Biochemistry

 Volume 28, Number 15 July 25, 1989

Perspectives in Biochemistry

Eye Lens ζ-Crystallin Relationships to the Family of "Long-Chain" Alcohol/Polyol Dehydrogenases. Protein Trimming and Conservation of Stable Parts†

Teresa Borrás,‡ Bengt Persson,§ and Hans Jörnvall*,§

Department of Chemistry I, Karolinska Institutet, S-104 01 Stockholm, Sweden, and Laboratory of Mechanisms of Ocular Diseases, National Eye Institute, National Institutes of Health, Bethesda, Maryland 20892

Received March 21, 1989

ABSTRACT: ζ-Crystallin of guinea pig lens is distantly related to the family of zinc-containing alcohol/polyol dehydrogenases. The amino acid residues binding the catalytic zinc atom in the alcohol dehydrogenase are exchanged in ζ-crystallin, explaining lack of known enzyme activity, and those residues binding the noncatalytic zinc in the dehydrogenase are located in a segment absent from the crystallin. Mammalian alcohol dehydrogenase, polyol dehydrogenase, and ζ-crystallin therefore constitute a series of proteins exhibiting successive changes in subunit metal content, from two to one and probably zero zinc atoms, respectively. In common with tetrameric dehydrogenases, the crystallin lacks a loop structure present in the dimeric dehydrogenases. Significantly, the crystallin is tetrameric, and a correlation between extra subunit interactions and lack of the loop segment is indicated. The lacking segment in crystallin is extended, encompassing a second loop in the dehydrogenase. The greatest conservation corresponds to the coenzyme-binding domain of the dehydrogenases, the central parts of which are remarkably similar to those in the crystallin. Glycine is by far the most conserved residue and corresponds to positions at bends in the conformation of the alcohol dehydrogenase. The conservation of the stable parts of the fold, the absence of the loop structure, the lack of the metal atoms, and the presence of only a small proportion of oxidation-sensitive cysteine residues in crystallin (5 versus 15 in the β_1 dehydrogenase subunit) suggest an increased stability of the lens protein and a derivation from the alcohol dehydrogenase family. This is compatible with the recruitment of stable enzyme structures for lens crystallin functions, with trimming of protein structures through these dehydrogenases or a yet unknown enzyme, and with multiple changes in the dehydrogenase family.

Analysis of mammalian ζ-crystallin in comparison with zinc-containing dehydrogenases reveals similarities and variations that are of interest for evaluation of properties of protein structures. The results show remarkable correlations between conserved residues and conformational properties. The variations in metal binding, absence of known enzyme activity, extreme stability of basic folds, and trimming of a parent protein through losses/nonacquirements of structurally nonessential parts are outlined after an account of the crystallins and alcohol/polyol dehydrogenases.

ζ-Crystallin and Other Crystallins

Crystallins in General. Crystallins constitute structural proteins of the eye lens and are present at high concentration in this organ, which has very little resynthesis capacity. Stable conformations, giving long half-lives, are essential. Apparently, nature has acquired these proteins by repeated recruitments of globular protein structures for incorporation as crystallins in the lens (Wistow & Piatigorsky, 1987; Wistow et al., 1987). Some crystallins, such as α, β, and γ, are present in the lenses of all vertebrates; α-crystallin shows a distant relationship to small heat-shock proteins, while β- and γ-crystallins belong to another family. A different group of crystallins appear to be taxon-specific and frequently involve common enzymes of

† This work was supported by Grant 03X-3532 from the Swedish Medical Research Council.

* Address correspondence to this author.
‡ National Eye Institute, NIH.
§ Karolinska Institutet.

0006-2960/89/0428-6133$01.50/0

well-known metabolic pathways [review in Wistow and Piatigorsky (1987)]. Presumably, the ancient origins and widespread occurrence of these enzymes or derivatives of them resulted in stable conformations, which made them suitable for functions also in the lens.

Taxon-Specific Mammalian Crystallins and the Present Case. Recently, two taxon-specific crystallins were purified from mammalian lenses, λ-crystallin from rabbit (Mulders et al., 1988) and ζ-crystallin from guinea pig (Huang et al., 1987). They are related to hydroxyacyl coenzyme A dehydrogenase (Mulders et al., 1988) and alcohol dehydrogenase (Rodokanaki et al., 1989), respectively. The ζ-crystallin/alcohol dehydrogenase case illustrates an apparent conservation of stable parts as presently described. The consequences are of interest for evolution of protein structures in general.

Nonmammalian Crystallins. Characterized relationships between enzymes and lens proteins in nonmammals involve duck ε-crystallin, identical with a lactate dehydrogenase (Wistow et al., 1987), and the following crystallins more or less different[1] from the parent proteins: frog ρ-crystallin and aldose reductase or prostaglandin synthase (Carper et al., 1987; Watanabe et al., 1988), reptile/avian δ-crystallins and argininosuccinate lyase [cf. R. F. Doolittle as quoted in Wistow and Piatigorsky (1987); Piatigorsky et al. (1988)], turtle τ-crystallin and enolases (Wistow & Piatigorsky, 1987), or invertebrate S_{III} crystallin and cytosolic glutathione transferase (Siezen & Shaw, 1982; Tomarev & Zinovieva, 1988; Wistow & Piatigorsky, 1987). These and the presently discussed dehydrogenase relationships indicate that the recruitments of old protein structures to serve new functions have frequently occurred via gene duplication(s), accumulation of mutational differences, and acquisition of different genetic regulatory mechanisms resulting in altered tissue expression (Wistow & Piatigorsky, 1987).

