Crystal Structure of η-Crystallin: Adaptation of a Class 1 Aldehyde Dehydrogenase for a New Role in the Eye Lens†

O. A. Bateman,‡,# A. G. Purkiss,‡,# R. van Montfort,‡,|| C. Slingsby,*,‡ C. Graham,§,⊥ and G. Wistow§

Birkbeck College, School of Crystallography, Malet Street, London WC1E 7HX, UK, Section on Molecular Structure and Function, National Eye Institute, Room 331, Building 6, National Institutes of Health, Bethesda, Maryland 20892-2740, and NCI/DBS/EIB, National Institutes of Health, Bethesda Maryland

Received December 17, 2002; Revised Manuscript Received February 20, 2003

ABSTRACT: η-Crystallin is a retinal dehydrogenase that has acquired a role as a structural protein in the eye lens of elephant shrews, members of an ancient order of mammals. While it retains some activity toward retinal, which is oxidized to retinoic acid, the protein has acquired a number of specific sequence changes that have presumably been selected to enhance the lens role. The crystal structure of η-crystallin, in common with class 1 and 2 ALDHs, is a dimer of dimers. It has a better-defined NAD binding site than those of related mammalian ALDH1 enzymes with the cofactor bound in the "hydride transfer" position in all four monomers with small differences about the dimer dyads. Although the active site is well conserved, the substrate-binding site is larger in η-crystallin, and there are some mutations to the substrate access tunnel that might affect binding or release of substrate and product. It is possible that η-crystallin has lost flexibility to improve its role in the lens. Enhanced binding of cofactor could enable it to act as a UV/blue light filter in the lens, improving visual acuity. The structure not only gives a view of a "natural mutant" of ALDH1 illustrating the adaptive conflict that can arise in multifunctional proteins, but also provides a well-ordered NAD binding site structure for this class of enzymes with important roles in development and health.

Aldehydes are key intermediates in a wide range of cellular biochemical pathways. Because of their high reactivity aldehydes, whether generated by endogenous oxidative reactions or from exogenous sources, are also potential toxins. An array of enzymes has evolved to catabolize aldehydes, either in biosynthetic reactions or in detoxification. Among these enzymes are the aldehyde dehydrogenases (EC 1.2.1.3) that catalyze oxidation of an aldehyde to a carboxylate in an NAD^+-dependent process. Their reaction mechanism involves binding of NAD^+, binding of the substrate aldehyde and formation of a thiohemiacetal intermediate with an active site cysteine. This intermediate then collapses to a thioester, followed by the stereospecific donation of a hydride ion to nicotinamide, thioester hydrolysis activated by a glutamate acting as a general base, release of the product acid and finally release of NADH (*1*).

Vertebrates have as many as 20 families of aldehyde dehydrogenases (ALDHs)[1] (*2*). Structural information is available for homotetrameric mammalian class 1 and class 2 enzymes of 220 kDa (4 × 55 kDa) and dimeric class 3 enzymes (2 × 50 kDa) (*3*). Cytosolic class 1 and mitochondrial class 2 liver enzymes oxidize retinaldehyde and acetaldehyde, respectively. Each monomer comprises an N-terminal βαβ nucleotide-binding domain (NBD) followed by a βαβ catalytic domain and a C-terminal small β-sheet oligomerization domain. Dimer formation involves contacts between monomer α-helices and β-sheet extension between the catalytic and oligomerization domains (*4*, *5*). NMR and fluorescence spectroscopy have shown that the nicotinamide portion samples a population of conformations, while the adenosine is less mobile (*6*), leading to the suggestion that coenzyme conformational mobility is beneficial to catalysis throughout the reaction cycle. X-ray crystallography has defined two conformations for the NAD (*5*, *7*, *8*); however, in one conformation it would prevent Glu 268 functioning as the general base, while in the other it is too far from the active site for efficient hydride transfer (*6*). In the tetrameric enzymes, the active sites lie at the end of a deep hydrophobic tunnel (*5*, *7*, *8*). The main difference between class 1 and class 2 enzymes is seen in the substrate entrance tunnel: the class 2 enzymes have a smaller substrate entrance tunnel, which correlates with their smaller acetaldehyde substrate, compared to the class 1 enzymes that oxidize retinaldehyde (*7*). Class 2 enzymes display half-of-the-sites reactivity, behavior considered as an extreme example of negative cooperativity. Although the structural details of this allosteric mechanism have not been fully defined, it thought to involve the α-helix G at the dimer interface (*9*, *10*).

† This work was funded by the Medical Research Council (London) G8303198 to CS.

* Corresponding author. E-mail: c.slingsby@bbk.ac.uk.
# These authors contributed equally.
‡ Birkbeck College.
§ National Institutes of Health.
⊥ NCI/DBS/EIB, NIH.
|| Present address: Astex Technology Ltd., 250 Cambridge Science Park, Milton Road, Cambridge CB4 0WE, UK.

[1] Abbreviations: aldehyde dehydrogenase (ALDH), cellular retinol-binding protein (CRBP), dithiothreitol (DTT), ethylenediaminetetraacetic acid (EDTA), nicotinamide adenine dinucleotide oxidized (NAD), nicotinamide adenine dinucleotide reduced (NADH), nucleotide-binding domain (NBD), nuclear magnetic resonance (NMR), Protein Data Bank (PDB), ultraviolet (UV).

10.1021/bi027367w CCC: $25.00
Published on Web 03/28/2003

Retinal is both the precursor of the potent morphogen and bioregulator, retinoic acid and also the essential ligand for light reception in the opsins (*11*). ALDHs that metabolize retinal are localized in many different tissues. While liver ALDHs are involved in oxidation of metabolic retinal (*12*, *13*), ALDHs also play a key role in embryonic development by controlling levels of retinoic acid. These enzymes show developmentally regulated expression in several tissues, including the eye, and play a key role in controlling the levels of retinoic acid (*14*), which is also a candidate regulator of eye growth in response to refractive errors (*15*). The two class 1 enzymes, ALDH1 and RALDH2, play essential but distinct roles in retinoic acid signaling (*2*, *16*) along with the spatially distinct ALDH6 (*17*). Understanding the specific structural features that confer the various functions of these isoenyzmes is important because of their impact on normal development and health.

In addition to their crucial roles in metabolism, ALDHs also belong to the group of proteins that have been recruited as structural eye lens proteins in both vertebrates and invertebrates. One of the taxon-specific enzyme crystallins in mammals is η-crystallin, which accounts for up to a quarter of total lens protein in elephant shrews (*18*, *19*), a group of diurnal, or partly diurnal, insectivores that belong to the ancient clade of the Afrotheria (*20*). η-Crystallin is a class 1 ALDH, but it is present in the lens at levels that far exceed any likely enzymatic requirement.

