Structural Comparison of the Enzymatically Active and Inactive Forms of δ Crystallin and the Role of Histidine 91[†,‡]

Mona Abu-Abed,[§,||] Mary A. Turner,[§] François Vallée,[§] Alan Simpson,[⊥,#] Christine Slingsby,[⊥] and P. Lynne Howell*[,§,||]

Division of Biochemistry Research, Hospital for Sick Children, 555 University Avenue, Toronto M5G 1X8, Ontario, Canada, Department of Biochemistry, Faculty of Medicine, University of Toronto, Toronto M5S 1A8, Ontario, Canada, and Department of Crystallography, Birkbeck College, Malet Street, London WC1E 7HX, U.K.

Received June 12, 1997; Revised Manuscript Received August 15, 1997[⊗]

ABSTRACT: The major soluble protein component of avian and reptilian eye lenses, δ crystallin, is highly homologous to the urea cycle enzyme, argininosuccinate lyase (ASL). In duck lenses there are two highly homologous δ crystallins, termed δ I and δ II, that are 94% identical in amino acid sequence. While δ II crystallin has been shown to exhibit ASL activity *in vitro*, δ I crystallin is inactive. The X-ray structure of a His to Asn mutant of duck δ II crystallin (H91N) has been determined to 2.5 Å resolution using the molecular replacement technique. The overall fold of the protein is similar to other members of the superfamily to which this protein belongs, with the active site located in a cleft between three different monomers of the tetrameric protein. A reexamination of the kinetic properties of the H91N mutant reveals that the mutant has 10% wild-type activity. The V_{max} of the mutant protein is identical to that of the wild-type protein, but a 10-fold increase in the Michaelis constant is seen, suggesting that His 91 is involved in binding the substrate. In an effort to determine the reasons for the loss of enzymatic activity in δ I crystallin, a structural comparison of the H91N mutant with the enzymatically inactive turkey δ I crystallin has been performed. This study revealed a remarkable similarity in the overall structures of the two proteins. Three regions of secondary structure do differ significantly between the two models; these include the N-terminal tail, a loop containing residues 76−91, and a *cis* versus *trans* peptide linkage at residue Thr 322. The *cis* to *trans* peptide variation appears to be an interspecies difference between turkey and duck and is therefore not directly involved in the loss of enzymatic activity. All the residues implicated in the catalytic mechanism are conserved in both the active and inactive proteins, and given the linearity of the relationship between the enzymatic activity of duck δ I/δ II heterotetramers and their δ II content (Piatigorsky & Horwitz, 1996), it is evident from the structure that only one of the three domains that contributes to the active site is responsible for the loss of activity in the δ I protein. Given the structural differences found in domain 1 (N-terminal tail and 76−91 loop), we postulate that these differences are responsible for the loss of catalytic activity in the δ I crystallin protein and that the δ I protein is inactive because it no longer binds the substrate.

Crystallins, the most abundant eye lens proteins, play a key role in determining and maintaining the transparency, rigidity, and refractive properties of the lens (Delaye & Tardieu, 1983; Fernald & Wright, 1983; Piatigorsky, 1984; Wistow & Piatigorsky, 1988). They constitute a diverse group of proteins with very little homology among its members (Piatigorsky, 1993). Crystallins are classified according to their distribution in vertebrate species: crystallins that are invariably incident in all vertebrates are known as the ubiquitous crystallins (α-, β-, and γ-crystallin), while those restricted in distribution to certain species are called taxon-specific (δ, ϵ, λ, etc.). Members of the taxon-specific class of crystallins are highly homologous to housekeeping enzymes (Piatigorsky, 1992; Piatigorsky & Wistow, 1989; Wistow & Piatigorsky, 1987), and in some cases the crystallins have even been shown to demonstrate *in vitro* the native enzymatic activity of their homologue (Chiou et al., 1991b; Hendriks et al., 1988; Kim & Wistow, 1993). The significance of the evolutionary relationship between crystallins and housekeeping enzymes is yet to be fully appreciated, but presumably, the recruitment of the enzyme to the eye lens is based on the necessity for the lens to acquire proteins which are thermodynamically stable and can accumulate in high concentrations (Piatigorsky et al., 1994; Wistow, 1993; Wistow & Piatigorsky, 1987). This phenomenon, i.e., two proteins with tissue-dependent functions, is called gene sharing (Piatigorsky, 1992; Piatigorsky & Wistow, 1989) and also accounts for the evolution of the ubiquitous crystallins (Piatigorsky et al., 1994).

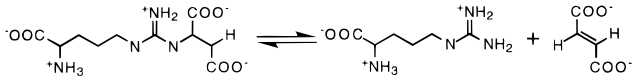
δ-Crystallin is the principal constituent of bird and some reptile eye lenses and accounts for 40−70% of all the soluble lenticular proteins. In the chicken and duck genomes, two distinct functional δ-crystallin genes, termed δ I and δ II, are present in tandem, the result of a duplication of the ancestral δ-crystallin gene (Nickerson et al., 1985, 1986; Wistow & Piatigorsky, 1990). Examination of the available

[†] Supported by a grant from the Medical Research Council of Canada.
[‡] The coordinates have been deposited in the Protein Data Bank, accession no. 1AUW.
* Corresponding author. Telephone: 416-813-5378. Fax: 416-813-5022. E-mail: howell@sickkids.on.ca.
[§] Hospital for Sick Children.
[||] University of Toronto.
[⊥] Birkbeck College.
[#] Present address: Purdue University, West Lafayette, IN 47907.
[⊗] Abstract published in *Advance ACS Abstracts,* October 15, 1997.

S0006-2960(97)01407-4 CCC: $14.00

protein sequences for δ I and δ II crystallins reveals an 88−94% identity among the various δ-crystallins and between 65% and 71% identity to the metabolic enzyme, argininosuccinate lyase (ASL) (EC 4.3.2.1) (Mori et al., 1990; O'Brien et al., 1986; Piatigorsky et al., 1988). Since both δ-crystallins are preferentially expressed in the lens tissue and as they are more identical to each other than to human ASL,[1] it appears that the recruitment of ASL to an eye lens role must have preceded the gene duplication event, demonstrating that gene duplication is not a mandatory prerequisite for the evolution of a new protein function (Hughes, 1994; Kimura & Ohta, 1974; Piatigorsky & Wistow, 1989, 1991).

Argininosuccinate lyase (ASL), the enzyme homologue of δ-crystallin, is ubiquitous to all organisms (O'Brien & Barr, 1981), where it is involved in the biosynthesis of arginine. The enzyme catalyzes the interconversion of argininosuccinate to arginine and fumarate:



L-Argininosuccinate L-Arginine Fumarate

ASL belongs to a superfamily of metabolic enzymes that all catalyze a reaction similar to that above. The tetrameric enzymes catalyze a β-elimination reaction through a general acid−base mechanism which cleaves a C_α−N or C_α−O bond and produces fumarate as one of the reaction products. Members of this superfamily include class II fumarase (Woods et al., 1988), aspartase (Takagi et al., 1986; Woods et al., 1986), adenylosuccinate lyase (Stone et al., 1993), 3-carboxy-*cis*,*cis*-muconate lactonizing enzyme (CMLE) (Williams et al., 1992), ASL, and δ-crystallin. Although these proteins are only distantly related, displaying low overall sequence homology (20−30%), there are three regions that are highly conserved (Figure 1 a). It was proposed—and later shown for two members of the superfamily—that these regions overlap in the quaternary structure forming a cleft region (Simpson et al., 1994; Weaver et al., 1995). In keeping with the overall similarity of the reactions catalyzed by members of the superfamily and the structure of *Escherichia coli* fumarase with a bound substrate−analogue (Weaver & Banaszak, 1996; Weaver et al., 1995), the active site has been proposed to lie within this cleft region, with a number of the highly conserved residues believed to interact with the fumarate moiety of the various substrates and to participate directly in catalysis.

δ I and δ II crystallins are the products of a gene duplication event of an ancestral ASL gene. Despite the high sequence identity between δ I and δ II crystallins, 91% between chicken and 94% between duck δ I and δ II crystallins, respectively, and the similarity between these crystallins and human ASL (64−71.5% identity), ASL activity has only been detected in the δ II isoform. No enzymatic activity is exhibited by any of the δ I crystallin proteins examined to date (Chiou et al., 1991a,b; Lee et al., 1992; Piatigorsky et al., 1988). The difference in the catalytic activity of these two extremely homologous proteins (δ I and δ II crystallins) must be the result of the variations in the primary structure of the proteins (Figure 1b). These sequence variations could cause a loss of enzymatic activity due to a mutation in one or more of the residues involved in the catalytic mechanism, or they could result in differences in higher levels of structure, for example, changes in the conformation, the overall fold, and/or subunit assembly.