Alcohol Dehydrogenase

Multiple Forms. The finding of ζ-crystallin as a distant relative of an alcohol dehydrogenase answers questions concerning which properties of a parent protein have been conserved in the new function and which have been altered. However, alcohol dehydrogenase, like crystallins, is not a single protein of simple origin. At least three separate types of alcohol dehydrogenase exist, with largely unrelated structures and mechanisms, the "long-chain" and "short-chain" alcohol dehydrogenases, with and without zinc at the active site, respectively (Jörnvall et al., 1981), and an iron-activated alcohol dehydrogenase (Scopes, 1983; Neale et al., 1986).

The Zinc-Enzyme, "Long-Chain" Type. The best known type, that of the zinc-containing, long-chain enzymes, is the one first characterized; it includes the common liver and yeast alcohol dehydrogenases and is the type related to ζ-crystallin. This type exhibits at least three levels of gene duplication (Jörnvall et al., 1987b). They have led to (1) different enzymes, alcohol and polyol dehydrogenases (Jörnvall et al., 1981; (2) different alcohol dehydrogenase classes, I, II, and III (Vallee & Bazzone, 1983); and (3) different intraclass isozymes [review in Jörnvall et al. (1987a)]. These separate levels of change in the zinc-enzyme type affect many properties, including metal content [different for polyol and alcohol dehydrogenases (Jeffery et al., 1984)], quaternary structure [also different (Jörnvall et al., 1987c)], enzymatic activities, and rates of evolutionary divergence [different among the alcohol dehydrogenase classes (Kaiser et al., 1989)].

Table I: ζ-Crystallin Relationships toward Each of the Zinc Dehydrogenase Lines Characterized[a]

line	residue identity with crystallin	spread
mammalian ADH	86 (27)	74–86 (23–27)
plant ADH	70 (22)	67–72 (21–23)
yeast ADH	67 (22)	60–68 (19–22)
mammalian SDH	64 (21)	64 (21)

[a] Column 2 gives the residue identity in number of positions and in percent (in parentheses) between crystallin and the alcohol dehydrogenase structures in Figure 1 (human β_1 in the case of the mammalian alcohol dehydrogenase line), while column 3 gives the values for crystallin compared with all characterized alcohol dehydrogenase structures within each line (only one alternative thus far reported for mammalian sorbitol dehydrogenase). ADH, alcohol dehydrogenase; SDH, sorbitol dehydrogenase.

Alignment

An alignment of the primary structure of ζ-crystallin with those of alcohol/polyol dehydrogenases is given in Figure 1. The analysis included 20 alcohol/polyol dehydrogenases [17 summarized in Jörnvall et al. (1987c), plus three class III structures (Kaiser et al., 1989)] and one ζ-crystallin structure (Rodokanaki et al., 1989), all compared in spans of variable size (Jörnvall et al., 1981). The zinc-containing dehydrogenases have been subgrouped into four lines: two dimeric (mammalian and plant alcohol dehydrogenase) and two tetrameric [yeast alcohol dehydrogenase and mammalian polyol dehydrogenase (Jörnvall et al., 1987c)]. The alignment given is based on maximal residue identities and minimal gap introductions and, in a manner similar to that for the dehydrogenase alignments (Jörnvall et al., 1978, 1987c; Jeffery et al., 1984; Brändén et al., 1984; Eklund et al., 1985, 1987), is compatible with the tertiary structure of horse liver alcohol dehydrogenase directly analyzed by X-ray crystallography [Eklund et al., 1976; recent summary in Eklund and Brändén (1987)]. At a few segments of the alignment, additional identities could have been obtained by further gap introductions, especially in the terminal parts where similarities are weak. Since variations within each dehydrogenase line are small (Table I), only representative structures discussed are shown. Figure 1 lists one dehydrogenase of each of the major lines (mammalian alcohol, plant alcohol, yeast alcohol, and mammalian polyol dehydrogenase), with the addition in the mammalian alcohol dehydrogenase line of alternatives to include all characterized classes of this enzyme line.

ζ-Crystallin exhibits the greatest overall similarity to the mammalian alcohol dehydrogenase line. Regarding the separate classes within this line, the extent of similarities does not differ much but is greatest with class I, intermediate with class III, and least with class II. Class I is the traditional and abundant liver type of the enzyme. The somewhat closer similarity of crystallin to this class than to the other two classes is noteworthy because the class I structure is the one for which most variation has been characterized thus far [about 3-fold more variable than class III (Kaiser et al., 1989)].

However, the crystallin structure also shows particular similarities with the structures of the tetrameric dehydrogenase lines (yeast alcohol and mammalian polyol dehydrogenases). Thus, many residues are unique to ζ-crystallin and these dehydrogenases (56 positions, versus 55 unique to the dimeric dehydrogenases and ζ-crystallin). Furthermore, a segment lacking in the crystallin has an identically positioned border with a segment lacking in yeast alcohol dehydrogenase and

[1] Some of these differences may turn out to represent close similarities or even identities although not yet established as such, since the crystallins and the enzymes have been characterized from separate species. Consequently, species variations and enzyme/crystallin variations are presently superimposed in the structures characterized.