During evolution, the composition of the vertebrate eye lens has been re-engineered to adapt to different visual environments. This has been achieved either by reducing or eliminating the expression of some genes, such as the γ-crystallins, or by the process known as gene recruitment whereby a preexisting enzyme gene acquires greatly enhanced expression, tissue-specifically in the lens. The enzyme then becomes a bulk component of the lens, where it is generally associated with softening of the lens and provides the ability to bind and sequester chromophores that may act as protective filters in the lens (*21*). Gene recruitment seems to begin with a single protein acquiring multiple functions (*22*) causing an "adaptive conflict" in which changes beneficial for one function may be deleterious for the other (*23*). This can be resolved by gene duplication and specialization. As a bulk component of the lens, η-crystallin may dilute the effect of the γ-crystallins that are associated with hard, myopic lenses. It may also have a protective filtering role in the elephant shrew eye through its ability to bind UV or blue light absorbing nicotinamide cofactors and retinoids.

Elephant shrew η-crystallin has the conserved catalytic cysteine (Cys 302) and glutamate (Glu 268) as well as the NAD binding site. Recombinant η-crystallin from Cape long-eared elephant shrew (*Elephantulus edwardii*) has been shown to be able to metabolize retinal, but as in the case of other enzymes that also serve as crystallins, it is likely that it has compromised its catalytic activity to enhance its function as a crystallin (*21*, *23*). This may be reflected in a group of mutations at positions that are well conserved in other ALDHs (*19*). η-Crystallin has been classified as ALDH1A2 a member of the subfamily that exhibits specificity for retinal and may actually prefer retinal sourced from cellular retinol-binding protein (CRBP) (*2*). In turn, CRBP has been recruited as a crystallin in diurnal geckoes (*24*) in which it binds a modified retinoid as chromophore, thought to be a functional adaptation to protect the eye and to increase visual acuity in bright light (*25*).

Remarkably, other ALDHs have also been recruited as crystallins in the lenses of some invertebrates and are present in abundance in the mammalian cornea. Ω-Crystallin, the predominant protein in the lens of cephalopods is an inactive aldehyde dehydrogenase equally related to class 1 and class 2 ALDHs (*26*). Another ALDH with no activity, also designated as Ω-crystallin, is a major component of the cellular eye lens in the scallop (*27*). ALDH3 is the major soluble protein of the mammalian cornea (*28*, *29*). Certain harmful bacteria contain gene clusters known as pathogenicity islands and these include an ALDH as well as a superfamily relative of the β- and γ-crystallins (*30*, *31*). This raises the interesting possibility of a functional linkage (perhaps a form of stress response) between groups of proteins that have been picked to serve as crystallins during evolution.

With their important roles in cellular metabolism, particularly the biosynthesis of retinoic acid, and in detoxification, ALDHs are important targets for drug design. Several structures have been determined, but many aspects of their binding activity, specificity, and mechanisms remain unclear (*9*). The natural experiments in structure and function that are provided by the evolutionary process are a valuable resource for insight into these issues. Here we describe the crystal structure of η-crystallin, which has a tightly bound NAD cofactor resulting in a better-ordered cofactor binding site and differences in the dimer interface. In addition, η-crystallin has a larger substrate access tunnel than other class 1 enzymes. Moreover, the η-crystallin structure provides only the second detailed structural analysis of an enzyme crystallin (the other being avian δ-crystallin, (*32*, *33*), illustrating more of the adaptations and adaptive conflicts that can fit a protein for a new role in evolution.

EXPERIMENTAL PROCEDURES (MATERIALS AND METHODS)

Isolation of Recombinant η-Crystallin. The cloning and expression of recombinant *E. edwardi* η-crystallin were described before (*19*). For crystallization experiments, the frozen pellet, resulting from centrifugation of 1l of bacterial culture, was thawed and resuspended in 20 mL of TEG [2.5 mL of 1 M Tris-HCl pH 8.0 + 2 mL of 0.5 M EDTA pH 8.0 + 5 mL of 1 M glucose] and made up to 100 mL. Ten microliters of protease inhibitor Merck Pefabloc S C (200 mg/mL aqueous solution) was added and the mixture was left at room temperature for 10 min. The cell suspension was sonicated on ice followed by the addition of DNAse 1 (dissolved in 0.2 M NaCl) and 1 M $MgCl_2$ to the suspension to final concentrations of 10 μg/mL and 10 mM, respectively. The mixture was then centrifuged at 20 000 rpm at 4 °C for 30 min. The supernatant was dialyzed overnight at 4 °C against 25 mM Tris-HCl, 1 mM EDTA, 1 mM DTT, pH 8.0, followed by centrifugation at 20 000 rpm at 4 °C for 30 min. Ten milliliters of 0.2 μm filtered supernatant was loaded on a Q Sepharose Hi-Load 16/10 High Performance column which was run at 4 mL/min. Following one bed volume of buffer A, a linear gradient of buffer B from 0 to 70% over five bed volumes was applied. Buffer A was the same as the dialysis buffer and buffer B was 1 M NaCl made up in

Table 1: X-ray Diffraction Statistics

parameter	value	
space group	*P*21	
unit cell parameters	$a = 80.9$ Å, $b = 136.4$ Å, $c = 85.0$ Å $\alpha = \gamma = 90°$, $\beta = 102.6°$	
no. of molecules in asymmetric unit	4	
wavelength used	0.933 Å (Beamline 14-3 ESRF)	
Bragg spacing	2.4 Å	
temp of data collection	100 K	
	all data (45.6−2.4 Å)	high-resolution bin (2.53−2.40 Å)
no. reflections measured	159192	20625
no. of unique reflections	67542	8412
redundancy	2.4	2.3
completeness	96.5%	94.7%
R_{merge}	6.8%	17.4%
$\langle I \rangle / \sigma \langle I \rangle$	12.2	7.2