A histidine residue has been implicated in the catalytic mechanism of both ASL (Garrard et al., 1985) and δ II crystallin (Chiou et al., 1991a), and since in the enzymatically inactive δ I crystallins histidine 91 is mutated to glutamine, it was previously postulated that histidine 91 played a direct role in catalysis (Barbosa et al., 1991b). In an attempt to test this hypothesis and identify the catalytic histidine, O'Brien and colleagues (Patejunas et al., 1995) constructed four single site-directed mutants in which the histidine residues at positions 91, 110, 162, and 178 were replaced with asparagine residues (H91N, H110N, H162N, and H178N). Preliminary kinetic analysis of these mutants revealed that the H91N and the H162N mutants were completely inactive, thus seemingly confirming the catalytic role proposed for His 91 (Patejunas et al., 1995). Our subsequent reevaluation of the enzymatic activity of the H91N δ II crystallin, presented in this paper, revealed that this mutant does exhibit low levels of activity, and the kinetic parameters for the H91N mutant have been determined and compared to those of the wild-type protein. A structural study of the H91N mutant of δ II crystallin was therefore undertaken to provide insight into the role of His 91 in the catalytic mechanism. While the structure presented here is that of a partially active mutant, it is structurally indistinguishable from the fully active H178N mutant and wild-type duck δ II crystallin, for which only low-resolution (∼3 Å) structures are presently available. The H91N model has therefore been used in a structural comparison of the enzymatically inactive turkey δ I protein to determine whether there were any major conformational differences between the δ II and δ I proteins that could account for the loss of enzymatic activity in the δ I protein.

MATERIALS AND METHODS

Gene Expression, Purification, Crystallization, and Data Collection. Duck δ II crystallin derived from the cDNA of duck (*Anas platyrhonchos*) eye lens was overexpressed in *E. coli* strain BL21(DE3)pLysS using a T7 polymerase system. Details of the expression of the wild-type and the H91N mutant of duck δ II crystallin (Patejunas et al., 1995) have been described previously (Barbosa et al., 1991b). Frozen cells from 2 L of culture were suspended in 125 mL of ice-cold column buffer consisting of 10 mM Tris-HCl, pH 7.5, 1 mM EDTA, and 1 mM dithiothreitol. The cell suspension was divided into five fractions prior to cell lysis. Each fraction was sonicated for a total of 2 min with 30 s pulses followed by 120 s of cooling on ice. These were centrifuged at 15 000 rpm for 30 min at 4 °C in a Beckman JA-20 rotor. The soluble fractions were then applied to a previously equilibrated DEAE-cellulose (DE-53 Whatman) ion-exchange column (1.6 cm × 40 cm) at a flow rate of 1 mL/min. The column was washed with 300 mL of column buffer to remove unbound proteins. A 500 mL linear gradient of 0−150 mM NaCl in column buffer was used to elute the protein. The protein eluted at approximately 75 mM NaCl. Fractions with a relatively high content of the

[1] Abbreviations: ASL, argininosuccinate lyase; NCS, noncrystallographic symmetry; H91N, histidine to asparagine mutation at residue 91; CMLE, 3-carboxy-*cis*,*cis*-muconate lactonizing enzyme.

a

```
                114     121 159       168 282             296
DDC II     SRNDQVVT  GYTHLQKAQP  GSSLMPQKKNPDSLE
DDC I      SRNEQVVT  GYTHLQKAQP  GSSLMPQKKNPDSLE
CDC II     SRNDQVVT  GYTHLQKALP  GSSLLPQKKNPDSLE
CDC I      SRNEQVVT  GYTHLQKALP  GSSLLPQKKNPDSLE
ASL_H      SRNDQVVT  GYTHLQRAQP  GSSLMPQKKNPDSLE
ASL_Y      SRNDQVVT  GYTHLQRAQP  GSSLMPQKKNADSLE
FUM_BS     SSNDTFPT  GRTHLQDATP  GSSIMPGKVNPTQSE
FUM_EC     SSNDVFPT  GRTHLQDATP  GSSIMPGKVNPTQCE
FUM_H      SSNDTFPT  GRTHTQDAVP  GSSIMPGKVNPTQCE
ASP_BS     SQNDVFPT  GRTHLQDAVP  GSSIMPGKVNPVMAE
ASP_EC     STNDAYPT  GRTQLQDAVP  GSSIMPAKVNPVVPE
ADL_BS     TSTDVVDT  GRTHGVHAEP  GSSAMPHKRNPIGSE
ADL_H      TSCYVGDN  GSTHFQPAQL  GSSAMPYKRNPMRSE
ADL_EC     TSEDINNL  SRTHGQPATP  GSSTMPHKVNPIDFE
CMLE_PP    TSQDAMDT  GRTWLQHATP  GSSTMPHKRNPVGAA

Consensus  S-ND-V-T  G-THLQ-A-P  GSS-MP-K-NP---E
```

b

```
DDC I      ----..---- m--------- ---qm-st-- st-------- --a-i----- ---------- ---------- 68
DDC II     ----ar---- -----s---- ----k----- ---------- ---------- ---------- ---------- 70
CDC I      --t-..---- l--------- ----i-s--- st----t--- --a------- ----s----- ---------- 68
CDC II     ----..---- -----s---- ----m----- ---------- ---------- ---------- ---------- 68
Consensus  MASE--GDKL WGGRFVGSTD PIME-LNSSI AYDQRLSEVD IQGSMAYAKA LEKAGILTKT ELEKILSGLE

DDC I      -----l---- i--t------ q--------- ---------- ---------- --------n- l--------- 138
DDC II     ---------- ---k------ ---------- ---------- ---------- --------n- l--------- 140
CDC I      -----s---- --mt------ q--i------ ---------q ---------- --------s- i--------- 138
CDC II     ---------- ---k------ ---------- ---------- ---------- --------s- i--------- 138
Consensus  KISEEWSKGV FVV-QSDEDI HTANERRLKE LIGDIAGKLH TGRSRNDQVV TDLKLFMK-S -SIISTHLLQ

DDC I      ---------- ---------- ---------- ---------- ---------- ---------- ---------- 208
DDC II     ---------- ---------- ---------- ---------- ---------- ---------- ---------- 210
CDC I      ---------- ---------- ------l--- ---------- ---------- ------t--- ---------- 208
CDC II     ---------- t--------- ------l--- ---------- --i------- -----ms--- ---------- 208
Consensus  LIKTLVERAA IEIDVILPGY THLQKAQPIR WSQFLLSHAV ALTRDSERLG EVKKRINVLP LGSGALAGNP

DDC I      ---------- ---------- ---------- ---------- ---------- -------s-- --l------- 278
DDC II     ---------- ---------- ---------- -----fl--- ---------- ---h---s-- --l------- 280
CDC I      ---------- ---mt--t-- -i-------- ----i----- ---------- -------t-- ---------- 278
CDC II     ---------- ---------- ---------- ---------- ---------- -------t-- ---------- 278
Consensus  LDIDRELLRS ELEFASISLN SMDAISERDF VVEFLSVATL LMIHLSKMAE DLIIYST-EF GFVTLSDAYS

DDC I      ---------- ---------- ---------- s--------- ---------- ----i----- ---------- 348
DDC II     ---------- ---------- ---------- s--------- ---------- ---------- ---------- 350
CDC I      ---------- ---------- ---------- a-------i- ---s------ ---------- ---------- 348
CDC II     ---------- ---------- ---------- a--------- ---------- ---------- ---------- 348
Consensus  TGSSLMPQKK NPDSLELIRS KAGRVFGRLA -ILMVLKGLP STYNKDLQED KEAVFDVVDT LTAVLQVATG

DDC I      -------s-- ---------- ---------- ---------- ---------- --------n- ---------- 418
DDC II     -------s-- ---------- ---------- ---v------ ---------- -----t--k- s--------- 420
CDC I      -------n-- ---------- -s-------- -----i---q ---------- -----t--n- ---------- 418
CDC II     -------n-- ---------- -s-------- ---------- v--------- --------k- ---------- 418
Consensus  VISTLQI-KE NMEKALTPEM LATDLALYLV RKGMPFRQAH TASGKAVHLA ETKGIAIN-L TLEDLKSISP

DDC I      ---------- ---------- ---------- -t-------- --------. 466
DDC II     q--------- ---------- --a------- -t-------- --------. 468
CDC I      --a------- sv-------- -v-------- -a-------- --------A 467
CDC II     --a------- -i-------- -v-------- -a-------- --------. 466
Consensus  LFSSDVSQVF NFVNSVEQYT ALGGTAKSSV T-QIEQLREL LKKQKEQA.
```

FIGURE 1: (a) Three regions of high sequence similarity among members of the superfamily. In duck δ II crystallin numbering conserved region 1 encompasses residues 114−121; conserved region 2, residues 159−168; and conserved region 3, residues 282−296. (b) Complete sequence alignment of duck and chicken δ I and II crystallin. Key to sequence alignments: DDC II, duck δ II crystallin; DDC I, duck δ I crystallin; CDC II, chicken δ II crystallin; CDC I, chicken δ I crystallin; ASL−H, human argininosuccinate lyase; ASL−Y, yeast argininosuccinate lyase; FUM−BS, *B. subtilis* fumarase C; FUM−EC, *E. coli* fumarase C; FUM−H, human fumarase C; ASP−BS, *B. subtilis* aspartase; ASP−EC, *E. coli* aspartase; ADL−BS, *B. subtilis* adenylosuccinate lyase; ADL−H, human adenylosuccinate lyase; ADL−EC, *E. coli* adenylosuccinate lyase; CLME, *P. putida* 3-carboxy-*cis*,*cis*-muconate lactonizing enzyme.

mutant δ II crystallin, as determined by the presence of a 50 kDa band on a SDS−PAGE gel, were pooled and concentrated. The concentrated sample was desalted and exchanged into 10 mM Tris-HCl, pH 7.5, by repeated dilution and further purified by chromatofocusing on a Mono-P FPLC column (HR 5/20). A concentrated sample (1.5 mL of 10 mg/mL) was applied to the column previously equilibrated with starting buffer (25 mM Bis-Tris, pH 7.1). The column was subsequently washed with 12 mL of starting buffer. A pH gradient from 7 to 4 was achieved in the Mono-P column with the use of 100 mL of eluting buffer (0.1 × Polybuffer 74, pH 4). The fractions that exhibited the highest purity were pooled and concentrated. This procedure yielded ~30 mg of >98% pure protein/2L of cells.