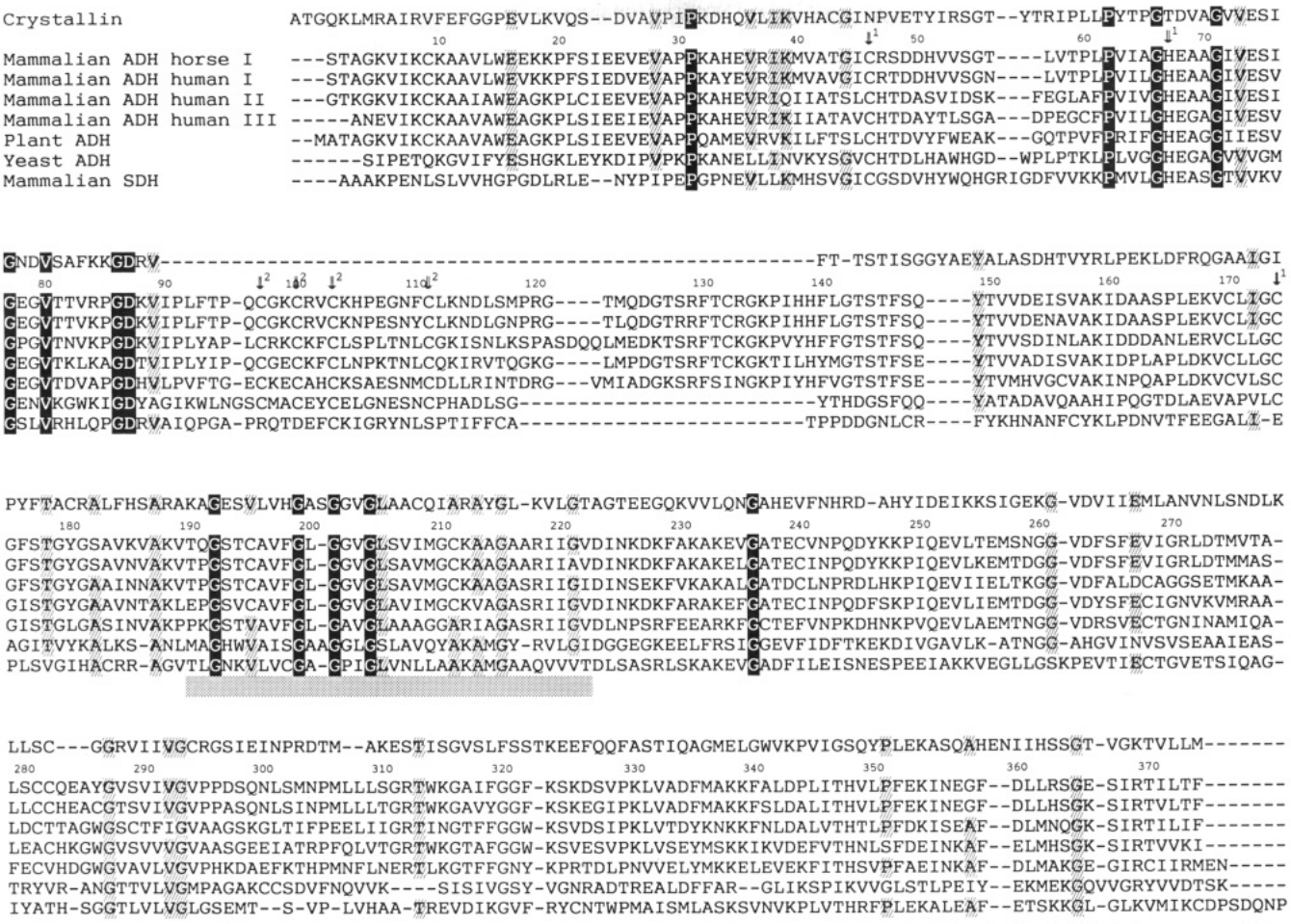

FIGURE 1: Structural comparison of ζ-crystallin with each of the four lines of zinc-containing long-chain alcohol/polyol dehydrogenases. The four evolutionary lines are from Jörnvall et al. (1987c), and in each case except the mammalian dehydrogenase line, the representative has been chosen that corresponds to the structure previously utilized for conformational comparisons, i.e., the horse E-type alcohol dehydrogenase subunit actually analyzed by X-ray crystallography (Eklund et al., 1976; Eklund & Brändén, 1987), the maize isozyme 1 structure (Brändén et al., 1984), the yeast (*Saccharomyces cerevisiae*) isozyme 1 structure (Jörnvall et al., 1978), and the sheep liver sorbitol dehydrogenase structure (Eklund et al., 1985). In addition, all three classes of human alcohol dehydrogenase have been included in the mammalian line (human β_1 subunit of class I, human π subunit of class II, and human χ subunit of class III) in order to allow judgment on class distinctions within the mammalian enzyme. Residues strictly conserved in all these lines are given against a black background, whereas residues conserved between the majority of the alternatives within at least three of the four dehydrogenase lines (column 4 in Table II) are given against a hatched background. Numbers given refer to the positions of the class I horse and human liver alcohol dehydrogenase subunits (Jörnvall et al., 1987c). ADH, alcohol dehydrogenase; SDH, sorbitol dehydrogenase. The stippling below the mammalian SDH line indicates the segment of maximal conservation and corresponds to the stippling in Figure 2B. Arrows marked 1 indicate the ligands to the active-site zinc atom, while arrows marked 2 indicate the ligands to the second zinc atom.

mammalian polyol dehydrogenase (Figure 1). Consequently, the overall residue identities with the mammalian alcohol dehydrogenase line and some of the particular similarities with the tetrameric lines emphasize different relationships. Both the similarities and differences can be further interpreted as discussed below.

