Crystal Structure of an Inactive Duck δ II Crystallin Mutant with Bound Argininosuccinate[†,‡]

Francois Vallée,[§] Mary A. Turner,[§] Peter L. Lindley,[||,⊥] and P. Lynne Howell*[,§,#]

Structural Biology and Biochemistry, Research Institute, 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 CCLRC, Daresbury Laboratory, Daresbury, Warrington, Cheshire WA4 4AD, United Kingdom

Received September 4, 1998; Revised Manuscript Received November 4, 1998

ABSTRACT: δ-Crystallin, the major soluble protein component of avian and reptilian eye lenses, is highly homologous to the urea cycle enzyme, argininosuccinate lyase (ASL). In duck lenses, there are two highly homologous δ crystallins, 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 is enzymatically inactive. The X-ray structure of a His to Asn mutant of duck δ II crystallin (H162N) with bound argininosuccinate has been determined to 2.3 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 formed by three different monomers in the tetramer. The active site of the H162N mutant structure reveals that the side chain of Glu 296 has a different orientation relative to the homologous residue in the H91N mutant structure [Abu-Abed et al. (1997) *Biochemistry 36*, 14012−14022]. This shift results in the loss of the hydrogen bond between His 162 and Glu 296 seen in the H91N and turkey δ I crystallin structures; this H-bond is believed to be crucial for the catalytic mechanism of ASL/δ II crystallin. Argininosuccinate was found to be bound to residues in each of the three monomers that form the active site. The fumarate moiety is oriented toward active site residues His 162 and Glu 296 and other residues that are part of two of the three highly conserved regions of amino acid sequence in the superfamily, while the arginine moiety of the substrate is oriented toward residues which belong to either domain 1 or domain 2. The analysis of the structure reveals that significant conformational changes occur on substrate binding. The comparison of this structure with the inactive turkey δ I crystallin reveals that the conformation of domain 1 is crucial for substrate affinity and that the δ I protein is almost certainly inactive because it can no longer bind the substrate.

Crystallins are soluble proteins found in vertebrate eye lenses where they are believed to play a structural role by contributing to its bulk protein framework and conferring special refractive properties to it (*2*−*4*). Crystallins which constitute a group of proteins with very little homology among its members (*5*) have been divided in two classes according to their distribution in vertebrate species. The ubiquitous crystallins (α-, β-, and γ-crystallins) are found in all vertebrates while the taxon-specific enzyme crystallins (δ, ε, λ, etc.) are restricted in distribution to certain species. Lens proteins are not renewed and are therefore among the longest lived proteins in vertebrates, resisting aggregation and degradation and maintaining their structure for a long period of time. Given these properties and the role that they play, it was surprising to discover that the taxon-specific crystallins are not specialized lens proteins but are directly related to housekeeping enzymes (*6*−*8*), and in some cases the crystallins have even been shown to demonstrate in vitro the native activity of their homologue (*9*−*11*).

δ-Crystallin is the principal constituent of bird and some reptile eye lenses and accounts for 60−70% of all the soluble lenticular protein. In the chicken and duck genomes, two distinct functional δ-crystallin genes, termed δ I and δ II, are present in tandem (*12*−*14*). Comparison of the protein sequences for δ I and δ II reveals 88−94% identity among the different δ-crystallins and 65−71% identity to argininosuccinate lyase (ASL,[1] EC 4.3.2.1) (*15*−*17*). The two crystallin isoforms are the products of a gene duplication event of the ancestral ASL gene. Despite their high sequence identity and the similarity between these crystallins and human ASL, ASL activity has only been detected in the δ II isoform. The loss of enzymatic activity observed for all the δ I crystallins characterized to date must be the result of sequence variations between the two isoforms. These se-

[†] Supported by a grant from the Medical Research Council of Canada to P.L.H.

[‡] PDB accession number 1DCN.

* Corresponding author: (416) 813-5378 (tel); (416) 813-5022 (fax); howell@sickkids.on.ca (e-mail).

[§] Hospital for Sick Children.

[||] Daresbury Laboratory.

[⊥] Present address: ESRF, BP-220, 41 Avenue des Martyrs, 38043 Grenoble Cedex 09, France.

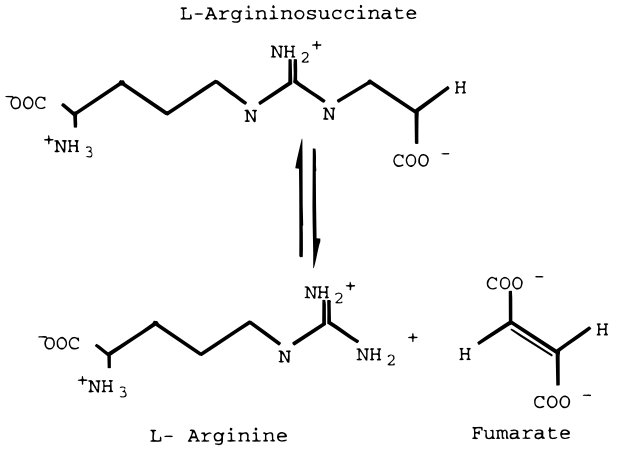
[#] University of Toronto.