Table 2: Refinement Statistics

parameter	value
final R_{free}	26.2%
final R	17.9%
no. of reflections	64099 (working)/ 3443 (test)
B_{iso} RMS of protein atoms	
main chain (bond/angle)	0.43/0.83 Å^2
side chain (bond/angle)	1.35/2.35 Å^2
no. of atoms	15677
r.m.s. deviations in bond length	0.011 Å
r.m.s. deviations in bond angles	1.28°
bulk solvent parameters	density = 0.36 e/Å^3, $B = 30.2$ Å^2
solvent content	40.2%
Ramachandran plot: most favored region	89.6%
Ramachandran plot: other allowed regions	10.4%

buffer A. The protein eluted around 25% buffer B. The progress of the purification was not only monitored by the use of SDS−PAGE but also by UV absorption A_{280}:A_{260} ratio with a ratio of 2:1 being suggested as a value when the nucleic acid concentration is negligible (*34*). At this stage of the purification, there were a number of minor impurities present in the sample and the A_{280}:A_{260} ratio was around 0.75:1. The second step was gel permeation chromatography on a Superose 12, 30/10 column. Two hundred microliters of concentrated protein sample from the Q-Sepharose was applied to the column and run at 0.5 mL/min with 25 mM Bis-Tris-HCl, 1 mM EDTA, 1 mM DTT, 0.2 M NaCl, pH 6.5. The major peak had an absorption ratio of around 1.4:1. The best crystals were obtained when the gel filtration procedure was performed twice. A sample of the protein obtained from the Q Sepharose column was run on the 1 mL Resource RPC column prior to analysis on the Micromass Platform ESMS. The purified protein had a measured mass of 54 421 Da, in reasonable agreement with the calculated sequence mass of 54 407 Da for the η-crystallin sequence (SWISS-PROT id: Q28399) minus the N-terminal methionine.

Crystallization of η-Crystallin. Purified η-crystallin was equilibrated into 25 mM Bis-Tris-HCl, 1 mM EDTA, 1 mM DTT, pH 6.5 and concentrated to approximately 8 mg/mL using an Amicon Microcon 30 microconcentrater. The protein solution was crystallized by hanging drop vapor diffusion. Crystals grew in 3 days at 4 °C from drops containing an equal mixture of protein solution and well solution consisting of 25 mM Bis-Tris-propane pH 7.5 with around 14% PEG 6 K. The crystals redissolved in less than two weeks and so had to be frozen as soon as they appeared.

Structure Solution and Refinement. Data collected at Beamline ID14-3 at the ESRF, Grenoble were integrated using MOSFLM (*35*) then merged and scaled using the CCP4 suite (*36*). The protein crystallized in space group $P2_1$ with one homotetramer in the asymmetric unit. Statistics for the diffraction data collected to 2.4 Å are shown in Table 1.

The structure was solved by molecular replacement using the tetrameric sheep class 1 ALDH coordinates (PDB code: 1BXS) as a starting model. AMORE (*37*) was used for the molecular replacement (correlation coefficient 50.4, R-factor 39.4). CNS SOLVE version 1.0 (*38*) was used for refinement using simulated annealing with a maximum likelihood target. Tight noncrystallographic symmetry restraints (500 kcal/mol) were applied during early rounds of refinement gradually reducing to zero for the final rounds. Parameters for NAD refinement were derived from the HIC-up server (*39*). Manual rebuilding was undertaken using XFIT 4.0 (*40*). Waters were located using the CCP4 programs PEAKMAX and WATPEAK; after several rounds of refinement, some of these water molecules were found to have B-factors of 1.0. Examination of the electron density showed that these peaks formed part of reduced DTT molecules trapped in the lattice between tetramers. REFMAC 5 (*41*) was used for the final rounds of refinement including these DTT molecules. The refinement statistics are shown in Table 2. Residues 7−500 are visible in the electron density for all four subunits in the tetramer. The oligomer is a dimer of dimers that are labeled AB and CD. The DTT, positioned between chain B and a symmetry related chain D, is in space that cannot accommodate the oxidized ring form and might explain the short lifetime of the crystals. The numbering of η-crystallin is based on the sequence minus the N-terminal methionine. The coordinates and observed structure factors have been submitted to the protein data bank (*42*), coordinates code: 1oj9, structure factors code: r1o9jsf.

Programs Used for Analysis. Figures were produced using MOLMOL version 2k1 (*43*) and rendered using POV-RAY version 3.5. Analysis of molecular contacts was undertaken using CONTAX (*44*). Helix distances and angles were calculated with PROMOTIF (*45*). Cavity volumes were calculated using VOIDOO (*46*) starting from an approximate cavity center within the substrate tunnel and using the van der Waals exclusion method.

RESULTS

Overall Structure of the Tetramer. η-Crystallin (Figure 1) has the same tetrameric structure as other X-ray structures of both class 1 cytoplasmic sheep liver (PDB code 1BXS, 2.35 Å resolution) and rat testis (1BI9, 2.7 Å enzymes (*7*, *47*)) and class 2 bovine (1A4Z, 2.65 Å) and human (1CW3 2.58 Å) liver mitochondrial enzymes (*5*, *8*). The η-crystallin tetramer superposes very well on the sheep class 1 and human class 2 enzymes (Table 3). There is also a close fit between

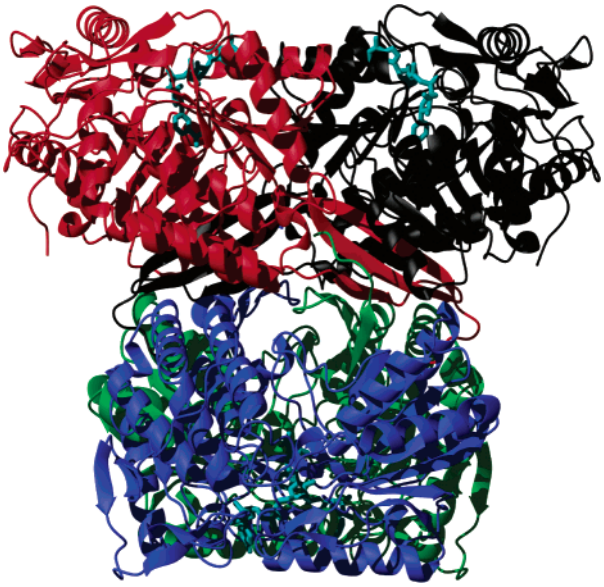

FIGURE 1: The η-crystallin tetramer showing the AB dimer in red and black and the CD dimer in blue and green. The bound NAD is in turquoise.

corresponding dimers (Table 3): for example there is an rmsd of 0.77 Å between η-crystallin AB dimer and sheep ALDH1 AB dimer. The monomer fold comprises three domains, the nucleotide-binding domain (NBD, residues 8−135 and 159−270), the catalytic domain (residues 271−470), and the oligomerization domain (residues 140−158 and 486−495) (Figure 2).