Conformational and Functional Consequences

The consequences of the alignment in Figure 1 are of special interest in relation to the characteristic properties of alcohol dehydrogenases, i.e., overall conformation, quaternary structure, active site, and coenzyme-binding fold.

Overall Conformation. The crystallin is homologous with the alcohol dehydrogenases over extensive parts of the protein chains (60–86 residues are identical; Table I). Furthermore, the large gap segment in the crystallin overlaps a gap segment in two of the alcohol dehydrogenase lines (and coincides at one end; cf. Figure 1, top and bottom two lines around positions 90/120–140). Consequently, the crystallin structure is concluded to exhibit an overall conformation related to those of the alcohol dehydrogenases. The latter have previously been suggested to be related in tertiary structure (Brändén et al., 1984; Eklund et al., 1985, 1987; Jörnvall et al., 1978), and the extent of residue identity between the four separate alcohol dehydrogenase lines was similar to those now observed with the crystallin (Table I).

The type and positions of the residue identities are of significance. Thus, glycine is by far the residue most strictly conserved (Table II). Such a distribution is a property typical of distantly related proteins with similar overall conformations. Furthermore, the actual positions of these glycine residues in the tertiary structure of the alcohol dehydrogenase show that a large proportion of them are located at reverse turns or other bends in the enzyme conformation (Figure 2A). This fact supports a conservation of the general characteristics of the dehydrogenase tertiary structures in that of the crystallin, with largely similar folding patterns. This conclusion applies to overall properties only, and localized deviations could well occur as detected below.

Deviating Patterns, Quaternary Structure. The overall similarity between crystallin and the mammalian alcohol dehydrogenase line (Table I) does not apply to the crystallin gap

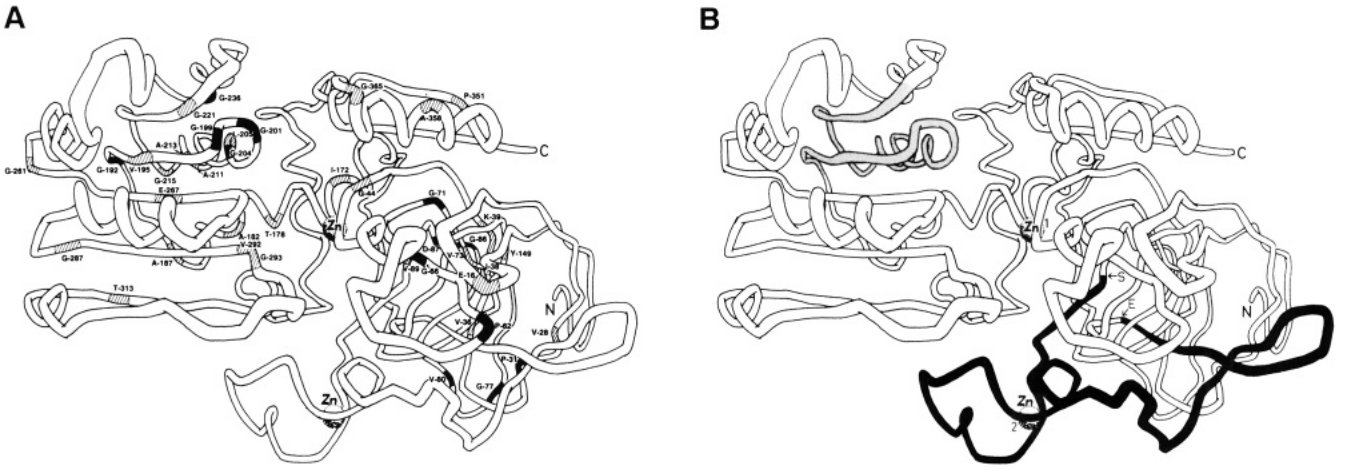

FIGURE 2: Characteristics of the crystallin structure in relation to the crystallographically analyzed (Eklund et al., 1976; Eklund & Brändén, 1987) horse liver alcohol dehydrogenase structure. (A) Positions of conserved residues. Residues strictly conserved (corresponding to column 2 in Table II) are shown black, and those conserved between the majority of the alternatives within at least three of the four dehydrogenase lines (column 4 in Table II) are hatched. The corresponding residues are in all cases also indicated by the one-letter code and the position in the horse enzyme. Data for the conserved residues are taken from Figure 1, while the conformation of the model enzyme is from Eklund et al. (1976) and Eklund and Brändén (1987). (B) Regions with the most conserved and variable segments. The long black region constitutes the gap segment in the crystallin of Figure 1, showing it to correspond exactly to surface loops of the dehydrogenase catalytic domain, including the structural zinc atom. A shortcut of the protein chain at the arrows marked S (for gap start) and E (for gap end) illustrates a possibility for direct continuity without large alterations of remaining chain folds. The stippled segment corresponds to the most conserved region shown in Table III at the center of the six β-pleated sheet strands of the stable conformation of the coenzyme-binding domain. The catalytic zinc atom is numbered 1 and the second (structural) zinc atom 2.