[1] Abbreviations: AS, argininosuccinate; ASL, argininosuccinate lyase; H91N, histidine 91 to asparagine mutant of duck δ II crystallin; H162N, histidine 162 to asparagine mutant of duck δ II crystallin; TDCI, turkey δ I crystallin; FUMC, *Escherichia coli* fumarase C; NCS, noncrystallographic symmetry.

10.1021/bi982149h CCC: $18.00
Published on Web 02/05/1999

quence variations could either affect residues involved in substrate binding and catalysis or residues important for maintaining the catalytically active conformation of the protein.

δ-Crystallin and ASL also belong to a superfamily of metabolic enzymes that all catalyze β-elimination reactions with the release of fumarate. Members of this superfamily include ASL, class II fumarase, adenylosuccinate lyase, 3-carboxy-*cis*,*cis*-muconate lactonizing enzyme, aspartase, and δ-crystallin. A general acid/base mechanism has been proposed for δ II crystallin/ASL with a histidine residue implicated as the general base (*9*, *18*) in the interconversion of argininosuccinate to arginine and fumarate:



:

Since in the enzymatically inactive δ I crystallins His 91 is mutated to Gln, it was previously suggested that His 91 played a direct role in the catalysis (*19*). Four single site-directed mutants in which the histidine residues at positions 91, 110, 162, and 178 were replaced with asparagine residues were subsequently constructed to test this hypothesis and to identify the catalytic histidine (*20*). Preliminary kinetic analysis of these mutants showed that the H91N and H162N mutants were completely inactive, thus seemingly confirming the catalytic role proposed for His 91. Subsequent studies on H91N, however, revealed that this δ-crystallin mutant exhibited low levels of activity ($\sim$10% of wild type), thus identifying His 162 as the most likely candidate for the catalytic base (*1*). Studies on *Escherichia coli* fumarase C (FUMC) have also shown that the equivalent histidine, His 188, is close to the active site, and Weaver et al. (*21*, *22*) have also proposed that this histidine plays a role in base catalysis. In this paper, we present the crystallization, three-dimensional structure determination, and refinement of the inactive duck δ II crystallin mutant H162N complexed with the substrate argininosuccinate. This represents the first structure of an inhibitor or substrate analogue bound to ASL/δ II crystallin or one of their mutants. The analysis of this structure and its comparison with the structures of H91N and turkey δ I crystallin (TDCI) have provided the following information: (1) the structural mapping of the residues involved in substrate binding, (2) a proposal for the mechanism of inactivation of the H162N protein, and (3) a confirmation, in part, of the hypothesis concerning the enzymatic inactivity of the δ I protein (*1*).

Table 1: Data Collection and Refinement Statistics

resolution limits of data (Å)	17−2.30
total no. of measurements	194258
no. of unique reflections	77291
completeness of data (%)	90.3 (96.4)[a]
R_{sym}[b]	7.9 (23.3)[a]
resolution range for refinement (Å)	17−2.30
no. of reflections in refinement	71225
no. of reflections used to compute R_{free}	7210
total no. of atoms	13310
total no. of solvent sites	476
R_{factor},[c] all data (%)	22.9
R_{free}[d] (%)	29.0
mean *B* values (Å^2)	
entire protein	44
monomer A, B, C, D	44/42/48/43
domain 1	71[e]/65[e]/58/47[f]
domain 2	33/31/36/32
domain 3	50/48/71/66
waters	39
substrate	58
rms deviation from ideality	
bond lengths (Å)	0.006
bond angles (deg)	1.2
dihedral angles (deg)	17.6
improper torsion angles (deg)	0.74

[a] Last resolution shell: 2.38−2.30 Å. [b] $R_{sym} = \sum|I - \langle I\rangle|/\sum I$, where I is the intensity measurement for symmetry-related reflections and $\langle I\rangle$ is the mean intensity for the reflection. [c] $R_{factor} = \sum(|F_o| - |F_c|)/\sum|F_o|$. [d] $R_{free} = \sum(|F_{o,s}| - |F_{c,s}|)/\sum|F_{o,s}|$, where s refers to a subset of data not used in refinement comprising 10% of the data. [e] Residues 17A−31A and 73B and 87B are not included in the model. [f] Average *B* values for domain 1, 2, or 3 for each monomer A/B/C/D.