The Dinucleotide Binding Site. The Rossmann fold in ALDHs is unusual (*4*) and may be related to the variety of nucleoside conformations that appear to be required for enzyme activity (*7*, *48*). In fact, the ALDH enzymes make fewer contacts with the bound coenzyme than other dehydrogenases allowing the orientation of the nicotinamide ring, relative to the adenosine ring, to pivot around the phosphodiester, with the adenosine ring remaining fixed while the nicotinamide shifts. There is a buried coenzyme position that is called either the "hydride transfer" or "catalytic" position (observed in bovine class 2 ALDHs) and a more exposed position, "the hydrolysis" position which has been observed as the major site in sheep ALDH1 with both positions being found in human ALDH2 (*8*). The η-crystallin NAD, in the "hydride transfer" position, occupies the nucleotide binding site in a single conformation (Figure 3). All the NAD atoms are visible, and the B factors for the nicotinamide ring are similar to the adenine (average B factor nicotinamide:adenine 24.1:20.0 Å^2) unlike other published structures (sheep ALDH1 41.7:21.6 Å^2, human ALDH2 60.2:29.0 Å^2). The enzyme-crystallin therefore differs from the known ALDH enzymes in having NAD that is better defined and in a single conformation.

The η-crystallin sequences from two species of elephant shrew [Swiss-Prot accession codes Q28399 and Q29490] were aligned with seven sequences of class 1 (human [P00352], horse [P15437], bovine [P48644], mouse [P24549], 2 rat isozymes [P51647 and P13601], sheep [P51977]) and seven sequences of class 2 mammalian enzymes (two human isozymes [P05091 and P30837], horse [P12762], bovine [P20000], mouse [P47738], rat [P11884], hamster [P81178]). For protein residues involved in NAD binding, there is generally sequence consensus in these ALDHs, but there are 3D structural changes associated with the conformation of the ribose-phosphate hinge region of the NAD. In this respect, η-crystallin more closely resembles the bovine class 2 structure that has NAD bound in the "hydride transfer" conformation. When this pair of structures is compared, there are only very minor changes in the active site region associated with the lack of definition (see next section) of the catalytic base (Glu 268) (Figure 4). The only sequence difference in the neighborhood of NAD is Phe 166 in η-crystallin where all the other proteins have Ile but as this residue interacts through the main chain oxygen, it does not appear to cause structural change at the binding site.

When comparing the sequences and known X-ray structures, the only local sequence change that might have some impact on the dinucleotide conformation is the sequence change at residue 348. This residue is a glutamate in the class 1 enzymes and the sheep X-ray structure shows it interacting with NAD in the "hydrolysis" position (Figure 4). However, in η-crystallin the replacement with alanine, similar to the threonine typical of class 2 enzymes, may result in stabilization of the "hydride transfer" conformation.

The Active Site. In η-crystallin all the side chains associated with catalysis are conserved, including the two most important Cys 302 and Glu 268. The nicotinamide ring divides the catalytic site, with Glu 268 on the *si* side and Cys 302 on the *re* side (Figure 4). The sulfur of Cys 302 is 2.7 Å from the C4 atom of NAD in the "hydride transfer" conformation. As in the sheep ALDH1 structure, the side chain electron density for Glu 268 is weak in η-crystallin (the OE atoms have an average B factor of 34.3 Å^2) and it forms no hydrogen bonds or ion pairs (Figure 4). This may be due to the presence of multiple conformations of its side chain and the side chain of Glu 476 (average B factor of OE atoms 33.7 Å^2). In each monomer of the bovine class 2 structure (with NAD in the hydride transfer position), there is an ordered water molecule, considered to be catalytic, located close to the nicotinamide carbonyl that hydrogen bonds with the side chain of Glu 476 whereas in the sheep class 1 structure (with NAD in the hydrolysis position), the only active site water has variable occupancy. In η-crystallin, there is no visible water molecule in a possible catalytic position and therefore η-crystallin differs from that of the bovine ALDH2 structure (*5*). The appearance of four well-defined water molecules at the active site appears to correlate with a small substrate-binding site (see next section) rather than NAD conformation.

The Substrate Binding Site and Tunnel. The substrate tunnel is part of the interface between the NBD and the catalytic domain, with a loop from the oligomerization domain of its dimer partner stacked above it. The substrate binding site and tunnel are larger in class 1 enzymes than in the class 2 enzymes that preferentially oxidizes acetaldehyde. Unlike the rat class 1 structure, the tunnel lining residues are well defined in η-crystallin and adopt similar conformations to the sheep class 1 enzyme (Figure 5A) (*7*, *47*). The volume of the substrate tunnel is larger in η-crystallin being 1015 Å^2 compared to 627 Å^2 in sheep ALDH1.

There are sequence differences in the substrate binding region as can be seen in the orthogonal views of the tunnel in Figure 5. The cavity of the tunnel near the active site in η-crystallin is larger than in other class 1 enzymes due to

Table 3: Comparison of Geometric Parameters of ALDH Oligomers

protein	separation of G-helices AB/CD (/Å)	angle between G-helices AB/CD (/°)	RMSD for tetramer to η-crystallin (/Å)	RMSD for dimers (to η-crystallin AB dimer) (/Å)	NAD ring position
η-crystallin	8.2/8.3	−175.4/−175.1	0.0	0.00 AB 0.26 CD	IN
sheep ALDH1	7.1/7.1	−174.8/−174.8	0.82	0.77 AB 0.78 CD	OUT/in
rat ALDH1	7.8/7.3	−168.6/−171.1	1.23 (Res 7−456,478−500)	1.11 AB 1.23 CD (7−456,478−500)	undefined
human ALDH2	7.2/7.1 ABCD 7.4/7.5 EFGH	−170.8/−171.1 −170.5/−170.4	0.92 (ABCD) 0.87 (EFGH)	0.81 AB 0.82 CD 0.77 EF 0.78 GH	OUT
bovine ALDH2	7.1/7.1	−170.1/−169.8	0.93 (8−500)	0.83 AB 0.84 CD	IN

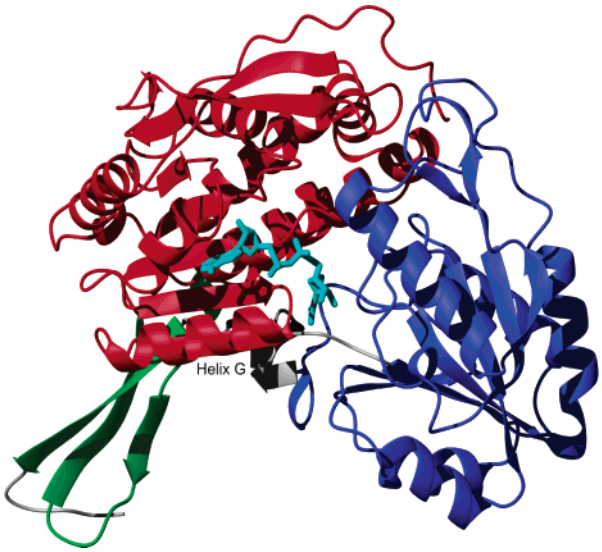

FIGURE 2: The secondary structure of the three domains of η-crystallin. The NAD binding domain is in red, the catalytic domain is in blue, and the oligomerization domain is in green. The NAD is shown in turquoise.