Table II: Conserved Residues between ζ-Crystallin and the Four Dehydrogenase Lines[a]

		conserved between crystallin and over half of all DHs in	
residue	strictly conserved	each of the four DH lines	at least three of the four DH lines
Asp	1	1	1
Thr			2
Glu			2
Pro	2	2	3
Gly	9	13	16
Ala		1	5
Val	1	3	7
Ile			2
Leu			1
Tyr			1
Lys			1
sum	13	20	41

[a] Residues in columns 2 and 4 are black and hatched, respectively, in Figure 2A. DH, dehydrogenase.

region covering positions 90–139 of the dehydrogenase. Instead, the crystallin gap segment coincides with a similarly positioned but smaller gap in the yeast alcohol and mammalian sorbitol dehydrogenases (Figure 1). The latter gap has previously been shown to correlate with the absence of one loop at a surface of the catalytic domain (Eklund et al., 1985; Jörnvall et al., 1978) and has been suggested as a possible reason (because of exposure of a novel subunit surface) why the two dehydrogenase lines lacking this loop represent tetrameric enzymes, while the two enzyme lines having this loop (mammals, plants) give dimeric enzymes. Significantly, the ζ-crystallin quaternary structure is also tetrameric (Huang et al., 1987; Rodokanaki et al., 1989), further supporting a correlation between this surface loop and dimer–dimer interactions.

In addition, the segment lacking from ζ-crystallin is considerably extended toward the N-terminal part, compared with the segment lacking from the yeast alcohol and mammalian sorbitol dehydrogenase lines (Figure 1). However, this extension corresponds to positions 90–118 in the mammalian alcohol dehydrogenase structure (Figure 1), which exactly constitutes a second loop, adjacent to the first and also located at the surface of the catalytic domain in the dehydrogenase structure (Figure 2B). Consequently, the extra segment lacking from the crystallin structure, though increasing the difference between crystallin and dehydrogenases, still confines the variations to the same surface of the subunit. The extended gap segment therefore supports the conclusion that this surface area may be involved in determining the quaternary structure and does not violate the conclusion about overall related conformations. In fact, the protein overall conformation can be obtained even with a "shortcut" between adjacent chain parts at the arrows indicated in Figure 2B, accounting for the absence of the two loops without large effects on remaining parts. As a consequence, the second zinc atom (the structural zinc atom), characteristic of the mammalian alcohol dehydrogenases, is absent in ζ-crystallin, exactly as in sorbitol dehydrogenase, where the lack has been shown by actual zinc analysis (Jeffery et al., 1984).

Active Site, Catalytic Properties. The three residues binding the catalytic zinc atom in all the dehydrogenase structures (Eklund et al., 1976, 1985, 1987; Jörnvall et al., 1978, 1987c; Jeffery et al., 1984; Brändén et al., 1984; Eklund & Brändén, 1987; Kaiser et al., 1989) are all absent in the ζ-crystallin structure (Figure 1). Thus, Cys-46 is replaced by Asn, His-67 by Thr, and Cys-174 in the alcohol dehydrogenases (Glu-154 in the sorbitol dehydrogenase) by Ile. None of these crystallin residues is typical of zinc-binding residues in proteins. Consequently, it can be concluded that ζ-crystallin lacks the catalytic zinc atom, thereby explaining the absence of alcohol/sorbitol dehydrogenase activity, which is in agreement with the direct enzymatic measurements (Zigler, unpublished results). Importantly, this change appears to be absolute. Thus, there is no Cys, His, or Glu residue that could easily replace the residues lost (Figure 1). Furthermore, the segment around the His-67 zinc-binding residue in the dehydrogenases, although being the most conserved region among the variable

Table III: Residue Identities between ζ-Crystallin and the Dehydrogenases for the Central Part of the Dehydrogenase Coenzyme-Binding Domain and the Whole Domain in Relation to Overall Protein Identities[a]

region	residue identity (%)
dehydrogenase positions 190–222	34 (31–52)
dehydrogenase coenzyme-binding domain	29 (19–29)
whole protein chains	27 (19–27)

[a] Values shown relate to the crystallin/human β_1 alcohol dehydrogenase subunit similarity and, within parentheses, those for the comparison of crystallin and any of the dehydrogenases.

dehydrogenases and the only segment where they have three consecutive residues strictly conserved (Jörnvall et al., 1987c), is also largely exchanged in the crystallin structure (having only Gly-66 conserved; Figure 1).

Strict Conservation, Coenzyme-Binding Fold. Despite the many exchanges between the crystallin and dehydrogenase structures (Table I), one segment is especially well conserved, the one corresponding to positions 190–222 in the dehydrogenase structure (Figure 1). This constitutes the center of the coenzyme-binding fold and encompasses central strands of the six-strand β-pleated sheet structure forming the backbone of the coenzyme-binding domain of the dehydrogenases (Eklund et al., 1976; Eklund & Brändén, 1987). The maximally conserved strands are highlighted in Figures 1 and 2B. This 33-residue segment is far more conserved than the whole protein chain and has many alcohol dehydrogenase typical residues, covering the βA–αB–βB center of the coenzyme-binding fold [Figure 2B; cf. Eklund et al. (1976) and Eklund and Brändén (1987)]. Consequently, the most characteristic dehydrogenase structure, constituting an evolutionarily old and heavily hydrogen-bonded structure (Rossman et al., 1978), is extensively conserved in the crystallin structure. Furthermore, this conservation is large (Table III), independent of whether the crystallin is compared to one or the other of the four dehydrogenase lines, and is largest (36–52%) toward the yeast enzyme, whereas overall residue similarities are most clearly visible toward the mammalian alcohol dehydrogenase (Table I).