EXPERIMENTAL PROCEDURES

Gene Expression, Protein Purification, Crystallization, and Data Collection. Duck δ II crystallin derived from the cDNA of duck (*Anas platyrhonchos*) eye lens (gift of W. E. O'Brien, Baylor College of Medicine, Houston, TX) was over-expressed in *E. coli* strain BL21(DE3)pLysS using a T7 polymerase system. Details of the expression of the wild type and the H162N mutant of duck δ II crystallin have been described previously (*19*, *20*). The procedure for protein purification and crystallization is the same as that described previously for H91N mutant (*1*). The crystals were grown using the hanging drop vapor diffusion method. Five microliter drops of a 14 mg mL^{-1} enzyme solution, 100 mM Tris-HCl (pH 8.5), and the precipitating solution, 18% (w/v) PEG MME 2K, 200 mM $MgCl_2$, and 100 mM Tris-HCl (pH 8.5), were suspended over a 1 mL reservoir of the precipitating solution. Platelike crystals grew within approximately 4 days to a maximal size of 0.7 mm $\times$ 0.5 mm $\times$ 0.2 mm. The crystals belong to space group $P2_1$ with four protein molecules per asymmetric unit and with unit cell dimensions $a = 94.58$ Å, $b = 99.59$ Å, $c = 107.14$ Å, and $\beta = 101.68°$. Crystals of the complex were obtained by soaking a crystal for 24 h in a drop containing 10 μL of well buffer and 1 μL of 40 mM argininosuccinate. Prior to data collection, a crystal was further soaked in a 15% v/v glycerol/mother liquor solution for 5 min and then flash frozen to 100 K. X-ray diffraction data were then collected using a Mar Research image plate (18 cm) detector at station 7.2 at the CCLRC Daresbury Laboratory synchrotron (U.K.). The crystals diffracted at least to 2.3 Å resolution, and the data measured were processed using the DENZO/SCALEPACK program package (*23*). The final data reduction statistics are listed in Table 1.

Structure Determination and Refinement. The structure of H162N was solved with the molecular replacement program package AMoRe (*24*) using a monomer of the H91N crystal structure (*1*) as a search model. Rotation functions were calculated over a resolution range of 10−4 Å, with a Patterson radius of 20 Å. The final molecular replacement solution with four monomers in the asymmetric unit gave a correlation coefficient of 61.7% and an *R*-factor of 32.7%.

The structure was refined using CNS (*25*−*27*) with a maximum likelihood target function, a flat bulk solvent correction and no low resolution or σ cutoff applied to the data. Ten percent of the structure factor amplitudes were randomly selected and excluded from the refinement and used to compute a free *R* (R_{free}). After each refinement step, $2|F_o| - |F_c|$ and $|F_o| - |F_c|$ electron density maps were computed. Fourier components were weighted to reduce model bias from an incomplete or partially incorrect structure with the program SigmaA (*28*). Corrections to the model were made after each round of refinement with the program TURBO−FRODO (*29*). Structure refinement consisted of 19 cycles of a combination of torsion angle refinement, Cartesian refinement, and conjugate gradient energy minimization. At the end of each round of refinement, grouped and individual *B*-factor refinement was performed. The refinement protocol reduced the crystallographic *R*-factor and *R*-free [R_{factor} (R_{free})] from 36.1% (36.8%) to 22.9% (29.0%) in the resolution range 17−2.3 Å.

During the refinement, the peptide linkage between residues Ser 321 and Thr 322 of each monomer was built as a cis peptide, as previously described in the H91N crystal structure (*1*). Noncrystallographic symmetry (NCS) restraints were applied separately to each domain of the four monomers with weights chosen to minimize the R_{free} value. As the difference between R_{factor} and R_{free} decreased during rounds 1−6 of the refinement, the NCS restraints were relaxed and were completely released after round 6. During rounds 6−10, manual corrections were carried out independently on the N-terminal segment and residues 276−281 and 291−293 of each of the four monomers. A total of 476 ordered water molecules obeying proper hydrogen-bonding conditions with electron densities greater than 1.0σ on a $2|F_o| - |F_c|$ and 2.5σ on a $|F_o| - |F_c|$ SigmaA-weighted maps were gradually included, and an additional stage of Cartesian refinement was added to the refinement protocol in order to refine the position of the water molecules which are kept fixed during torsion angle refinement. Argininosuccinate, built using QUANTA (Molecular Simulations Inc., 1996) and XPLO2D (*30*), was modeled into the electron density during the final stages of refinement (rounds 11−19). The final model comprises 13290 protein atoms, 20 inhibitor atoms, and 476 solvent molecules. The stereochemical quality of the model was checked by PROCHECK (*31*), and the Ramachandran plot showed that 90.3% of the residues lie in the most favored region and that no residues reside in the disallowed regions. The final refinement parameters are shown in Table 1.

Structural Alignment and Comparison. Structural alignment was carried out by superimposing either the H91N or TDCI structure on the H162N δ II crystallin structure or each of the catalytic sites B, C, and D of the H162N structure on catalytic site A using the RIGID option in TURBO−FRODO (*29*). The first step is a rough rigid body alignment based on the coordinates of eight structurally equivalent atoms located at the beginning and end of each helix of domain 2. Starting from this position, an iterative least-squares fitting procedure was carried out between all the Cα atoms lying at a distance progressively decreasing from 5.0 to 0.1 Å. After this fitting procedure, the Cα rms deviations of all the Cα atoms were calculated in TURBO−FRODO. It should be noted that while the data for the H162N mutant were collected at liquid nitrogen temperature and those for H91N and TDCI were collected at 273 K, we do not think that the difference in temperature has influenced the structural comparison presented here, as the structural differences found between H162N and TDCI are similar to the differences found previously between H91N and TDCI (*1*) and the root-mean-square deviation between H162N and H91N is only 0.54 Å.