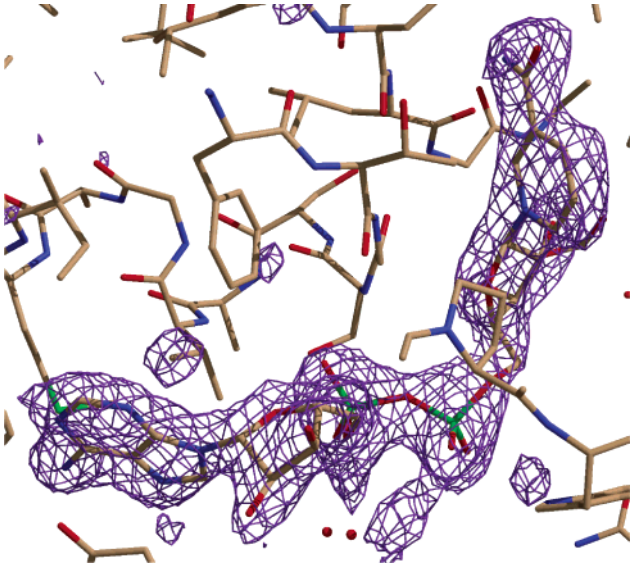

FIGURE 3: Difference map (NAD omitted during map calculation) showing clear density for a single NAD conformation in η-crystallin.

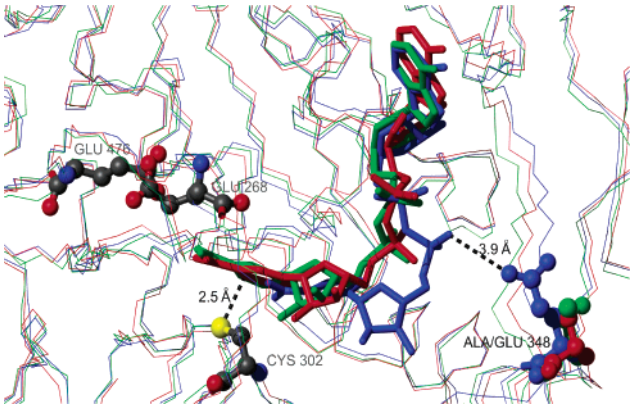

FIGURE 4: A comparison of the conformations of NAD in the three species, η-crystallin (red), bovine ALDH2 (green), and sheep ALDH1 (blue), within the context of the active site residues of η-crystallin. The interaction between Glu 348 and NAD NO1 in sheep ALDH1 (blue) is lost in η-crytallin (red) due to the mutation to Ala. This is likely to contribute to the stabilization of NAD in the "hydride transfer" conformation in η-crystallin. This view also shows how the nicotinamide group from NAD blocks the access of substrate or Glu 268 to Cys 302 (SG−NAD NC4 distance shown).

the presence of Ala 170, Cys 177, and Thr 296 in η-crystallin (Figure 5B). The η-specific replacement of Ala 465 for Phe places a smaller side chain on the roof of the tunnel, away from the entrance (Figure 5A). There is also a significant sequence change close to the tunnel entrance that could restrict access to the active site. In η-crystallin, Tyr 124 (on helix C) in the entrance to the tunnel replaces glycine or another small residue in the class 1 sequences (Figure 5A). This is similar to the bovine class 2 sequence, which has leucine at this position (Figure 5A). In addition, Leu 457 in η-crystallin replaces serine or glycine in some other class 1 enzymes, although rat and mouse (as well as the elephant shrew liver enzyme) have methionine or isoleucine at this position (Figure 5A). The entrance restriction is not relieved by the replacement of the conserved aromatic at position 296 by threonine in the η-crystallins, as the aromatic residue is packed against the side of the pocket. Overall, as a result of many η-crystallin-specific mutations, its tunnel entrance is more restricted than in the sheep ALDH1 structure (making it more like ALDH2) and leads into a tunnel cavity that is larger than ALDH1 (Figure 5).

Symmetry of the Oligomer. Class 2 enzyme kinetics imply only half of the sites are active in any single catalytic cycle, behavior considered as an extreme example of negative cooperativity (*10*). It is considered that conformational transitions play key roles in coenzyme binding and catalysis in ALDH2 and that this may be common to the superfamily (*10*). The X-ray structure of a mutant ALDH2 shown to exhibit positive cooperativity toward coenzyme binding shows a shift in the position of residues 246−260 (helix G that is both one side of the adenosine binding site and part of the dimer dyad interface) on cofactor binding that is not seen in native, as well as a disorder-to-order transition in residues 246−270 and residues 466−480 (a loop that forms

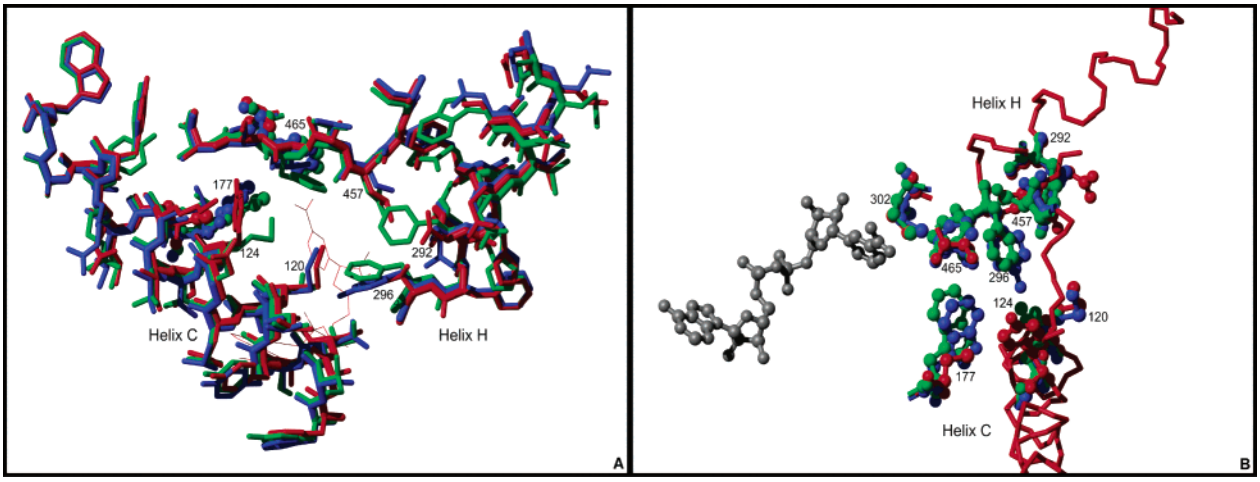

FIGURE 5: (A) The substrate binding tunnel in η-crystallin (red), sheep ALDH1 (blue), and bovine ALDH2 (green) viewed from the entrance. The variable tunnel entrance residues 120, 124, 292, 296, 457, and tunnel lining residues 177 and 465, (ball-and-stick) are labeled. Mutated entrance residues Y 124, L 457, T 296, and tunnel lining residues C 177 and A 465 in η-crystallin could affect substrate binding. Right at the back of the tunnel is the nicotinamide. Note how the tunnel wall η-crystallin helix H superposes more closely on the corresponding helix from class 1 than from class 2. (B) The substrate binding tunnel from the "side", η-crystallin in red (NAD in gray), sheep ALDH1 in blue, bovine ALDH2 in green. The mutants W177C, F465A, and F/Y296T have opened up the substrate-binding tunnel giving a much larger cavity volume.