Consequences of the Relationships

The comparisons clearly show that mammalian ζ-crystallin and alcohol/polyol dehydrogenases are structurally related. The homology extends over the whole protein chains (Figure 1); the most conserved residues are glycine residues (Table I), largely confined to bends in the tertiary structure (Figure 2A). The single long segment lacking in the crystallin corresponds to surface loops and is compatible with an otherwise largely unchanged chain fold through a polypeptide "shortcut" (Figure 2B). The absence of the structural zinc atom coincides with lack of this segment and the absence of the catalytic zinc atom with the lack of known enzyme activity. Finally, the most extensively conserved structure is represented by a central segment in one domain, corresponding to stable secondary structures with β-strands (Figure 2B). Naturally, the overall strict agreements do not exclude further localized dissimilarities than those regarding the zinc atoms and the surface loop. Nevertheless, the overall agreement establishes the common origin through an ancestral gene duplication and places ζ-crystallin as a deviating member of the alcohol/polyol long-chain dehydrogenase superfamily (Jörnvall et al., 1981, 1987c). The inclusion of the crystallin in this protein group has consequences for conclusions on both the crystallin structure and the dehydrogenase structures.

Consequences for the Crystallin Protein. The fact that ζ-crystallin is related to alcohol dehydrogenase strongly supports the recruitment theory (Wistow & Piatigorsky, 1987; Wistow et al., 1987) for the taxon-specific crystallins of the eye lens. These recruitments have occurred repeatedly, explaining why different crystallin forms are taxon-specific and showing separate relationships between crystallin forms and lactate dehydrogenase, argininosuccinate lyase, enolase, glutathione transferase, prostaglandin synthase, aldose reductase, hydroxyacyl coenzyme A dehydrogenase, and alcohol dehydrogenase [Carper et al., 1987; Doolittle as quoted in Wistow and Piatigorsky (1987); Huang et al., 1987; Mulders et al., 1988; Piatigorsky et al., 1988; Rodokanaki et al., 1989; Siezen & Shaw, 1982; Tomarev & Zinovieva, 1988; Watanabe et al., 1988; Wistow et al., 1987; Wistow & Piatigorsky, 1987]. In addition, the present case illustrates that the recruitment apparently involves conservation of the most stable structural characteristics, but lack of structures little stabilized. Thus, in the case of the alcohol dehydrogenase/ζ-crystallin pair, the crystallin has conserved the hydrogen-bonded most central part of the coenzyme-binding domain that constitutes a very old and common region of many dehydrogenase and kinase structures. Similarly, the crystallin structure lacks the two zinc atoms per alcohol dehydrogenase subunit and six zinc-binding cysteine residues, which, when present, are more or less reactive (Johansson et al., unpublished results) and hence sensitive to oxidation or other reactions. The crystallin structure has only 5 thiols per subunit, versus 15 in human alcohol dehydrogenase. Finally, the loop segment absent in crystallin corresponds to the only segment in the entire alcohol dehydrogenase structure that lacks stabilization by an extensive secondary structure (Eklund et al., 1976; Eklund & Brändén, 1987). In summary, the crystallin structure has all properties of a stable conformation with much secondary structure and with few sensitive residues, as expected from efficient utilization of a gene product recruited from a metabolic enzyme structure.

Regarding the question from which line the recruitment occurred, different possibilities seem to exist. Since ζ-crystallin is quite different from alcohol/sorbitol dehydrogenase, one possibility is that it represents recruitment from yet another line, hitherto unknown, of this enzyme superfamily. If so, that enzyme could already have lost/never acquired the zinc atoms and have a different enzymatic mechanism and metabolic role. No such enzyme is known and its possible activity or existence is hypothetical. However, it should be noted that a novel protein similar to ζ-crystallin has been detected in liver from several mammalian species (Zigler & Du, 1989). Also, the time of guinea pig taxon-specific divergence appears short in relation to the large differences observed. Consequently, a recruitment from such a hypothetical enzyme line is possible, provided the demonstration of the novel protein in liver does not reflect an ectopic expression of the ζ-crystallin gene.

The other alternative is that ζ-crystallin has been recruited from the alcohol/sorbitol dehydrogenase line. If so, the properties of the crystallin are noteworthy. Thus, the ζ-crystallin shows partly different similarities. The lacking segment and one set of unique residue identities suggest relationships with the tetrameric alcohol/polyol dehydrogenases, while overall residue identities and another set of positions suggest relationships with the dimeric enzymes. This mixed pattern appears to indicate the existence of both divergence and convergence in the evolution of ζ-crystallin: divergence from an original line and convergence toward the other(s), presumably because of structural requirements. Since the ζ-crystallin is hystricomorphic and guinea pig proteins may deviate to a considerable extent from those of other mam-

malian forms [cf. Persson et al. (1989) for another enzyme], it would be of interest to know the structures of guinea pig alcohol and sorbitol dehydrogenases. However, guinea pig alcohol dehydrogenase has been purified and its properties, including the presence of different enzyme classes, resemble those of other mammalian alcohol dehydrogenases (Keung & Fong, 1988). Consequently, guinea pig alcohol dehydrogenase has not evolved differently than other alcohol dehydrogenases and most likely does not constitute the novel hypothetical enzyme eluded to above. Thus, the two possibilities remain, both equally unexpected, demonstrating recruitment involving either an exceptionally large trimming of a parent alcohol/sorbitol dehydrogenase structure or the existence of a hitherto uncharacterized liver enzyme with an unknown substrate.