RESULTS AND DISCUSSION

Overall Architecture. H162N δ II crystallin has the same overall global fold as turkey δ I crystallin (TDCI) (*32*), H91N duck δ II crystallin (*1*), human ASL (*33*), and to a lesser extent FUMC (*22*) and *E. coli* L-aspartate ammonia-lyase (*34*). While each monomer should consist of 463 amino acids, the first 16 amino-terminal residues of monomers B, C, and D, the first 31 residues of monomer A, segment 76−89 in monomer B, and a variable number of residues located in the region 276−293 of each monomer have not been included in the model as the electron density for these residues was of insufficient quality for the conformation to be determined, presumably due to their conformational flexibility. The secondary structure of the protein is predominantly helical with a total of 21 helices per monomer. There are three distinct domains per subunit (Figure 1a). 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 them 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 1b). Two monomers associate via helices h8, h11, and h12 (Figure 1c) while the tetramer is the result of the association of two such dimers. The central h12 helices of each monomer form a four-helix bundle at the core of the protein. The overall tetramer is held together by mainly hydrophobic interactions along the length of the three helices h8, h11, and h12. The overall tetramer exhibits 222 symmetry.

Catalytic Site and Loss of Catalytic Activity in the H162N Mutant. The location of the active site was first identified in the structure of TDCI (*32*) and later confirmed by several structural studies on FUMC (*21, 22, 35*). The most highly conserved stretches of amino acid sequence across the superfamily (denoted c1, residues 114_1-121_1; c2, residues 159_2-168_2; and c3, residues 282_3-296_3) (see Figure 1c) which are dispersed in each monomer (Figure 1a) are brought together in the tetramer to form the active site cleft (see Figure 1b). No major structural differences were found to exist between the H162N and H91N crystal structures (*1*) as pairwise superposition of the Cα atom of both structures yields a root-mean-square (rms) deviation of only 0.54 Å. Significant differences were found, however, in the confor-