Crystallization and Data Collection. The initial crystallization conditions found for duck δ-crystallin have been described elsewhere (Abu-Abed et al., 1994). However, prior to data collection at the Cornell High Energy Synchrotron

Table 1: Data Collection and Processing Statistics

space group	$P2_1$
unit cell dimensions	94.1 Å, 99.9 Å, 108.7 Å, $\beta = 102°$

	set 1	set 2[a]
max resolution (Å)	3.0	2.5
total no. of observations	78170	239855
unique reflections	39012	65461
completeness (%)		
all data	99 (85.4)[b]	95 (92.6)[b]
$I \geq 2\ \sigma(I)$	54 (12.1)[b]	82 (60.0)[b]
$R_{sym}(I)$[c]	0.137 (0.45)[b]	0.103 (0.33)[b]

[a] This data set was measured from nine crystals. [b] The numbers given in parentheses denote the values in the highest resolution shell, 3.15−3.0 Å and 2.6−2.5 Å for data sets 1 and 2, respectively. [c] $R_{sym}(I) = \Sigma|I - \langle I\rangle|\Sigma I$, where I is the intensity measurement for symmetry-related reflections and $\langle I\rangle$ is the mean intensity for the reflection.

Source, these conditions were further optimized to yield thicker and more uniform crystals. This was achieved by using a higher concentration of protein, a slightly different precipitating agent, monomethyl ether PEG 2K instead of PEG 4K, and a higher concentration of $MgCl_2$. All crystals were grown by the vapor diffusion technique in hanging drops at room temperature. Five microliters of protein (14−16 mg/mL, 100 mM Tris-HCl, pH 8.5) was mixed with an equal volume of the precipitating solution [18−20% (w/v) PEG MME 2K, 300 mM $MgCl_2$, 100 mM Tris-HCl, pH 8.5] and suspended over a 1 mL reservoir. Flat plate-like crystals grew within approximately 4 days to a maximal size of 0.8 mm × 0.45 mm × 0.25 mm. An initial low-resolution 3 Å data set (set 1, Table 1) was collected at room temperature on a Siemen's Xentronics area detector with a Rigaku RU200 rotating anode X-ray generator ($\lambda = 1.541$ Å) and processed using the XDS software package (Blum et al., 1987; Kabsch, 1988a,b). A second higher resolution data set (set 2, Table 1) was subsequently collected at room temperature on Fuji image plates at the Cornell High Energy Synchrotron Source (station F_1, $\lambda = 0.91$Å). The image plates were read using a Fuji BAS 2000 scanner and the data processed using the DENZO/SCALEPACK software package (Otwinowski, 1993). When first exposed to X-rays the crystals were found to diffract to ∼2 Å but were susceptible to radiation damage, and therefore using the criteria that the data in the final resolution shell must be >90% complete overall with at least 50% of the intensities over $2\sigma(I)$, the effective resolution of the data was considered to be 2.5 Å.

Structure Solution. The structure of H91N δ II crystallin was solved by molecular replacement using the program X-PLOR, version 3.1 (Brünger, 1990), and the coordinates of the enzymatically inactive turkey δ I crystallin as the search model (Simpson et al., 1994). Since the sequence of turkey δ I crystallin is unknown, the sequence of chicken δ I crystallin has been used in the turkey δ I model. Data between 8 and 4 Å resolution (set 1, Table 1) and the coordinates of all atoms of the tetrameric turkey δ I crystallin were used in all rotation and translational searches. The rotation searches and Patterson correlation refinement yielded four significant peaks. These peaks were equivalent to each other and found to be related by the 222 noncrystallographic symmetry found in the search model. The translational search produced a unique solution, 9.1 standard deviations above the mean and 6σ higher than the next peak. The final rotation/translation solution had an *R*-factor of 43.77% [*R*-free (Brünger, 1992) 43.84%] for data between 8 and 4 Å resolution.

Refinement. Rigid-body refinement was carried out using the program X-PLOR (version 3.1). Initially the orientation and position of each of the four subunits in the tetramer were refined. Each subunit was then further divided into four rigid bodies, consisting of residues 17−112, 113−361, 362−431, and 432−465, which correspond to the three domains of the protein and the C-terminal helix. This reduced the *R*-factor to 33.1% (*R*-free 33.5%) for data between 8 and 3 Å resolution. The necessary amino acid substitutions were made at this stage to incorporate the correct sequence into the model which was then subjected to eight rounds of simulated annealing X-PLOR refinement (X-PLOR versions 3.1 and 3.8). The first two stages of the refinement used the data between 8 and 3 Å resolution (set 1, Table 1) with noncrystallographic symmetry (NCS) restraints (force constant = 200 kcal mol^{-1} Å^{-2}) imposed on all non-hydrogen atoms in each subunit and an overall *B*-factor of 15 Å^2. Later rounds of refinement used the higher resolution data set and data between 8 and 2.5 Å (set 2, Table 1). During round 3, the NCS restraints were released and individual isotropic *B*-factors refined. All rounds of simulated annealing refinement were alternated with sessions of rebuilding using the program O (Jones et al., 1991) or Turbo-Frodo (Cambillau et al., 1996) and electron density maps with coefficients $2F_o - F_c$ and $F_o - F_c$ or the equivalent sigmaA weighted maps (Hodel et al., 1992; Read, 1986). Simulated annealing omit trials were used to remove model bias and to validate the correctness of certain sections of the model. Solvent molecules were added gradually starting at round 5. The computer program SRCHXPL (Larson, unpublished program) was used to analyze the $F_o - F_c$ map such that densities at least 2.5σ above the mean were picked and their *x*, *y*, *z* coordinates output. Each potential water molecule was subsequently examined using the program O or Turbo-Frodo in conjunction with the model of the protein and both $2F_o - F_c$ and $F_o - F_c$ or the equivalent sigmaA weighted electron density maps. Water molecules were only included if the both the $2F_o - F_c$ (1σ above mean) and $F_o - F_c$ density warranted, if they made at least one hydrogen bond to a protein or other solvent atom and their *B*-factor, after one round of simulated annealing refinement, was less than 70 Å^2. Due to density of insufficient quality, the first 16 N-terminal and last 3 C-terminal residues have been omitted from the model.

An assessment of the quality of the structure was made using the program PROCHECK (Laskowski et al., 1993), and the geometry was generally found to be excellent. The Ramachandran plot showed that >94% of the residues reside in the most favored regions and that no residues reside in the disallowed regions. Residue Leu 204 of each monomer falls in the generously allowed region and is found to be located in a type IV β-turn. Considerable difficulty was encountered in model building Ser 321 and Thr 322 into the available density. While the sigmaA weighted $2F_o - F_c$ and simulated annealing omit maps show good contiguous density for these residues, all attempts to rebuild the residues with a *trans* peptide and refine their geometry resulted in a few atoms (e.g., carbonyl oxygen of Thr 322) which poorly fit the density. Model building the peptide, however, as a *cis* peptide resulted in a model in which all atoms fit the density (Figure 2). A summary of the final refinement

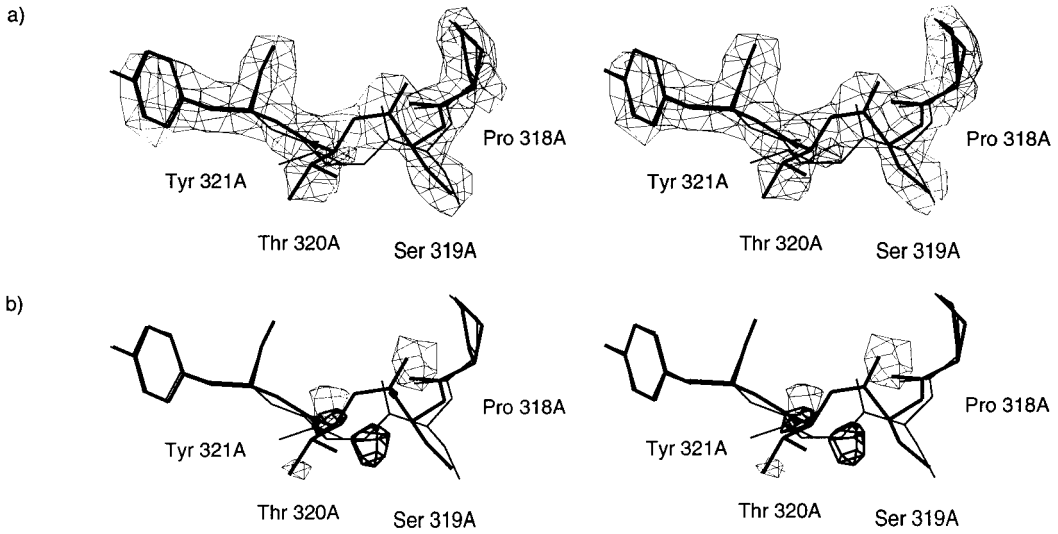


FIGURE 2: (a) Omit map of residues Pro 320A to Tyr 323A showing the *cis* and *trans* configurations of the Thr 322 peptide linkage. In both (a) and (b) the *cis* configuration is represented by the thick lines and the *trans* configuration by the thin lines. (b) $F_o - F_c$ difference density for residues Pro 320A to Tyr 323A. The map is contoured at $+2.5\sigma$ (thin lines) and -2.5σ (thick lines).