Consequences for Alcohol Dehydrogenase Relationships. Three levels of duplicatory events have been traced in the evolution of the zinc-containing long-chain alcohol dehydrogenases (Jörnvall et al., 1987b,c). At the first level, the alcohol dehydrogenase and polyol dehydrogenase lines separated; these two types of structure now exhibit overall residue identities of only about 25% (Jörnvall et al., 1987b). At the most recent level, the separate isozymes of mammalian alcohol dehydrogenases became apparent, presently exhibiting residue identities of more than 90%. The intermediate level is the one giving rise to the three classes of mammalian zinc-containing alcohol dehydrogenases. These three classes constitute the abundant, classical liver-type (class I) enzyme and the separate alcohol dehydrogenases of classes II and III, which differ in structure, substrate specificity, and organ distribution (Vallee & Bazzone, 1983; Jörnvall et al., 1987b,c). Species variations within the classes suggest that class III is highly constant in structure, whereas class I exhibits a variation about 3-fold more divergent (Kaiser et al., 1989). Furthermore, the three classes differ at the active site, the area for subunit interactions, and in the segment around the second zinc atom, suggesting that they constitute separate enzymes rather than simple isozymes (Jörnvall et al., 1987b). Noticeably, the same types of variation, i.e., variations affecting areas of subunit interactions and the loop around the second zinc atom, are exactly those now found to affect the crystallin structure in relation to those of the alcohol dehydrogenases. Thus, the crystallin shows that the interclass dehydrogenase variability is not an isolated phenomenon for just the alcohol dehydrogenases, but affects similarly other members of this protein superfamily.

Furthermore, the crystallin relationship can illustrate the origin of the separate classes. This question is important, because the three classes differ in substrate specificity (Vallee & Bazzone, 1983) and neither class has a clearly defined, known metabolic role. It is now apparent that the crystallin relationship starts to show a distinction between class II and the other classes of alcohol dehydrogenase (23% residue identity toward class II versus 26–27% toward class I and 25–26% toward class III). Thus, class II appears to be the class least similar to crystallin. It is the only class for which species variations have thus far not been characterized and for which the extent of divergence is therefore unknown. In contrast, classes I and III (Kaiser et al., 1989) reveal different extents of changes and separate properties. Separate properties are now also extended to class II, showing its relation to crystallin not to resemble those of either of the other two classes.

Finally, the successive changes in metal content within the alcohol dehydrogenase/polyol dehydrogenase/crystallin family are noteworthy. Thus, the corresponding subunits exhibit two zinc atoms, one zinc atom, and zero zinc atoms, respectively, as judged from previous analyses of the dehydrogenases (Jeffery et al., 1984) and the present comparisons for the crystallin. The successive differences in metal content were noteworthy already when only the alcohol and polyol dehydrogenases were characterized, but are now even more apparent, showing a lineage of related proteins with three different stages regarding metal content.

Conclusions

The following properties are of particular interest for conclusions on protein structures in general.

(1) The relationships add the "long-chain" zinc-containing dehydrogenase family to the group of protein families that contain highly divergent members involving both enzymes and other proteins. The early cases of lysozyme/α-lactalbumin (Hill & Brew, 1975) and serine proteases/haptoglobin (Kurosky et al., 1974), together with recent additions of, for example, lipases/vitellogenins (Bownes et al., 1988; Persson et al., 1989) and enzymes/crystallins (Wistow & Piatigorsky, 1987), with the present case in particular, show that many protein families are large with wide relationships across conventional borders. Apparently, protein structures exhibit limited variability. The crystallin member now detailed of the "long-chain, zinc-containing dehydrogenase" family has neither the zinc ligands (hence not the zinc) nor the same substrates if at all any enzyme activity. Successive variations in metalloprotein metal content are well illustrated by alcohol dehydrogenase/sorbitol dehydrogenase/ζ-crystallin. The lack of known enzyme activity can be ascribed to the absence of single residues (here active-site zinc ligands) in the same manner as for the loss of Glu-35 in lysozyme, Ser-195 in chymotrypsin, and a putative functional serine in lipases regarding the other examples of enzymes/nonenzymes mentioned above. Thus, small protein changes can have large effects on important properties at the same time as extreme divergence does apparently not alter basic folds.

(2) The most conserved segments are central regions of an evolutionarily old domain, particularly involving glycine residues at critical positions, as shown early for cytochrome *c* (Smith & Margoliash, 1964) and known within the alcohol/polyol dehydrogenases themselves (Jörnvall et al., 1984). Apparently, such structures are stable, making them suitable also for recruitments as structural proteins.

(3) Natural protein trimming occurs and suggests the removal (if ζ-crystallin is derived from alcohol dehydrogenase) or nonacquirement (if derived from a novel, hitherto unknown protein already having these characteristics) of unstable parts, such as oxidation-sensitive cysteine residues, superficial segments, and loops nonstabilized by elements of secondary structure. Obviously, observations on the naturally occurring variants complement those possible to obtain from directed mutagenesis. Native variability exhibits successive changes, and as presently shown, extensive differences between enzymes and structural proteins can be correlated with particular residues or segments in critical regions. Combined, observations on both natural variants and directed mutants facilitate delineation of basic folds. Regarding differences in metal content, quaternary structure, and biological activity, the crystallin/alcohol dehydrogenase/polyol dehydrogenase protein family illustrates exceptionally well several types of relationship, emphasizing stability and multiplicity.

Added in Proof

A report has recently appeared, extending the family of long-chain zinc-containing dehydrogenases to include also prokaryotic threonine dehydrogenase (Aronson et al., 1989).

One further line can therefore be added to Figure 1 and Table I, and the family can now be summarized as alcohol/polyol/threonine dehydrogenases/ζ-crystallin.

ACKNOWLEDGMENTS

We are grateful to Carina Palmberg and Eva Andersson for figures and to Bo Furugren for the original drawing of the schematic chain fold of the horse enzyme crystallographically analyzed by Carl-Ivar Brändén, Hans Eklund, and co-workers.