part of the active site of its own subunit and makes contacts with residues 264 and 487 from the neighboring subunit). These regions were suggested to underpin the cooperativity (*10*) with one explanation being that the binding of coenzyme to one subunit stabilizes the position of helix G. All the available structures show full occupancy of the coenzyme site in each monomer of the tetramer, implying that two active sites per tetramer do not turn over even though they bind coenzyme; however, there is some evidence for structural asymmetry in the active sites (*10*).

In the η-crystallin adenine binding pocket, there are differences in the protein B-factors between the two monomers within each dimer. For the AB dimer for example, residues in helix F (227−233, and distal from oligomer interfaces and lattice contacts) have average B-factors of 23.2 Å^2 for the A chain and 16.2 Å^2 for the B chain. Residues in helix G also show similar, but smaller differences (28.3 Å^2 in chain A, 27.0 Å^2 in chain B). In η-crystallin, this pattern of B-factor differences is also seen in average values for the whole monomers (chain A: 27.2 Å^2, B: 22.6 Å^2, C: 29.4 Å^2, D: 21.3 Å^2). There appears to be little difference in the crystal lattice environment of these chains to explain the differences, and thus the asymmetry in the AB dimer B-factors provides some structural evidence that class 1 enzymes may also be allosteric.

There are structural and sequence differences in the interface between chains A and B (and C and D) in η-crystallin when compared to the other enzymes (Figure 6). The G helices across the dimer interface (Figure 6) are further apart in η-crystallin compared with in sheep ALDH1 and the other enzymes (Table 3). The most obvious sequence difference is the class 2-like glutamine in η-crystallin replacing Lys 254 in the other class 1 enzymes. This substitution in helix G on the AB dyad (and CD dyad) has the effect of removing two positive charges and placing the glutamine in a position to potentially hydrogen bond to its partner across the interface (Figure 6 inset). There are three other significant sequence changes close to the dimer dyad. From sequence alignments, the ALDH enzymes have conserved residues at Gly 258 (at the end of helix G), Gly 271, and Ser 273 (at the junction between the NBD and catalytic domains). In η-crystallin, these are mutated to Ala 258, Ala 271, and Asn 273, respectively, and are likely to render the enzyme-crystallin less flexible.

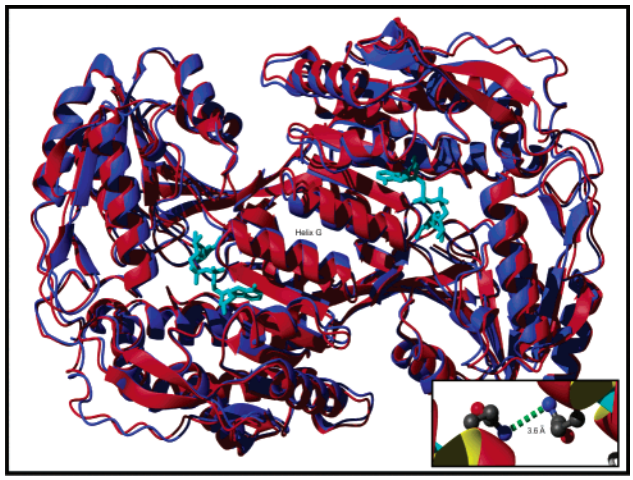

FIGURE 6: A superposition of the AB dimer of η-crystallin (red) and sheep ALDH 1 (blue) showing the dimer dyad passes between the G helices which can be seen to be further apart in the η-crystallin structure. The inset shows the replacement of Gln 254 in the G helix in η-crystallin is close to the dimer dyad axis.

DISCUSSION

When an enzyme has acquired a new role as a crystallin, enzyme activity may be compromised. In some cases, this adaptive conflict may be resolved by gene duplication and specialization, separating the competing protein functions, as illustrated by argininosuccinate lyase/δ-crystallin (*21*, *23*). In fact, there is evidence that gene duplication may also have occurred in association with the recruitment of η-crystallin, since a very similar ALDH1 is also expressed in liver and other tissues in the short-eared elephant shrew *Macroscelides proboscideus* (*19*). Even so, η-crystallin may not (yet) be fully specialized for a role in the lens since it is still expressed

in other elephant shrew tissues, particularly elsewhere in the eye (*19*). η-Crystallin sequence changes that have presumably been selected in evolution to improve the function of a class 1 ALDH as a crystallin in the lens may include modifications in three broad areas: to enhance stability (since crystallins are extremely long-lived); to avoid deleterious interactions with other crystallins and other components of the lens at the high concentrations achieved by crystallins; and to enhance binding of chromophores that can protect the eye and reduce glare (thereby enhancing visual acuity) by filtering UV and blue light.

The most notable overall difference between the η-crystallin structure and that of other class 1 ALDHs is the less mobile cofactor binding region and a possibly related shift about the dimer dyad via helix G. Even though NAD was not specifically added to the crystallization and the protein had undergone several chromatographic and dialysis steps during its purification, all the pockets were fully occupied with the dinucleotide, derived from the bacterial host, in a single conformation. This is the first time that a class 1-type structure has been solved with NAD solely in the "hydride transfer" position. When comparing the sequences and known X-ray structures, the only local sequence change that might have some impact on the stabilization of the "hydride transfer" conformation is the sequence change at residue 348. However, it seems likely that the cofactor binding is also influenced by the larger scale changes in the dynamic flexibility of the protein.