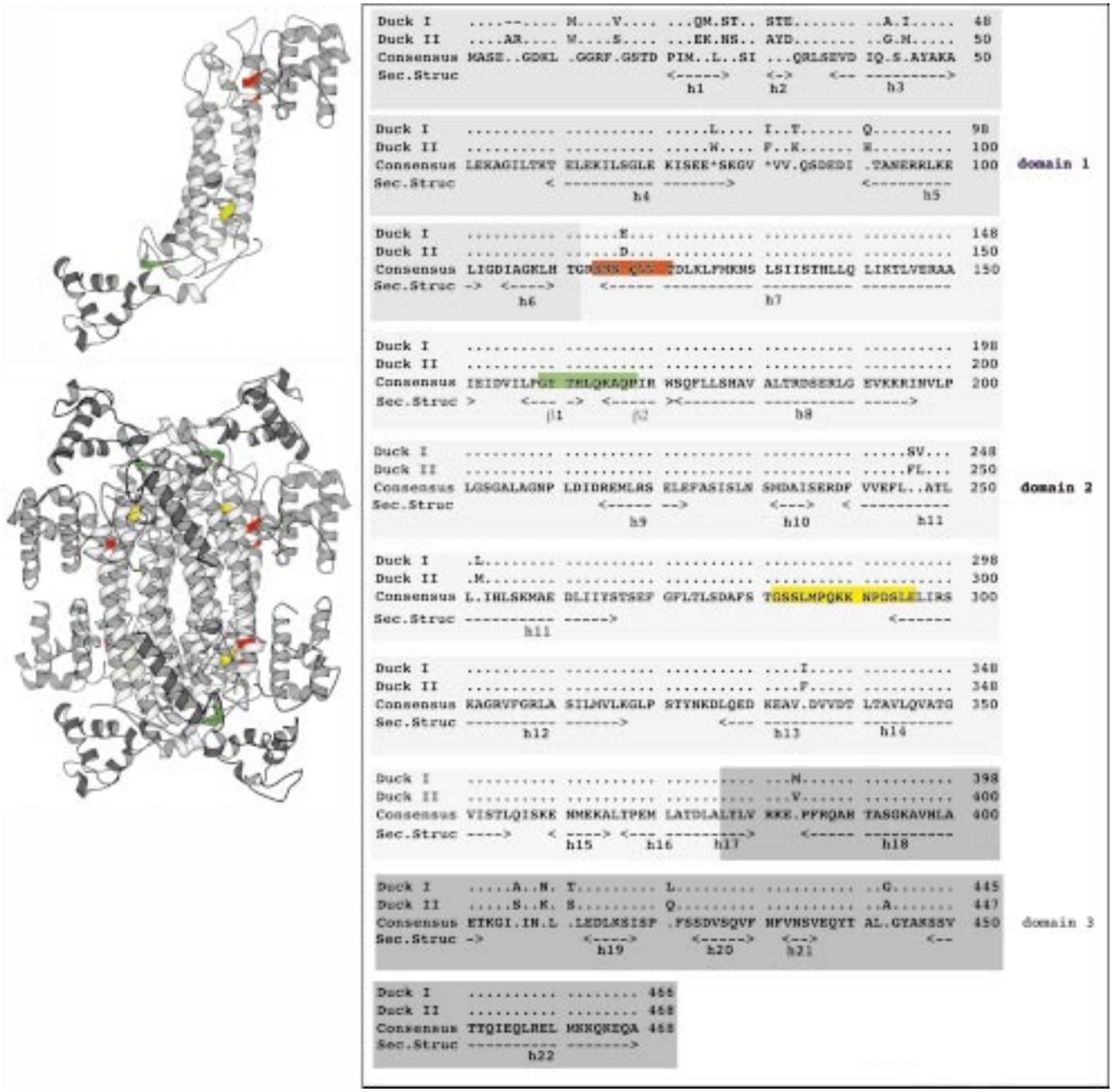


FIGURE 1: Schematic diagram showing the three-dimensional topology of the duck δ II crystallin (a, top left) monomer and (b, bottom left) tetramer as well as (c, right) its primary structure compared to the inactive duck δ I crystallin. (a) Schematic diagram of the three-dimensional topology of a monomer. (b) Schematic diagram of the tertiary structure of the tetramer of H162N δ II crystallin. (c) Complete sequence alignment of duck δ I and δ II crystallin. Key to sequence alignment: Duck I, duck δ I crystallin; Duck II, duck δ II crystallin; Sec.Struc., secondary structure (h*x*, helix number *x*; β*x*, β-sheet number *x*). In the three panels, domain 1 is colored in medium gray, domain 2 in light gray, and domain 3 in dark gray, and the highly conserved regions across the superfamily are colored according to their localization: region c1, red; region c2, green, and region c3, yellow. Panels a and b were prepared using the program Molscript (*42*).

mation of the side chains of residues located in each of the four active sites.

During the initial rounds of refinement, the sigmaA-weighted Fourier difference electron density maps clearly indicated that the orientation of the side chain of the mutated residue, Asn 162, was significantly altered relative to the position of His 162 found in the H91N structure. This difference in the conformation is the result of a hydrogen bond interaction between Asn 162_2 Oδ1 and Lys 325_3 N (see Figure 2).[2] Although H91N is a mutant structure, residue 91 is located ~14.5 Å from residue 162, and therefore this mutant can be considered as "native" in the active site region. In the H91N structure, the Nε1 of His 162_2 interacts with Glu 296_3 Oε1. In the H162N mutant, it is clear that the average 1.15 Å shift of the mutated side chain (measured between the Cγ atoms of the superimposed side chains) results in a rotation of the side chain of Glu 296_3 which now interacts with Pro 292_3 O, Arg 299_3 Nη2, Asn 162_2 O, Asp

[2] Notations: Each of the four catalytic sites is made up of residues from three different monomers (termed monomers 1, 2, and 3). The four active sites can be formed without distinction by monomers A, B, and D; B, A, and C; D, C, and A; or C, D, and B. The four active sites have been labeled A, B, C, and D according to which monomer contributes residue 162, i.e., active site A contains Asn 162_A, etc. The notation X_i indicates that the residue X belongs to monomer *i* with i = 1, 2, or 3 or, more precisely if necessary, with i = A, B, C, or D.