Table 2: Summary of the Final Model Refinement Statistics

resolution (Å)	8−2.5
R-factor[a] [$F > 2\ \sigma(F)$]	0.161
R-factor[a] (all data)	0.178
R-free[b]	0.235
no. of non-hydrogen atoms	13912 (1796 residues)
no. of solvent atoms	343
rms deviation from ideality	
bond lengths (Å)	0.006
bond angles (deg)	1.16
dihedral angles (deg)	19.92[c]
improper (deg)	2.86
average *B*-factors (Å^2)	
main chain	29.96
side chain	33.37
all protein atoms	31.91
solvent	32.47

[a] *R*-factor $= \sum(|F_o| - |F_c|)/\sum|F_o|$. [b] *R*-free $= \sum(|F_{o,s}| - |F_{c,s}|)/\sum|F_{o,s}|$, where s refers to a subset of data not used in refinement comprising 10% of the data. [c] This value excludes the *cis* peptide bond.

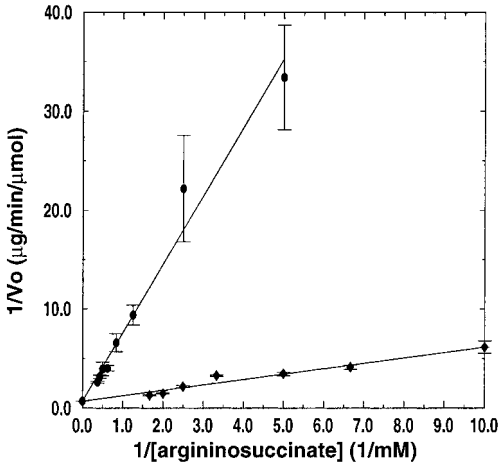


FIGURE 3: Kinetic analysis of the H91N and wild-type duck δ II crystallin: Lineweaver−Burk plots for the forward reaction, catalyzed by wild-type (◆) and H91N (●) duck δ II crystallin, respectively. The mean of three measurements is plotted, and the error bars correspond to the standard error of the mean.

statistics is given in Table 2. A comparison of each of the individual monomers in the tetramer revealed that the rms deviation between all backbone atoms in four NCS-related monomers is 0.29 ± 0.05 Å. There are a number of regions that show significant differences between the monomers; these all occur in loop regions, where differences between structures are more typically expected due to the poorer density found for these regions.

Kinetic Analysis. The ASL activity of wild-type and H91N duck δ II crystallin was assayed at 22 °C according to the method described previously (Havir et al., 1965). Stock solutions of potassium argininosuccinate were prepared by adding a solution of K_2SO_4 to barium argininosuccinate (Sigma Chemical Co.) in a 1:1.25 molar ratio and centrifuging to remove the $BaSO_4$ precipitate. The reaction mixture was prepared by diluting the stock solution to the desired concentration in 50 mM Tris-HCl, pH 7.5. Reactions were initiated by adding 5−20 μL of a dilute protein solution (1−2 mg/mL) to 1 mL of the reaction mixture. After brief mixing, the production of fumarate was monitored by measuring the increase in UV absorption at 240 nm on a Milton Roy 3000 spectrophotometer. The rates of the forward reaction were measured for substrate concentrations ranging from 0.2 to 2.67 mM with three replicate measurements taken at each concentration. Initial velocities were normalized to the amount of protein used in each reaction and kinetic parameters calculated from a Lineweaver−Burk plot (Figure 3). Wild-type duck δ II crystallin was similarly characterized for comparison purposes.

RESULTS AND DISCUSSION

Tertiary and Quaternary Folds and Similarity to Other Superfamily Members. H91N δ II crystallin has the same overall global fold as turkey δ I crystallin (Simpson et al., 1994) and human ASL (Turner et al., 1997). In the present structure, each of the four identical subunits consists of 449 amino acids, corresponding to residues 17−465. The first 16 amino-terminal and the last 3 carboxy-terminal residues (466−468) are not included in the current model as the electron density for these residues was of insufficient quality for the conformation to be determined, implying that these residues are conformationally flexible. The secondary structure is predominantly helical (~60% of the sequence) with a total of 21 helices per subunit. A schematic

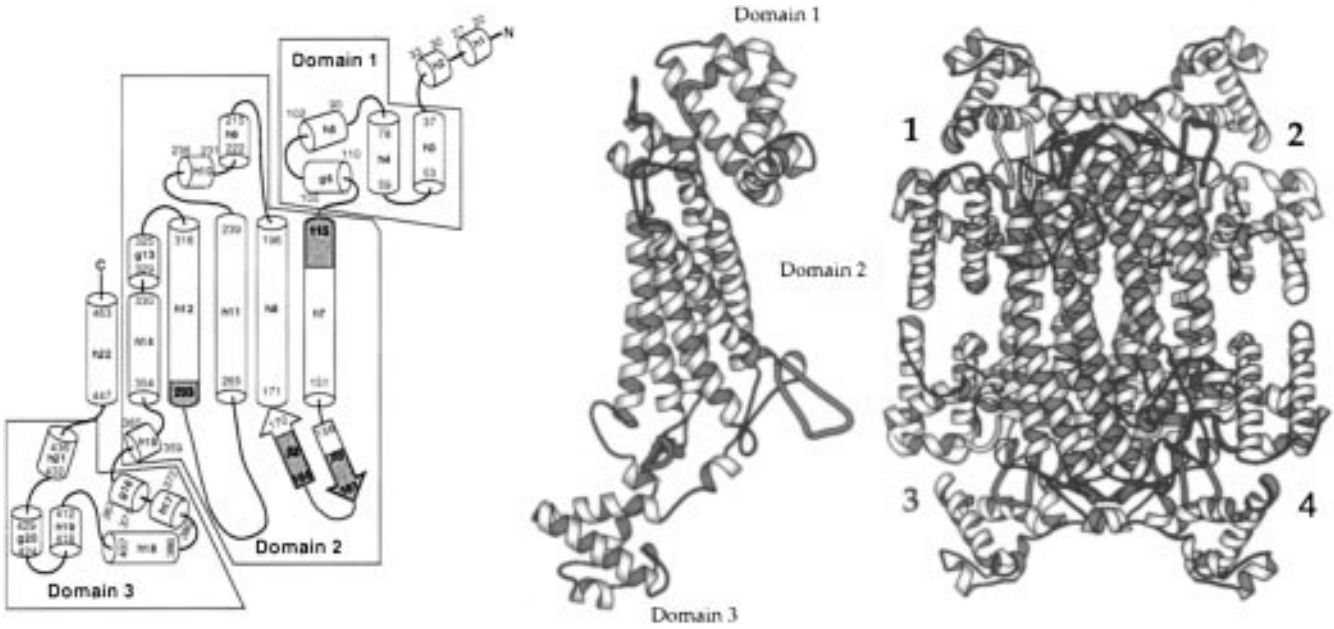


FIGURE 4: Schematic diagrams showing (a, left) the two-dimensional and three-dimensional topology of the duck δ II crystallin (b, center) monomer and (c, right) tetramer. (a) Diagram showing the extent of the secondary structural elements with helical segments depicted as cylinders (h, α-helix; g, 3_{10} helix), β-sheets (β) by arrows, and loop regions by dark lines. The residues at the beginning and end of each segment of secondary structure are numbered accordingly. The first 16 N-terminal and the last 3 C-terminal residues have been omitted from this diagram (see text). The secondary structure elements were assigned using the program DSSP (Kabsch & Sander, 1983). The dashed lines represent the extent of the three domains into which this monomer can be divided. The regions of highly conserved sequence across the superfamily given in Figure 1a are shaded in red in both panel a and panel b. (b) Schematic diagram of the three-dimensional topology of a monomer. (c) Schematic ribbon diagram of the tertiary structure of the tetramer of H91N δ II crystallin. The highly conserved regions across the superfamily are colored according to the monomer in which they are located: monomer A, red; monomer B, blue; monomer C, green; monomer D, yellow. The numbers indicate the location of the four active sites. Panels b and c were prepared using the program Molscript (Kraulis, 1991).

illustration of the overall two- and three-dimensional topology of the subunit is given in Figure 4 a,b. The subunit is approximately 95 Å × 40 Å × 35 Å. There are three distinct domains in each subunit, which are flanked by N- and C-terminal arms. Domains 1 and 3 have a similar overall topology with each domain consisting of two helix−turn−helix motifs, arranged mutually perpendicular to each other. Domain 2 is comprised of nine helical segments, five of which are coaxially aligned in an up−down−up−down−up orientation, to form a central five-helix bundle.