Efforts to identify the structural basis of cooperativity in the class 2 ALDH enzymes and its contribution to the enzyme mechanism have focused on cofactor binding regions and the dimer dyad interface (*10*). Less is known about the mechanism for the class 1 enzymes, but it is of interest to note that η-crystallin has an amino acid substitution at the dyad interface which renders it more class 2-like. The substitution of Gln 254 for lysine replaces a charge repulsion with a possible hydrogen bond that could stabilize the dimer interaction within the tetramer. Other changes in the same region replace glycines with other small residues and are likely to reduce main chain flexibility and these sequence changes together with altered dimer geometry and flexibility in η-crystallin may perturb any allosteric mechanisms. If flexibility is important for ALDH function, we might expect that the activity of η-crystallin had indeed been compromised. Recombinant η-crystallin exhibits no activity toward simple aldehydes but has been shown to have retinal dehydrogenase activity using an in-gel, qualitative assay (*19*). Thus, although there is some evidence that the enzyme crystallin has activity that is generally consistent with its designation as a class ALDH1A2 enzyme, quantitative data are lacking.

What can η-crystallin structure tell us about ALDH function? The structure may represent a detailed view of one of a family of conformations that are sampled by ALDHs during an active cycle that relies on flexibility, both in the protein and in the cofactor binding site. The decreased flexibility of η-crystallin may have "frozen-in" a conformation that is also used by other, more conventional members of the family. One critical piece of information that is missing in this and other solved ALDH structures is the binding of retinal itself. Retinoids are insoluble molecules that provide real challenges for introduction into proteins for crystallization. Although the recombinant η-crystallin picked up a cofactor from the bacterial host, this route cannot be used to acquire retinoids.

All the residues of the active site machinery are conserved in η-crystallin, consistent with the retention of enzyme function. However, the substrate-binding site is larger in η-crystallin, relative to other structurally characterized ALDHs. It seems unlikely that this increase could contribute to protein stability, so it may have been selected for another purpose related to the lens role. For example, this modification could allow the binding of different or larger molecules at the active site. Another difference between η-crystallin and the other solved class 1 ALDHs is in the tunnel entrance where a well-conserved glycine at position 124 has been replaced by a tyrosine that protrudes into the tunnel. The change in η-crystallin may serve to restrict passage of substrate/product molecules. If η-crystallin has in fact developed a chromophore binding function in the lens, this modification may serve to retain the chromophore once it is bound. A parallel case is the recruited ι-crystallin derived from a CRBP; however, instead of binding all-trans-retinol, the ligand of ι-crystallin is modified 3-dehydroretinol (vitamin A2) (*24*). It is unknown what, if any, chromophore might be bound by η-crystallin in the elephant shrew lens, but some of the post mortem samples of elephant shrew eyes were distinctly yellow in color (*18*). Many of the proteins that have a taxon-specific distribution in vertebrate lens have at least the potential to bind and sequester chromophores in the lens. A vitamin A derivative bound at the substrate-binding site could provide the sort of yellow filter that is seen in other diurnal lenses, while an NAD derivative (particularly NADH) could filter light in the near UV.

ACKNOWLEDGMENT

We thank the station staff at the ESRF, Grenoble, France, and Sunita Sardiwal for sequence alignments.

REFERENCES

1. Sheikh, S., Ni, L., Hurley, T. D., and Weiner, H. (1997) The potential roles of the conserved amino acids in human liver mitochondrial aldehyde dehydrogenase. *J. Biol. Chem.* 272, 18817−18822.
2. Vasiliou, V., Pappa, A., and Petersen, D. R. (2000) Role of aldehyde dehydrogenases in endogenous and xenobiotic metabolism. *Chem.-Biol. Interact. 129*, 1−19.
3. Hempel, J., Nicholas, H., and Lindahl, R. (1993) Aldehyde dehydrogenases: widespread structural and functional diversity within a shared framework. *Protein Sci. 2*, 1890−1900.
4. Liu, Z.-J., Sun, Y.-J., Rose, J., Chung, Y.-J., Hsiao, C.-D., Chang, W.-R., Kuo, I., Perozich, J., Lindahl, R., Hempel, J., and Wang, B-C. (1997) The first structure of an aldehyde dehydrogenase reveals novel interactions between NAD and the Rossmann fold. *Nat. Struct. Biol. 4*, 317−326.
5. Steinmetz, C. G., Xie, P., Weiner, H., and Hurley, T. D. (1997) Structure of mitochondrial aldehyde dehydrogenase: the genetic component of ethanol aversion. *Structure 5*, 701−711.
6. Hammen, P. K., Allali-Hassani, A., Hallenga, K., Hurley, T. D., and Weiner, H. (2002) Multiple conformations of NAD and NADH when bound to human cytosolic and mitochondrial aldehyde dehydrogenase. *Biochemistry 41*, 7156−7168.
7. Moore, S. A., Baker, H. M., Blythe, T. J., Kitson, K. E., Kitson, T. M., and Baker, E. N. (1998) Sheep liver cytosolic aldehyde dehydrogenase: the structure reveals the basis for the retinal specificity of class 1 aldehyde dehydrogenases. *Structure 6*, 1541−1551.
8. Ni, L., Zhou, J., Hurley, T. D., and Weiner, H. (1999) Human liver mitochondrial aldehyde dehydrogenase: Three-dimensional structure and the restoration of solubility and activity of chimeric forms. *Protein Sci. 8*, 2784−2791.