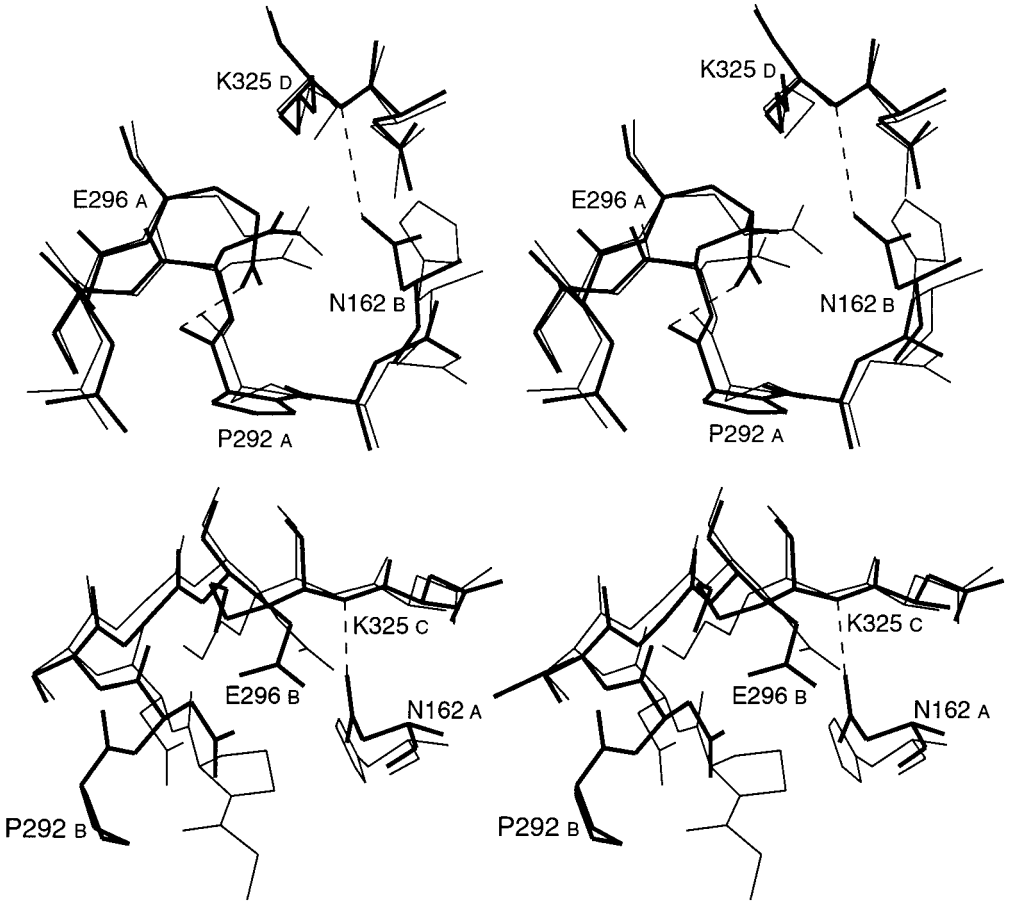


FIGURE 2: Stereoview of the catalytic site for H162N (thick lines) and H91N (thin lines) duck δ II crystallin for active sites A and D (a, top) and B and C (b, bottom), respectively. The interactions between Asn 162_2 Oδ1 and Lys 325_3 N and between Glu 296_3 Oδ1 and Pro 292_3 O are represented by black dashed lines. This figure was prepared with the program TURBO−FRODO (*29*). For clarity, in this and the following figures, side chains are labeled with the appropriate one-letter code and amino acid number.

293_3 O, and/or Asn 291_3 Oδ1. The last interaction (Asn 291_3) is only found in monomer A. It remains unclear why the interaction between Asn 162_2 Oδ1 and Lys 325_3 N is preferred to the potential interaction between Asn 162_2 Oδ1 and Glu 296_3 Oϵ1 which would mimic the interaction between His 162_2 Nϵ1 and Glu 296_3 Oϵ1 found in the H91N and the homologuous TDCI and FUMC structures. The His 162_2−Glu 296_3 interaction is believed to be crucial for the catalytic mechanism of ASL/δ II crystallin. Glu 296_3 is involved in a charge relay system with His 162_2, which as a result of the interaction is rendered more nucleophilic and has thus been proposed to be the catalytic base involved in initiating the reaction (*1, 22*).

Conformational Variation of the 280's Loop. The reorientation of Glu 296_3's side chain in the H162N structure is closely related to conformational changes observed in the highly conserved region c3, residues 282−296. This region is believed to be important in the catalytic reaction. The absolute conservation of Lys 289 throughout the superfamily and the observation that mutation of this Lys to Arg in aspartase (*36*) completely abolishes the binding of the substrate to the enzyme suggest that this residue is involved in stabilizing the enzyme−substrate complex either by hydrogen bonding to one of the COO^- groups of the fumarate moiety or by stabilizing the negatively charged carbanion intermediate. The evidence of the importance of this region is also shown by the disease causing Q286R mutation in human ASL (288 in δ II), which exhibits only 0.01% of wild-type activity (*33, 37*). In each of the four active sites, a variable number of residues in the loop 276−292 could not be located, a result of their conformational flexibility. After complete relaxation of the NCS restraints it was possible to build independently residues in each monomer belonging to segments 276−281 and 290−292, which represent the lower and upper part of the loop, respectively, and to characterize two different types of conformations for segment 292−294 (see Figure 2).

The first conformation of the 280's loop is seen in active sites B and C (containing Asn 162_B and Asn 162_C, respectively). In active sites B and C, Glu $296_{A(D)}$ Oϵ1 makes a hydrogen bond with Pro $292_{A(D)}$ O (Figure 2a), which is shifted approximately 1.0 Å relative to its position in the H91N structure. This interaction, not found in the H91N structure, is the consequence of the Glu 296 side-chain reorientation. The second conformation for the 280's loop (Figure 2b) is seen in active sites A and D (containing Asn 162_A and Asn 162_D, respectively). A complete reorientation of Pro 292_B and Pro 292_C abolishes the interaction observed previously between Pro 292 and Glu 296 in active sites B and C (see Figure 2a). The increased thermal motion observed for the lower part of both loops, which can only be modeled to residues 276_B and 278_C, respectively, suggests that the new conformation of this proline induces a much larger conformational change in the 280's loop than the conformation of the proline seen in active sites B and C.