The tetramer is approximately 110 Å × 95 Å × 95 Å and can be thought of as two dimers of closely interacting monomers (Figure 4c). The closely interacting monomers associate via helices h8, h11, and h12, while the tetramer is formed by the association of two such dimers via helix h12 of each monomer. The four central h12 helices form a classical four-helix bundle at the core of the protein. Both the closely associated monomers and the tetramer are held together by mainly hydrophobic interactions along the length of helices h8, h11, and h12. The overall structure of the tetramer exhibits 222 noncrystallographic symmetry as previously observed in the self-rotation function (Abu-Abed et al., 1994).

In addition to the structural similarity with turkey δ I crystallin (Simpson et al., 1994) and human ASL (Turner et al., 1997), duck δ II crystallin is also structurally similar to fumarase C from *E. coli* (Weaver et al., 1995), hence reinforcing the hypothesis that all members of this enzyme superfamily will share a common structural fold. Some variation in the topology of the δ crystallins/ASL and fumarase C does occur in domain 1, with differences seen in the connectivity of some of the helical segments and in the presence of an N-terminal antiparallel β-sheet in the fumarase C structure which is absent in the crystallin structures.

Location of the Active Site. The active site has been defined previously (Simpson et al., 1994; Weaver et al., 1995) for the structures of turkey δ I crystallin and fumarase C, respectively. The three regions of highly conserved residues (Figure 1a), which are well dispersed in the structure of the monomer (Figure 4a,b), come together in the tetramer (Figure 4c) to form the putative active site cleft (Figure 5). The location of the active site has been confirmed for fumarase C (Weaver & Banaszak, 1996; Weaver et al., 1995, 1997). In the structure solution the fumarate analogue, tungstate, was used as a heavy atom derivative and was found to be located close to the highly conserved regions. Subsequent studies of fumarase C with the inhibitor 1,2,4,5-tetracarboxybenzene and citrate have provided further evidence as to the location of the active site (Weaver & Banaszak, 1996) as have the structures of two mutant forms of *E. coli* fumarase C (Weaver et al., 1997). These studies also provided evidence for the presence of a second dicarboxylic binding site which Weaver et al. (1997) suggest is a potential activation site for the fumarase enzyme. There is no evidence, to date, that such an activation site is present in the duck δ II crystallin or human ASL structures.

In δ II crystallin, the conserved regions c2 from monomer B and c3 from monomer C are close together, with the $N_{\delta 1}$ of His 162B within hydrogen-bonding distance of the $O_{\epsilon 2}$ of Glu 296C or, if the orientation of the histidine residue was inverted, the Oδ1 of Asn 289C. Other interactions between the two conserved regions occur between the $O_{\gamma}1$ and carbonyl oxygen of Thr 161B and the N_{ζ} of Lys 289C

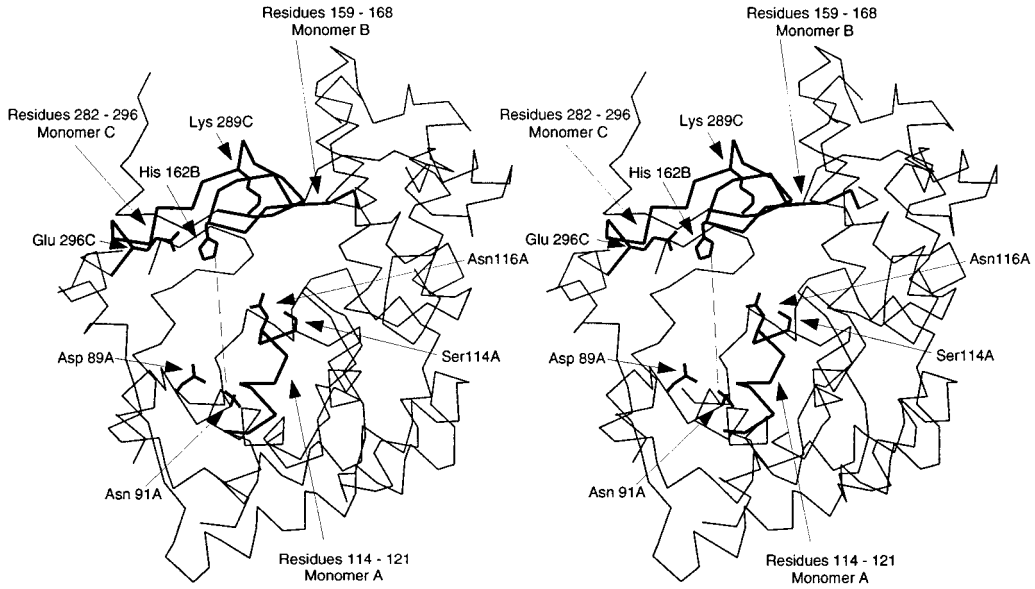


FIGURE 5: Stereoview of the putative active site of duck δ II crystallin. Conserved regions 1, 2, and 3 are contributed to the active site by different monomers and are represented by thick lines in the diagram; each is labeled according to the monomer to which they belong. The distance between Asn 91 and His 162 is 13.8 Å.

and Gln 164B $N_{\epsilon 2}$ and the carbonyl oxygen of Lys 290C. The conserved residues of region c1 (114−120) do not interact directly with either of the other two regions (c2 or c3) (see Figure 5). Only Asn 116 and Ser 114, two of the most conserved residues across the superfamily in this region (Figure 1a and 5), are oriented toward the highly conserved His 162, Glu 296, and Lys 289 residues but are ∼6 and ∼9 Å away from His 162, respectively. Although it had been speculated, based on the fumarase structure (Weaver et al., 1995), that Ser 114 has a catalytic role, there is no experimental evidence that supports this hypothesis. If the conserved residues in region c1 play any role in the catalytic mechanism, it appears more likely that it is structural, involved potentially in the binding of the substrate rather than directly in catalysis, a hypothesis that is supported by the lack of homology across the superfamily in this region (Figure 1a) and the structure of fumarase C with citrate bound (Weaver & Banaszak, 1996).

Catalytic Mechanism. Kinetic studies of the bovine ASL (Garrard et al., 1985) and duck δ II crystallin (Lee et al., 1993) have implicated a histidine residue in the catalytic mechanism. The proximity of His 188 (equivalent to His 162 in duck II crystallin, region c2) in fumarase to the tungstate ion (Weaver et al., 1995) would suggest that His 162 is the catalytic histidine in ASL/δ II crystallin. Site-directed mutagenesis of His 162 to an asparagine residue (Patejunas et al., 1995) confirms the importance of this residue in the catalytic mechanism as enzymatic activity is completely abolished. However, until the structure of a substrate analogue or inhibitor−protein complex for ASL/δ II crystallin has been determined, it is still not clear what role His 162 plays in the reaction mechanism and what other residues may be involved in binding the substrate and catalysis.

The presence of a hydrogen bond between His 162 ($N_{\delta 1}$) and Glu 296 $O_{\epsilon 1}$ and the absence of any well-defined water molecule coordinated to the histidine would seem to support the original "charge relay" hypothesis of Weaver et al. (1995) and would suggest that His 162 was the catalytic base. The absence of any well-ordered water molecule from the δ II crystallin structure in this location does not, of course, rule out the possibility that this His 162 could be responsible for activating a water molecule which would subsequently act as the base as suggested by Weaver et al. in their subsequent study of a number of inhibitor molecules bound to fumarase C (Weaver & Banaszak, 1996). In both scenarios, the hydrogen bond between His 162 and Glu 296 would make His 162 more nucleophilic and therefore more capable of abstracting a proton either from the C_β position of the substrate to initiate the reaction or from the water molecule. The hydrogen bond interaction would also render His 162 less likely to carry a positive charge and therefore less likely to offer stabilization of the negatively charged carbanion intermediate.

One argument against His 162 having a direct role in catalysis is that if histidine 162 was the catalytic base, then the reaction mechanism is not conserved across all members of the superfamily, as both CMLE and *E. coli* aspartase lack a histidine at this position. The histidine has been replaced by tryptophan and glutamine in CMLE and *E. coli* aspartase, respectively (Figure 1a). In CMLE, Glu 296 is also not conserved. *E. coli* aspartase also appears to be unique in the superfamily, as the enzyme requires a divalent metal ion for activity at alkaline pH (Flazone et al., 1988; Ida & Tokushige, 1985).