9. Wei, B. X., Ni, L., Hurley, T. D., and Weiner, H. (2000) Cooperativity in niotinamide adenine dinucleotide binding induced by mutations of arginine 475 located at the interface in the human mitochondrial class 2 aldehyde dehydrogenase. *Biochemistry 39*, 5295−5302.
10. Hurley, T. D., Perez-Miller, S., and Breen, H. (2001) Order and disorder in mitochondrial aldehyde dehydrogenase. *Chem.-Biol. Interact. 130*−*132*, 3−14.
11. McBee, J. K., Palczewski, K., Baehr, W., and Pepperberg, D. R. (2001) Confronting complexity: the interlink of phototransduction and retinoid metabolism in the vertebrate retina. *Prog. Retinal Eye Res. 20*, 469−529.
12. Lin, M., and Napoli, J. L. (2000) cDNA cloning and expression of a human aldehyde dehydrogenase (ALDH) active with 9-*cis*-retinal and identification of a rat ortholog, ALDH12. *J. Biol. Chem. 275*, 40106−40112.
13. Lindahl, R. (1992) Aldehyde dehydrogenases and their role in carcinogenesis. *Crit. Rev. Biochem. Mol. Biol. 27*, 283−335.
14. Niederreither, K., Subbarayan, V., Dolle, P., and Chambon, P. (1999) Embryonic retinoic acid synthesis is essential for early mouse post-implantation development. *Nat. Genet. 21*, 444−448.
15. Bitzer, M., Feldkaemeper, M., and Schaeffel, F. (2002) Visually induced changes in components of the retinoic acid system in fundal layers of the chick. *Exp. Eye Res. 70*, 97−106.
16. Haselbeck, R. J., Hoffmann, I., and Duester, G. (1999) Distinct functions for Aldh1 and Raldh2 in the control of ligand production for embryonic retinoid signaling pathways. *Dev. Genet. 25*, 353−364.
17. Grün, F., Hirose, Y., Kawauch, S., Ogura, T., and Umesono, K. (2000) Aldehyde dehydrogenase 6, a cytosolic retinaldehyde dehydrogenase prominently expressed in sensory neuroepithelia during development. *J. Biol. Chem. 275*, 41210−41218.
18. Wistow, G., and Kim, H. (1991) Lens protein expression in mammals − taxon-specificity and the recruitment of crystallins. *J. Mol. Evol. 32*, 262−269.
19. Graham, C., Hodin, J., and Wistow, G. (1996) A retinaldehyde dehydrogenase as a structural protein in a mammalian eye lens. *J. Biol. Chem. 271*, 15623−15628.
20. van Dijk, M. A., Madsen, O., Catzeflis, F., Stanhope, M. J., de Jong, W. W., and Pagel, M. (2001) Protein sequence signatures support the African clade of mammals. *Proc. Natl. Acad. Sci. U.S.A. 98*, 188−93.
21. Wistow G. J. (1995) in *Molecular Biology and Evolution of Crystallins: Gene Recruitment and Multifunctional Proteins in the Eye Lens*, R. G. Landes, Austin, TX.
22. Piatigorsky, J., and Wistow, G. (1991) The recruitment of crystallins: new functions precede gene duplication. *Science 252*, 1078−1079.
23. Wistow, G. (1993) Lens crystallins: gene recruitment and evolutionary dynamism. *Trends Biochem. Sci. 18*, 301−6.
24. Roll B., Amons R., and de Jong, W. W. (1996) Vitamin A2 bound to cellular retinol-binding protein as ultraviolet filter in the eye lens of the gecko *Lygodactylus picturatus*. *J. Biol. Chem. 271*, 10437−10440.
25. Werten, P. J. L., Roll, B., van Aalten, D. M. F., and de Jong, W. W. (2000) Gecko iota-crystallin: how cellular retinol-binding protein became an eye lens ultraviolet filter. *Proc. Natl. Acad. Sci. U.S.A. 97*, 3282−3287.
26. Zinovieva, R. D., Tomarev, S. I., and Piatigorsky, J. (1993) Aldehyde dehydrogenase-derived omega-crystallins of squid and octopus − specialization for lens expression. *J. Biol. Chem. 268*, 11449−11455.
27. Piatigorsky, J., Kozmik, Z., Horwitz, J., Ding, L., Carosa, E., Robison, W. G. Jr., Steinbach, P. J., and Tamm, E. R. (2000) Ω-Crystallin of the scallop lens: a dimeric aldehyde dehydrogenase class 1/2 enzyme-crystallin. *J. Biol. Chem. 275*, 41064−41073.
28. Cooper, D. L., Isola, N. R., Stevenson, K., and Baptist, E. W. (1993) Members of the ALDH gene family are lens and corneal crystallins. *Adv. Exp. Med. Biol. 328*, 169−179.
29. Piatigorsky, J. (1998) Gene sharing in lens and cornea: facts and implications. *Prog. Retinal Eye Res. 17*, 145−174.
30. Karaolis, D. K. R., Johnson, J. A., Bailey, C. C., Boedeker, E. C., Kaper, J. B., and Reeves, P. R. (1998) A *Vibrio cholerae* pathogenicity island associated with epidemic and pandemic strains. *Proc. Natl. Acad. Sci. U.S.A. 95*, 3134−3139.
31. Clout, N. J., Kretschmar, M., Jaenicke, R., and Slingsby, C. (2001) Crystal structure of the calcium-loaded spherulin 3a dimer sheds light on the evolution of the eye lens $\beta\gamma$-crystallin domain fold. *Structure 9*, 115−124.
32. Simpson, A., Bateman, O., Driessen, H., Lindley, P., Moss, D., Mylvaganam, S., Narebor, E., and Slingsby, C. (1994) The structure of avian eye lens delta-crystallin reveals a new fold for a superfamily of oligomeric enzymes. *Nat. Struct. Biol. 1*, 724−734.
33. Sampaleanu, L. M., Vallee, F., Slingsby, C., and Howell, P. L. (2001) Structural studies of duck delta 1 and delta 2 Crystallin suggest conformational changes occur during catalysis. *Biochemistry 40*, 2732−2742.
34. Bollag, D. M., and Edelstein, S. J. (1991) *Protein Methods*, Wiley-Liss, New York.
35. Leslie, A. G. W. (1992) in *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography*, *No. 26*.
36. Number 4 Collaborative Computational Project (1994) The CCP4 suite: Programs for the protein crystallography. *Acta Crystallogr*. D50, 760−763.
37. Navaza, J. (1994) AMoRe: An automated package for molecular replacement. *Acta Crystallogr*. *A50*, 157−163.
38. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software system for macromolecular structure determination. *Acta Crystallogr. D54*, 905−921.
39. Kleywegt, G. J., and Jones, T. A. (1998) Databases in protein crystallography. *Acta Crystallogr. D54*, 1119−1131.
40. McRee, D. E. (1999) XtalView/Xfit − A Versatile Program for Manipulating Atomic Coordinates and Electron Density. *J. Struct. Biol. 125*, 156−165.
41. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallogr. D53*, 240−255.
42. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank, *Nucleic Acids Res. 28*, 235−242.
43. Koradi, R., Billeter, M., and Wuthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics 14*, 51−55.
44. Purkiss (2000) URL: http://people.cryst.bbk.ac.uk/~bpurk01/contax/index.html.
45. Hutchinson, E. G., and Thornton, J. M. (1996) PROMOTIF − A program to identify and analyze structural motifs in proteins. *Protein Sci. 5*, 212−220.
46. Kleywegt, G. J., and Jones, T. A (1994) Detection, delineation, measurement and display of cavities in macromolecular structures. *Acta Crystallogr. D50*, 178−185.
47. Lamb, A. L., and Newcomer, M. E. (1999) The structure of retinal dehydrogenase type II at 2.7 Å resolution: implications for retinal specificity. *Biochemistry 38*, 6003−6011.
48. Cobessi, D., Favier-Tete, F., Marchal, S., Branlant, G., and Aubry, A. (2000) Structural and biochemical investigations of the catalytic mechanism of an NADP-dependent aldehyde dehydrogenase from *Streptococcus mutans J. Mol. Biol. 300*, 141−152.

BI027367W