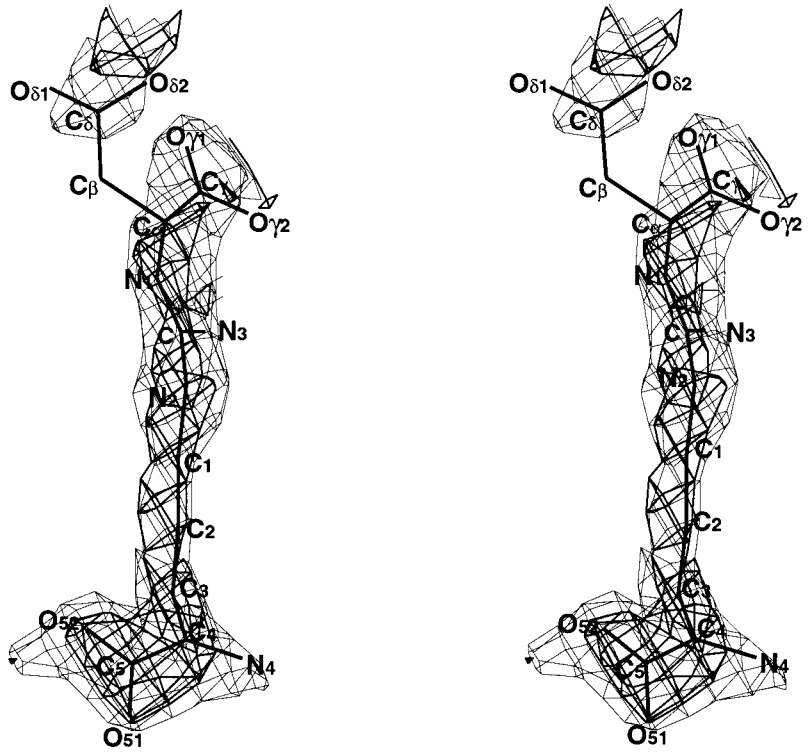


FIGURE 3: Stereoview of the substrate argininosuccinate with the final $2F_o - F_c$ ($+1\sigma$, thin lines) map obtained after the last round of refinement and the initial $F_o - F_c$ ($+2\sigma$, thick lines) omit map calculated after round 1 of refinement.

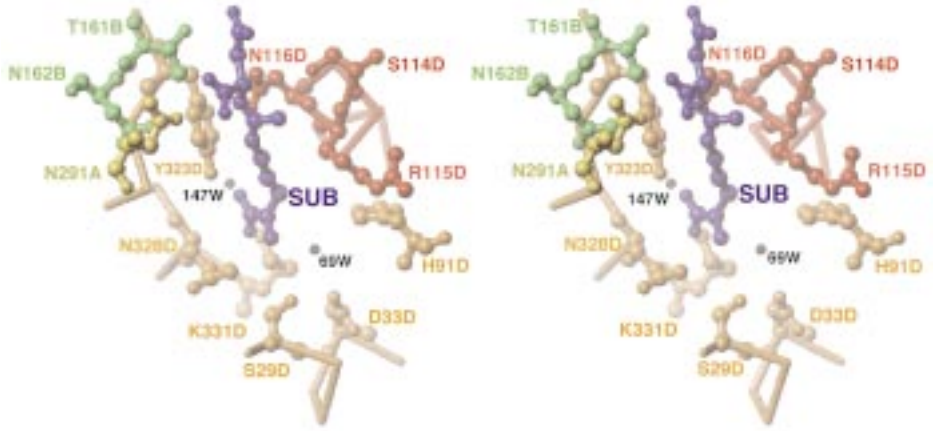


FIGURE 4: Stereoview of the argininosuccinate binding site formed by monomers A, B, and D. Regions c1, c2, and c3 are colored as previously described in Figure 1. Argininosuccinate (AS) is colored in blue and domain 1 (with the exception of region c1) is colored in orange. The amino acid side chains, with the exception of R115, interact either directly or indirectly via H-bonds with the substrate. R115 is within van der Waals distance and has been shown to be critical for substrate binding (*38*). The figure was prepared with the program TURBO−FRODO.

This hypothesis is confirmed, in part, by the superposition of the H91N and H162N mutant structures, which shows that the reoriented proline in active sites A and D in the H162N structure would sterically clash with residue Ser 280 in the H91N structure.

These results show the intrinsic flexibility of the protein in this region. However, it is difficult at this stage to know if these conformational changes are the direct result of the H162N mutation or the characterization of two specific conformations among a family of conformations available for this loop.

Argininosuccinate Binding. Argininosuccinate was found to be bound to the enzyme in active site B (see Figure 3). Figure 4 shows that argininosuccinate, a conformationally flexible molecule, lies across the catalytic site such that the fumarate moiety is oriented toward Asn 291_A, Asn 162_B, and residues located in the conserved regions c2 and c3, while the $-NH_3^+$ and $-COO^-$ groups of the arginine moiety are oriented toward residues Ser 29_D, His 91_D, Tyr 323_D, Gln 328_D, and Lys 331_D (Table 2). These residues belong to either domain 1 or domain 2 of monomer D. The thermal factors obtained for the modeled substrate (average 58 Å^2) are compatible with the substrate being present at full occupancy in the structure. A closer examination of the values of the thermal factors along the substrate reveals that the arginine moiety is more stable than the fumarate moiety. Indeed, the second carboxylate of the fumarate moiety group cannot be modeled in the electron density map (see Figure

Table 2: H-Bond Interactions between Argininosuccinate and H162N δ II Crystallin

argininosuccinate	H162N		distance (Å)
AS O51	Gln 328 D Nε2		3.7
AS O51	Lys 331 D Nζ		2.8
AS O52	Tyr 323 D OH		2.7
AS O52	H_2O 147 W		2.8
	H_2O 147 W	Leu 327 D O	2.9
	H_2O 147 W	Arg 238 D NH2	2.5
	H_2O 147 W	Asp 330 D Oδ2	3.1
	H_2O 147 W	Lys 331 D Nζ	3.6
AS N4	Gln 328 D Nε2		3.0
AS N4	Ser 29 D Oγ		3.8
AS N4	H_2O 69 W		3.2
	H_2O 69 W	His 91 D Nε2	3.3
	H_2O 69 W	Ser 29 D Oγ	3.4
	H_2O 69 W	Asp 33 D Oδ1	2.9
AS N1	Asn 116 D Oδ1		3.4
AS N1	Asn 116 D Nδ2		3.9
AS Oδ1	Asn 291 A Oδ1		2.6
AS Oδ2	Asn 291 A Oδ1		2.8
AS Oγ1	Thr 161 B Oγ1		3.4

3). The failure to locate these residues is almost certainly due to the conformational flexibility of the adjacent Asn 291_A/Asp 293_A segment.