Also located in the active site cleft are the highly conserved residues, Lys 289 and Asn 291, and Thr 161. The N_ξ of Lys 289 is positioned within hydrogen-bonding distance of $O_{\delta 1}$ of Asn 291 and the carbonyl oxygen and $O_{\gamma 1}$ of Thr161. The absolute conservation of Lys 289 throughout the superfamily and the observation that mutation of this Lys to Arg in aspartase (Saribas et al., 1994) completely abolishes the binding of the substrate to the enzyme strongly suggest that this residue is involved in stabilizing the enzyme−substrate complex, either by hydrogen bonding to one of the carboxylate moieties on the substrate or by stabilizing the negatively charged carbanion intermediate (Simpson et al., 1994). Alternatively, Lys 289 could be the catalytic acid as suggested by Weaver and Banaszak (1996).

Potential Role of His 91. Since a histidine residue had been implicated in the ASL/ δ crystallin catalytic mechanism and as the enzymatically inactive chicken and duck δ I crystallins were found to possess a histidine to glutamine mutation at position 91, it had previously been suggested that His 91 was directly involved in catalysis and that this mutation was the primary cause for the loss of enzymatic activity in δ I crystallin (Barbosa et al., 1991b). However, contrary to earlier reports (Patejunas et al., 1995) that the His 91 to Asn mutant (H91N) of duck δ II crystallin is completely inactive, our kinetic analyses show that the mutant in fact exhibits approximately 10% of the wild-type δ II crystallin activity. The Lineweaver−Burk plots (Figure 3) for the rates of the forward reaction as catalyzed by the H91N mutant and wild-type δ II crystallin are linear with no evidence of any negative cooperativity. The kinetic parameters, K_m and V_{max}, calculated from these plots are 0.83 ± 0.11 mM and 1.54 ± 0.14 μmol of fumarate formed min^{-1} (mg of protein)$^{-1}$ and 8.25 ± 1.64 mM and 1.37 ± 0.06 μmol of fumarate formed min^{-1} (mg of protein)$^{-1}$ for wild-type and the H91N mutant, respectively. While the two proteins exhibit comparable maximal velocities, their Michaelis constants for argininosuccinate differ significantly, with a 10-fold increase being observed in the K_m value for the H91N mutant. As there are no significant structural differences relative to the wild-type protein (Patejunas et al., 1995), the kinetic data suggest that the low levels of enzymatic activity exhibited by this mutant are the result of impaired substrate binding and implicate histidine 91 in a substrate binding role rather than actual catalysis.

The location of His 91 in the active site (Figure 5) supports the hypothesis that His 91 is involved in substrate binding and not directly in catalysis as the residue is ∼13.8 Å away from His 162. Given this distance and provided that there are no large structural changes in this area upon substrate binding, it appears unlikely that His 91 could participate directly in the catalytic mechanism. However, it is possible that His 91 could interact with the peptidic end of the substrate. Argininosuccinate, a conformationally flexible molecule, could lie across the active site cleft such that the fumarate moiety is oriented toward His 162 and the conserved regions c2 and c3, while the $-NH_3^+$ and $-COO^-$ groups of the arginine moiety are oriented toward His 91. While this hypothesis has yet to be confirmed with the structure of an enzyme−inhibitor complex, it is consistent with our kinetic data, as the Michaelis constant of the H91N mutant is the only kinetic parameter affected by the mutation. As His 91 is not conserved across the superfamily, but is only conserved in ASL's and enzymatically active δ II crystallins, we postulate that its role is in defining the substrate specificity of these members of the enzyme superfamily.

In addition, these kinetic results are now consistent with early studies of the human ASL protein, where histidine 89 (equivalent to histidine 91 in duck δ II crystallin) was mutated to a glutamine residue. This H89Q ASL mutant was also found to exhibit 10% of wild-type activity (Barbosa et al., 1991a). The mutation of histidine 91 to glutamine in the enzymatically inactive δ I crystallins could therefore account for a significant portion, perhaps as much as 90%, of the loss of catalytic activity. However, since this mutation is just one of the 28 amino acids that differ between the enzymatically inactive and active forms of the duck eye lens protein, other point mutations (Figure 1b) or conformational changes as a consequence of the point mutations must also contribute to the inactivation of the protein.

Structural Differences between δ I and δ II Crystallin. δ I crystallin is enzymatically inactive, despite the presence of all the residues implicated in the catalytic mechanism (Figure 5) and the preservation of almost all of the highly conserved residues across the superfamily (Figure 1a). The only variation in the conserved regions is a Glu to Asp mutation at position 117 between δ I and δ II crystallin, respectively. The replacement of His 91 with Gln in δ I crystallin could explain, in part, the loss of its catalytic activity since His 91 appears to be crucial for substrate binding. However, unlike the H91N mutant of δ II crystallin, δ I crystallin is completely inactive, and therefore additional factors have to be responsible for the lost catalytic activity. A structural comparison of the enzymatically inactive turkey δ I crystallin (Simpson et al., 1994) and the duck H91N δ II crystallin mutant was therefore undertaken. While this comparison should ideally be between the wild-type duck δ I and δ II proteins, as this would eliminate any structural differences that may occur between the species that do not relate to the loss of enzymatic activity, the structures for these proteins are unavailable at present. The resolution of the data available for the wild-type duck δ II structure is considerably less than that of the H91N mutant (3.0 vs 2.5 Å). However, comparison of the structures of the wild-type δ II and the H91N mutant shows no major structural rearrangements as a result of the H91N mutation. The comparison of the H91N mutant structure with that of turkey δ I crystallin should therefore yield results similar to a comparison of the wild-type duck δ II protein and the turkey structure. It should be noted that the sequence of turkey δ I crystallin is unknown and that the sequence of chicken δ I crystallin (Figure 1b) has been used in the turkey δ I model. It is expected that the sequence identity between chicken and turkey δ I crystallin will be greater than 90% (Simpson et al., 1994).

A comparison of the main chain positions of the individual monomers of each structure to every monomer in the second structure reveals an average rms deviation of 0.56 ± 0.02 Å. This value is slightly larger than the average rms deviations seen between each of the individual monomers of the two structures. The average rms deviation for all monomers is 0.29 ± 0.05 and 0.27 ± 0.04 Å for the duck and turkey crystallins, respectively. The small deviations observed for the NCS related molecules reflect the fact that strict noncrystallographic symmetry was not maintained in the final stages of refinement. While most of the differences are the result of small changes in the backbone conformation, two regions in each monomer differ by more than 1 Å. These regions are adjacent in the structure and include residues 25−32 and residues 76−91. The backbone of δ II crystallin has shifted approximately 5.0 Å and is between 3 and 4 Å relative to the δ I structure for the 25−32 and 76−91 regions, respectively. In each of the regions there are two nonconservative sequence changes: Ile (δ I) to Lys 25 (δ II), Thr (δ I) to Tyr 32 (δ II), Ser (δ I) to Trp 76 (δ II), and Leu (δ I) to Phe 81 (δ II). Each of these changes involves the substitution of a small side chain (δ I) by a larger, bulkier one (δ II), with the introduction of a positive charge at position 25. In order to accommodate the substitutions, the

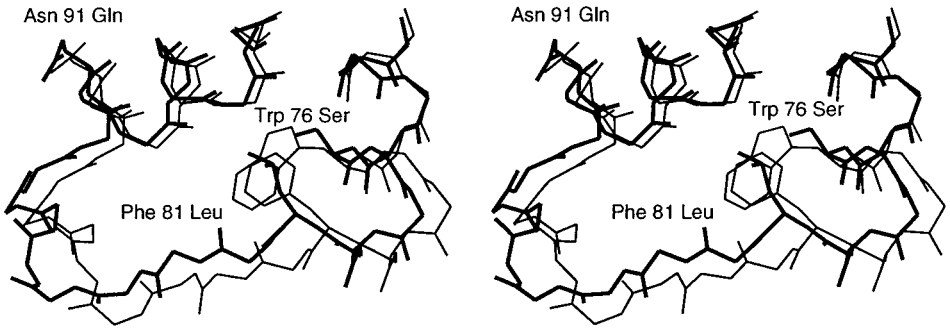


FIGURE 6: Stereoview of the 76−91 loop for turkey δ I (thick lines) and duck δ II (thin lines) crystallin. Side chains for residues 76, 81, and 91 have been included as these residues are not conserved between the δ I and δ II sequences (see Figure 1b).