Registry No. ADH, 9031-72-5; SDH, 9028-21-1; Zn, 7440-66-6; polyol dehydrogenase, 80448-98-2.

REFERENCES

Aronson, B. D., Somerville, R. L., Epperly, B. R., & Dekker, E. E. (1989) *J. Biol. Chem. 264*, 5226–5232.

Bownes, M., Shirras, A., Blair, M., Collins, J., & Coulson, A. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1554–1557.

Brändén, C.-I., Eklund, H., Cambillau, C., & Pryor, A. J. (1984) *EMBO J. 3*, 1307–1310.

Carper, D., Nishimura, C., Shinohara, T., Dietzchold, B., Wistow, G., Craft, C., Kador, P., & Kinoshita, J. H. (1987) *FEBS Lett. 220*, 209–213.

Eklund, H., & Brändén, C.-I. (1987) in *Biological Macromolecules and Assemblies* (Jurnak, F., & McPherson, A., Eds.) Vol. 3, pp 74–142, Wiley, New York.

Eklund, H., Nordström, B., Zeppezauer, E., Söderlund, G., Ohlsson, I., Boiwe, T., Söderberg, B.-O., Tapia, O., Brändén, C.-I., & Åkeson, Å. (1976) *J. Mol. Biol. 102*, 27–59.

Eklund, H., Horjales, E., Jörnvall, H., Brändén, C.-I., & Jeffery, J. (1985) *Biochemistry 24*, 8005–8012.

Eklund, H., Horjales, E., Vallee, B. L., & Jörnvall, H. (1987) *Eur. J. Biochem. 167*, 185–193.

Hill, R. L., & Brew, K. (1975) *Adv. Enzymol. Relat. Areas Mol. Biol. 43*, 411–490.

Huang, Q.-L., Russell, P., Stone, S. H., & Zigler, J. S., Jr. (1987) *Curr. Eye Res. 6*, 725–732.

Jeffery, J., Chester, J., Mills, C., Sadler, P. J., & Jörnvall, H. (1984) *EMBO J. 3*, 357–360.

Jörnvall, H., Eklund, H., & Brändén, C.-I. (1978) *J. Biol. Chem. 253*, 8414–8419.

Jörnvall, H., Persson, M., & Jeffery, J. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 4226–4230.

Jörnvall, H., von Bahr-Lindström, H., & Jeffery, J. (1984) *Eur. J. Biochem. 140*, 17–23.

Jörnvall, H., Hempel, J., & Vallee, B. L. (1987a) *Enzyme 37*, 5–18.

Jörnvall, H., Höög, J.-O., von Bahr-Lindström, H., & Vallee, B. L. (1987b) *Proc. Natl. Acad. Sci. U.S.A. 84*, 2580–2584.

Jörnvall, H., Persson, B., & Jeffery, J. (1987c) *Eur. J. Biochem. 167*, 195–201.

Kaiser, R., Holmquist, B., Vallee, B. L., & Jörnvall, H. (1989) *Biochemistry* (in press).

Keung, W.-M., & Fong, W.-P. (1988) *Comp. Biochem. Physiol. 89B*, 85–89.

Kurosky, A., Barnett, D. R., Rasco, M. A., Lee, T.-H., & Bowman, B. H. (1974) *Biochem. Genet. 11*, 279–293.

Mulders, J. W. M., Hendriks, W., Blankesteijn, W. H., Bloemendal, H., & de Jong, W. W. (1988) *J. Biol. Chem. 263*, 15462–15466.

Neale, A. D., Scopes, R. K., Kelley, J. M., & Wettenhall, R. E. H. (1986) *Eur. J. Biochem. 154*, 119–124.

Persson, B., Bengtsson-Olivecrona, G., Enerbäck, S., Olivecrona, T., & Jörnvall, H. (1989) *Eur. J. Biochem. 179*, 39–45.

Piatigorsky, J., O'Brien, W. E., Norman, B. L., Kalumuck, K., Wistow, G. J., Borrás, T., Nickerson, J. M., & Wawrousek, E. F. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 3479–3483.

Rodokanaki, A., Holmes, R., & Borrás, T. (1989) *Gene 78*, 215–224.

Rossmann, M. G., Liljas, A., Brändén, C.-I., & Banaszak, L. J. (1978) *Enzymes (3rd Ed.) 11*, 61–102.

Scopes, R. K. (1983) *FEBS Lett. 156*, 303–306.

Siezen, R. J., & Shaw, D. C. (1982) *Biochim. Biophys. Acta 704*, 304–320.

Smith, E. L., & Margoliash, E. (1964) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 23*, 1243–1247.

Tomarev, S. I., & Zinovieva, R. D. (1988) *Nature 336*, 86–88.

Vallee, B. L., & Bazzone, T. J. (1983) *Isozymes: Curr. Top. Biol. Med. Res. 8*, 219–244.

Watanabe, K., Fujii, Y., Nakayama, K., Ohkubo, H., Kuramitsu, S., Kagamiyama, H., Nakanishi, S., & Hayaishi, O. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 11–15.

Wistow, G., & Piatigorsky, J. (1987) *Science 236*, 1554–1556.

Wistow, G. J., Mulders, J. W. M., & de Jong, W. W. (1987) *Nature 326*, 622–624.

Zigler, J. S., Jr., & Du, X.-y. (1989) *Invest. Ophthalmol. Visual Sci. 30 (Suppl.)*, 266.