermoanaerobium brockii*. Biochemistry 28:6549–6555

Plapp BV, Ganzhorn AJ, Gould RM, Green DW, Jacobi T, Warth E, Kratzer DA (1991) Catalysis by yeast alcohol dehydrogenase. Adv Exp Med Biol 284:241–251

Rao PV, Krishna CM, Zigler JS (1992) Identification and characterization of the enzymatic activity of ζ-crystallin from guinea pig lens. J Biol Chem 267:96–102

Rees DC, Lewis M, Honzatko RB, Lipscomb WN, Hardman KD (1981) Zinc enviroment and cis peptide bonds in carboxypeptidase A at 1.75 Å resolution. Proc Natl Acad Sci USA 78:3408–3412

Ricard B, Mocquot B, Fournier A, Delseny M, Pradet A (1986) Expression of alcohol dehydrogenase in rice embryos under anoxia. Plant Mol Biol 7:321–329

Russell DW, Smith M, Williamson VM, Young ET (1983) Nucleotide sequence of the yeast alcohol dehydrogenase II gene. J Biol Chem 258:2674–2682

Russell PR, Hall BD (1983) The primary structure of the alcohol dehydrogenase gene from the fission yeast *Schizosaccharomyces pombe*. J Biol Chem 258:143–149

Saliola M, Shuster JR, Falcone C (1990) The alcohol dehydrogenase system in the yeast, *Kluyveromyces lactis*. Yeast 6:193–204

Smith M, Hopkinson DA, Harris H (1971) Development changes and polymorphism in human alcohol dehydrogenase. Ann Hum Genet 34:251–271

Trezise AEO, Godfrey EA, Holmes RS, Beacham IR (1989) Cloning and sequencing of cDNA encoding baboon liver alcohol dehydrogenase: evidence for a common ancestral lineage with the human alcohol dehydrogenase β subunit for class I alcohol dehydrogenase gene duplications predating primate radiation. Proc Natl Acad Sci USA 86:5454–5458

Trick M, Dennis ES, Edwards KJR, Peacock WJ (1988) Molecular analysis of the alcohol dehydrogenase gene family of barley. Plant Mol Biol 11:147–160

Vallee BL, Auld DS (1990) Active-site zinc ligands and activated H_2O of zinc enzymes. Proc Natl Acad Sci USA 87:220–224

Vallee BL, Bazzone TJ (1983) Isozymes of human liver alcohol dehydrogenase. In: Rattazzi MC, Scandalios JG, Witt GS (eds) Current topics in biological medical research, vol 8. Liss, New York, pp 219–244

von Bahr-Lindström H, Höög JO, Hedén LO, Kaiser R, Fleetwood L, Larsson K, Lake M, Holmquist B, Holmgren A, Hempel J, Vallee BL, Jörnvall H (1986) cDNA and protein structure for the α subunit of human liver alcohol dehydrogenase. Biochemistry 25:2465–2470

Wolyn DJ, Jelenkovic G (1990) Nucleotide sequence of an alcohol dehydrogenase gene in octoploid strawberry. Plant Mol Biol 14:855–857

Xie Y, Wu R (1989) Rice alcohol dehydrogenase genes: anaerobic induction, organ specific expression and characterization of cDNA clones. Plant Mol Biol 13:53–68

Xie Y, Wu R (1990) Molecular analysis of an alcohol dehydrogenase-encoding genomic clone (*adh2*) from rice. Gene 87:185–191

Yasunami M, Chen CS, Yoshida A (1990) Multiplication of the class I alcohol dehydrogenase locus in mammalian evolution. Biochem Genet 28:591–599

Yasunami M, Chen CS, Yoshida A (1991) A human alcohol dehydrogenase gene (ADH6) encoding an additional class of isozyme. Proc Natl Acad Sci USA 88:7610–7614

Yokoyama S, Yokoyama R, Kinlaw C, Harry DE (1990) Molecular evolution of the zinc-containing long-chain alcohol dehydrogenase genes. Mol Biol Evol 7:143–154

Young ET, Pilgrim D (1985) Isolation and DNA sequence of *ADH3*, a nuclear gene encoding the mitochondrial isozyme of alcohol dehydrogenase in *Saccharomyces cerevisiae*. Mol Cell Biol 5:3024–3034

Received November 7, 1991/Revised January 8, 1992