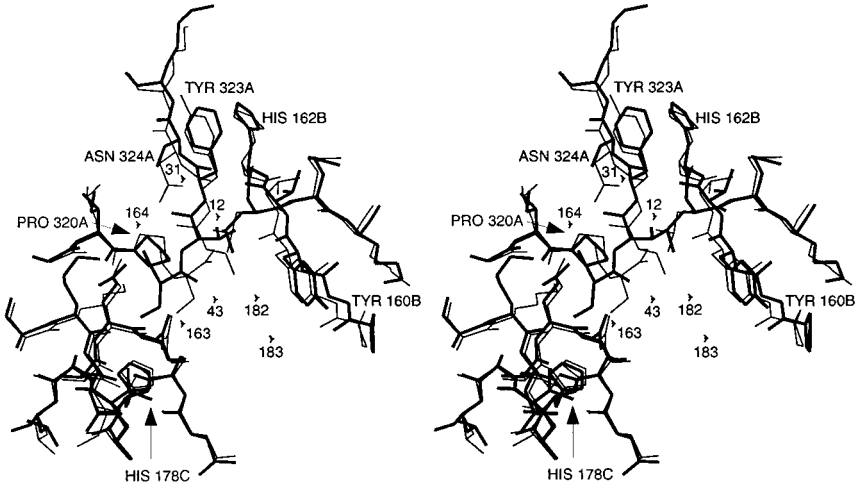


FIGURE 7: Stereoview of the loop containing residues 320−324 and the surrounding residues for turkey δ I (thick lines) and duck δ II (thin lines) crystallin.

backbone conformation of the δ II structure has had to change.

Both structures have rather disordered N-termini. The conformational change for residues 25−32 may therefore just be the result of the flexibility of the region as indicated by the higher than average thermal parameters. Examination of residues 76−91, however, shows that residues Trp 76 and Phe 81 in δ II crystallin (Figure 6) are involved in a ring-stacking interaction, and in order to accommodate the bulkier side chains at these locations, the backbone has shifted downward. This shift does not alter significantly the position of residue 91, where the His → Asn mutation has been introduced to the sequence of duck δ II crystallin. The electron density for region 76−91 is also noisy but unequivocally shows the changes between the two structures. Since both of these regions lie at the lower corner of the putative active site (Figure 5) and are quite distant from the highly conserved residues, it is not immediately obvious how these structural changes could affect the catalytic activity. However, it is interesting to note that both the active duck and chicken δ II crystallins contain the large aromatic residues at positions 76 and 81. Goose crystallin, which has been shown to be enzymatically active (Yu & Chiou, 1993), is more similar in amino acid sequence to duck δ I (97% identity) than to duck δ II (93.8% identity) crystallin. While only one of the two genes for goose δ crystallin has been sequenced, it is interesting to note that in the sequence available residues 76, 81, and 91 are tryptophan, phenylalanine, and histidine, respectively. These are identical to the residues found in the active duck and chicken δ II crystallin and hence support the suggestion that these residues are important in conferring enzymatic activity on the enzyme.

The only other significant difference between turkey and duck δ crystallin structures is a difference in the peptide linkage between residues Ser 321 and Thr 322 (Figures 2 and 7). In the duck δ II crystallin structure this peptide linkage is in the *cis* conformation while in the turkey δ I crystallin structure it is in the *trans* conformation. Given the infrequency with which non-proline cis peptides occur (Stewart et al., 1990) and their role in the catalytic activity of the protein when they do, it is tempting to suggest that this difference is the key to the loss of enzymatic activity in δ I crystallin. However, if the *cis* peptide does play a role in the enzymatic mechanism, it would have to be an indirect one as the residues are completely buried, although the 320−324 loop of one monomer does make a number of hydrogen bond contacts to the residues on the 159−168 loop (conserved region 2) of a second monomer both directly (Table 3) and via a network of water molecules. The 158−168 loop contains His 162, a residue that appears critical for catalytic activity. The 320−324 loop in which the *cis* peptide occurs is stabilized in duck δ II crystallin by a hydrogen bond between $N_{\delta 2}$ of Asn 324 and the main chain O of Pro 320. Asn 324 is mutated to serine in the turkey δ crystallin, and while the potential for a hydrogen bond does exist, the distance between the two groups (Ser 324A O_γ and O of Pro 320A) is 5.3 Å. The amino acid sequence in this region for duck δ I crystallin is identical to that of duck δ II crystallin (Figure 1b), and therefore duck δ I crystallin should be able to stabilize this loop in a manner similar to that seen

Table 3: Summary of Hydrogen Bond Contacts for Residues 320−324

		distance (Å)	
residue and atom	interaction	δ I crystallin	δ II crystallin
Pro 320A			
O	$N_{\delta 2}$ Asn 324A		2.7
Ser 321A			
N	$O_{\delta 1}$ Asp 261B	2.9	2.8
	$O_{\delta 2}$ Asp 261B	3.6	2.5
$O_{\gamma 1}$	$N_{\epsilon 2}$ His 178B	2.9	*a*
Thr 322A			
N	O Tyr 160B		2.9
$O_{\gamma 1}$	O Tyr 160B	2.9	3.5
Tyr 323A			
O	N His 162B		3.0
	O Thr 161B[b]	2.5	
Asn 324A			
$N_{\delta 2}$	O Pro 320A		2.7
O	N Asp 326A	3.0	2.9
	N Leu 327A	3.2	3.0

[a] While the $O_{\gamma 1}$ of Ser 321A makes no direct interaction with His 178B in duck δ II crystallin, it does makes an indirect inaction via water molecule 163 (2.77 Å) to the $O_{\gamma 1}$ of Ser 177B (distance O 163 to $O_{\gamma 1}$ Ser 177B 2.89 Å). [b] The difference in the hydrogen bonds between Tyr 323 (δ II) or Phe 323 (δ I) is due to the difference in the orientation of the peptide group between residues Thr 161B and His 162B. In the present orientation in turkey δ I crystallin, the interaction quoted is between two carbonyl oxygens and is therefore unfavorable. However, if the peptide group was flipped, the O of Phe 323 would make a good hydrogen bond to the N of His 162B.

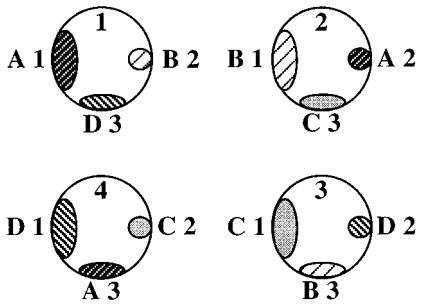


FIGURE 8: Schematic diagram of the four active sites of δ II crystallin. Each circle represents an active site and is numbered 1−4. In each active site the three conserved regions have been represented by a different shape and are shaded according to which monomer they belong. The labels indicate which region/domain (1−3) of which monomer (A−D) makes up the active site. For example, active site 1 is made from monomer A, conserved region 1; monomer B, conserved region 2; and monomer D, conserved region 3.

Table 4: Effect of δ I Monomers on the Number of Active Sites in a Heterotetrameric Mixture of δ I and δ II

ratio of δ I to δ II	3 conserved regions affected	2 conserved regions affected	1 conserved region affected
4:0	0	0	0
3:1	0	0	1
2:2	0	2 or 0[a]	2
1:3	1	2	3
0:4	4	4	4

[a] The exact number of native active sites that will be present will depend on how the individual monomers of each type (δ I or δ II) form the tetramer. Due to the symmetry of the system there are certain combinations that will yield two active sites while others will result in no native-like active sites. For example, if monomers A and B (or C and D) are δ I and conserved regions 1 and 2 affected the enzymatic activity, then two native active sites will be formed (see Figure 8). Similarly, the same applies for monomers A and D (or B and C), with conserved regions 1 and 3 affecting the activity. Other combinations of monomer and conserved regions will result in no native active sites being formed.

for duck δ II crystallin. Confirmation of the presence of a *cis* peptide in duck δ I crystallin will require the structure to be determined. However, this suggests that the structural variation seen in this region is the result of differences between the duck and chicken species and not a difference between the enzymatically inactive and active forms of the protein.

Role of Domain 1. In a study of embryonic duck lens by Piatigorsky and Horwitz (1996), it was discovered that five different isoforms of the δ crystallin protein exist. The five isoforms are the result of varying numbers of δ I and δ II monomers present in each tetramer (i.e., 4:0, 3:1, 2:2, 1:3, and 0:4, δ I to δ II crystallin monomers). Under saturating conditions the V_{max} for each isoform was directly related to the number of δ II monomers present in the isoform, with the tetramer of four δ II monomers being the most active and the tetramer with 4 δ I monomers being inactive. Examination of Figure 8 shows a schematic diagram of the four active sites that would be found in the protein, indicating which monomer (A−D) and which conserved region (1−3) contributes to each site. Using this diagram we can predict how many "native" active sites will be present in a heterotetrameric mixture of δ I and δ II as a function of the number of conserved regions or domains that could be affecting activity. For example, in the case where you have a ratio of 3 δ I to 1 δ II monomers in a tetramer and all three regions that make up the active site contribute to the loss of catalytic activity, then regions A1, A2, A3, B1, B2, B3, C1, C2, and C3 would all compromise the activity in some way and no active sites could be formed that did not contain at least one of these regions. If all three regions contribute to the loss of enzymatic activity, we would have expected this heterotetramer to be catalytically inactive. Since it was found to exhibit activity, all three regions cannot contribute to the loss of catalytic activity. A similar analysis can be performed assuming that either two regions or one of the regions that contribute to the active site is involved in conferring loss of enzymatic activity. The number of native active sites that would result from this analysis is represented in Table 4. What is obvious from Table 4 and the kinetic results of the five different isoforms is that the linear relationship between the number of δ I monomers and the reduction in catalytic activity can be best explained if sequence variations in only one of the three regions that constitute the active site affect the activity of the protein.

We postulate that the region/domain that confers the loss of enzymatic activity is the first domain in which we see the large conformation change in loop 76−91 and where the His → Gln mutation at position 91 is found. Our kinetic results suggest that the His → Gln mutation only affects the binding of the substrate and although, in the absence of a structure of the duck δ II protein with an inhibitor or substrate analogue bound, it is difficult to assess how the conformational changes seen in the 76−91 loop affect the catalytic activity, it is interesting to note that most of the amino acid differences between duck δ I and δ II crystallin do occur in this domain (Figure 1b). Since almost all residues that have been implicated in the catalytic mechanism and that lie in the conserved regions are conserved between the δ I and δ II crystallins and there are very few structural differences

that occur in these regions, this would suggest that the conformation of domain 1 and the amino acids in this region are critical for substrate binding and that therefore δ I crystallin is inactive because it can no longer bind to substrate.

ACKNOWLEDGMENT

The authors thank Bill O'Brien for providing the expression vectors of the wild-type and mutant duck δ II crystallins and Steve Ealick and his colleagues at CHESS for access to the facility.

REFERENCES

Abu-Abed, M., Turner, M. A., Atkinson, J., Kiran, D., & Howell, P. L. (1994) *J. Mol. Biol. 243*, 944−946.
Barbosa, P., Cialkowski, M., & O'Brien, W. E. (1991a) *J. Biol. Chem. 266*, 5286−5290.
Barbosa, P., Wistow, G. J., Cialkowski, M., Piatigorsky, J., & O'Brien, W. (1991b) *J. Biol. Chem. 266*, 22319−22322.
Blum, M., Metcalf, P., Harrison, S. C., & Wiley, D. C. (1987) *J. Appl. Crystallogr. 20*, 235−242.
Brünger, A. (1992) *Nature 355*, 472−475.
Brünger, A. T. (1990) *X-PLOR: A System for Crystallography and NMR*, Yale University Press, New Haven, CT.
Cambillau, C., Roussel, A., Inisan, A. G., & Knoops-Mouthuy (1996) *TURBO-FRODO Marseille*, Version 5.5 release b.
Chiou, S., Lee, H., Lai, T., & Chang, G. (1991a) *Biochem. Int. 25*, 705−713.
Chiou, S., Lo, C., Chang, C., Itoh, T., Kaji, H., & Samejima, T. (1991b) *Biochem. J. 273*, 295−300.
Delaye, M., & Tardieu, A. (1983) *Nature 302*, 415−417.
Fernald, R. D., & Wright, S. E. (1983) *Nature 301*, 618−620.
Flazone, C. J., Karsten, W. E., Conley, J. D., & Viola, R. (1988) *Biochemistry 27*, 9089−9093.
Garrard, L. J., Bui, Q. T. N., Nygaard, R., & Raushel, F. M. (1985) *J. Biol. Chem. 260*, 5548−5553.
Havir, E. A., Tamir, H., Ratner, S., & Warner, R. (1965) *J. Biol. Chem. 240*, 3079−3087.
Hendriks, W., Mulders, J. W., Bibby, M. A., Slingsby, C., Bloemendal, H., & de Jong, W. W. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 7114−7118.
Hodel, A., Kim, S.-H., & Brunger, A. T. (1992) *Acta Crystallogr. A48*, 851−859.
Hughes, A. (1994) *Proc. R. Soc. London, B 256*, 119−124.
Ida, N., & Tokushige, M. (1985) *J. Biochem. 98*, 793−797.
Jones, T. A., Zou, J.-Y., Cowan, S. W., & Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
Kabsch, W. (1988a) *J. Appl. Crystallogr. 21*, 67−71.
Kabsch, W. (1988b) *J. Appl. Crystallogr. 21*, 916−924.
Kabsch, W., & Sander, C. (1983) *Biopolymers 22*, 2577−2637.
Kim, R., & Wistow, G. J. (1993) *FASEB J. 7*, 464−469.
Kimura, M., & Ohta, T. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 2848−2852.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
Lee, H., Chiou, S., & Chang, G. (1992) *Biochem. J. 283*, 597−603.
Lee, H.-J., Chiou, S.-H., & Chang, G.-G. (1993) *Biochem. J. 293*, 537−544.
Mori, M., Matsubasa, T., Amaya, Y., & Takiguchi, M. (1990) *Prog. Clin. Biol. Res. 344*, 683−699.
Nickerson, J. M., Wawrousek, E. F., Hawkins, J. W., Wakil, A. S., Wistow, G. J., Thomas, G., Norman, B. L., & Piatigorsky, J. (1985) *J. Biol. Chem. 260*, 9100−9105.
Nickerson, J. M., Wawrousuk, E. F., Borras, T., Hawkins, J. W., Norman, B. L., Filupa, D. R., Nagle, J. W., Ally, A. H., & Piatigorsky, J. (1986) *J. Biol. Chem. 261*, 552−557.
O'Brien, W. E., & Barr, R. H. (1981) *Biochemistry 20*, 2056−2060.
O'Brien, W. E., McInnes, R., Kalumuck, K., & Adcock, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 7211−7215.
Otwinowski, Z. (1993) in *Data Collection and Processing, Proceedings of the CCP4 Study Weekend* (Sawyer, L., Isaacs, N., & Baily, S., Eds.) pp 56−62, SERC Daresbury Laboratory, Warrington, U.K.
Patejunas, G., Barbosa, P., Lacombe, M., & O'Brien, W. E. (1995) *Exp. Eye Res. 61*, 151−154.
Piatigorsky, J. (1984) *Cell 38*, 620−621.
Piatigorsky, J. (1992) *J. Biol. Chem. 267*, 4277−4280.
Piatigorsky, J. (1993) *Dev. Dyn. 169*, 196−272.
Piatigorsky, J., & Wistow, G. J. (1989) *Cell 57*, 197−199.
Piatigorsky, J., & Wistow, G. J. (1991) *Science 252*, 1078−1079.
Piatigorsky, J., & Horwitz, J. (1996) *Biochim. Biophys. Acta 1295*, 158−164.
Piatigorsky, J., O'Brien, W. E., Norman, B. L., Kalumuck, K., Wistow, G. J., Borras, T., Nickerson, J. M., & Wawrousek, E. F. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 3479−3483.
Piatigorsky, J., Kantorow, M., Gopal-Srivastava, R., & Tomarev, S. I. (1994) in *Toward a Molecular Basis of Alcohol Use and Abuse* (Jansson, B., Jörnvall, H., Rydberg, U., Terenius, L., & Vallee, B. L., Eds.) pp 241−250, Birkhäuser Verlag, Basel, Switzerland.
Read, R. J. (1986) *Acta Crystallogr. A42*, 140−149.
Saribas, A. S., Schindler, J. F., & Viola, R. E. (1994) *J. Biol. Chem. 296*, 6313−6319.
Simpson, A., Bateman, O., Dreissen, H., Lindley, P., Moss, D., Mylvaganam, M., Narebor, E., & Slingsby, C. (1994) *Nat. Struct. Biol. 1*, 724−734.
Stewart, D. E., Sarkar, A., & Wampler, J. E. (1990) *J. Mol. Biol. 214*, 253−260.
Stone, R. L., Zalkin, H., & Dixon, J. E. (1993) *J. Biol. Chem. 268*, 19710−19716.
Takagi, J. S., Tokushige, M., Shimura, Y., & Kanehisa, M. (1986) *Biochem. Biophys. Res. Commun. 138*, 568−572.
Turner, M. A., Simpson, A., McInnes, R. R., & Howell, P. L. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 9063−9068.
Weaver, T., & Banaszak, L. (1996) *Biochemistry 35*, 13955 - 13965.
Weaver, T. M., Levitt, D. G., Donnelly, M. I., Wilkens-Stevens, P. P., & Banaszak, L. J. (1995) *Nat. Struct. Biol. 2*, 654−662.
Weaver, T., Lees, M., & Banaszak, L. (1997) *Protein Sci. 6*, 834−842.
Williams, S. E., Woolridge, E. M., Ransom, S. C., Landro, J. A., Babbitt, P. C., & Kozarich, J. W. (1992) *Biochemistry 31*, 9768−9776.
Wistow, G. (1993) *Trends Biol. Sci. 18*, 301−306.
Wistow, G., & Piatigorsky, J. (1987) *Science 236*, 1554−1556.
Wistow, G. J., & Piatigorsky, J. (1988) *Annu. Rev. Biochem. 57*, 479−504.
Wistow, G. J., & Piatigorsky, J. (1990) *Gene 96*, 263−270.
Woods, S. A., Miles, J. S., Roberts, R. E., & Guest, J. R. (1986) *Biochem. J. 237*, 547−557.
Woods, S. A., Schwarzbach, S. D., & Guest, J. R. (1988) *Biochim. Biophys. Acta 954*, 14−26.
Yu, C.-W., & Chiou, S.-H. (1993) *Biochem. Biophys. Res. Commun. 192*, 948−953.

BI971407S