Although we would have expected the substrate to be bound to all four active sites, the fact that only one active site was found to contain the substrate could be explained, in part, if the various conformations of the 280's loop found in the model affect the ability of the enzyme to bind the substrate. Table 2 shows that the Asn 291_A side chain is involved in stabilizing the substrate. This interaction is not possible if the conformational change seen at Pro 292_A in active sites A and D occurs. While the conformation of the 280's loop in active site C is comparable to that found in active site B in which the substrate was found, a closer examination of this active site (i.e., C) reveals that Arg 115 is in a different conformation. Instead of interacting via hydrogen bonds with residues Glu 95, His 110, and Gln 118 and via van der Waals interactions with His 91, Arg 115 now interacts with Asn 116. The residue lies across active site C, preventing access to the substrate binding site and prohibiting both Asn 116 (see Table 2) and itself from interacting with the substrate. The Cβ of Arg 115 makes van der Waals contacts with the aliphatic part of the substrate's arginine moiety. The importance of both of these residues in substrate binding has been demonstrated by site-directed mutagenesis (*38*).

The partially undefined fumarate moiety makes discussion about the catalytic mechanism difficult. On the basis of the results of previous FUMC−inhibitor complex structures, the position of the fumarate moiety in the catalytic site is unexpected and could suggest that the enzyme−substrate complex seen here does not reflect the true binding of the fumarate moiety of the substrate to the wild-type enzyme. Due to the orientation of the fumarate moiety and the absence of His 162, no water molecule is found in the active site. We can therefore not comment on the hypothesis suggested for FUMC that a water molecule acts as the catalytic base (*21, 35, 39*). The water molecule is believed to be activated by the charge-relay pair Glu 331−His 188 which are equivalent to Glu 296 and His 162 in this structure. If the current hypothesis regarding the catalytic mechanism of ASL/δ II crystallin is correct (*9, 40*), a reorientation of the fumarate moiety of the substrate would be required to enable the Cβ hydrogen to be within hydrogen abstraction distance of His 162. This reorientation is not possible in the current structure, due to the position of the residue Asn 291_A. However, this part of the model (the 280's loop) has been shown to be very flexible and could thus probably accommodate a different orientation of this part of the substrate.

Despite the poor definition of the fumarate part of the substrate, it is however evident that the stabilization of the central part and the arginine moiety of the substrate involves residues which are located in each of the three conserved regions, c1, c2, and c3. The central part of the substrate, linking the arginine and fumarate moieties (atoms N1, N2, and N3), interacts directly with Ser 114_D and Asn 116_D of region c1 in domain 1. These residues are two of the most conserved residues across the superfamily in this region. It has even been speculated in FUMC that Ser 114 could be the catalytic acid (*21*). However, both the structure presented here and that of FUMC with citrate bound are more consistent with Ser 114 playing a structural role (*21*). The interaction between Ser 114 Oγ1 and AS N3 is 4.1 Å. In the current structure, the nitrogen of the scissible bond N1−C is close to Asn 116 (3.4 Å). However, since it is unlikely that an aspargine can act as the catalytic acid, it appears that this residue is involved in substrate binding and stabilizing the conformation of Tyr 323 (also involved in substrate binding) (see the accompanying paper for more detailed discussion of the role of these residues) (*38*).

Interestingly, the interaction network created by the binding of the substrate helps to stabilize domain 1 of the protein. Comparison between the *B*-factors of the first domain of the four different monomers unambiguously shows that the domain 1 involved in substrate binding (domain 1, monomer D) is more stable than the domain 1 of the other three monomers [overall *B*-factor of 47 A^2 (for monomer D) versus 57 A^2 (for monomer C), 65 A^2 (for monomer B), and 71 A^2 (for monomer A)] (Table 1). Two water molecules (69_W and 147_W) which participate in this network of interactions seem to be essential to the stabilization. These water molecules bridge residues Ser 29_D, Asp 33_D, and His 91_D and residues Leu 327_D, Asp 330_D, Lys 331_D, and Arg 338_D, with the $-COO^-$ and $-NH_3^+$ groups of the arginine moiety of the substrate, respectively. The *B*-factors (28.4 and 21.2 A^2 for 69_W and 147_W, respectively) for these water molecules are significantly lower than the average *B*-factor calculated for all solvent molecules (39.5 A^2) and are therefore in good agreement with their putative role in the stabilization of the substrate. One of the water molecules bridges His 91 Nε2 and AS N4. This histidine has been shown to be critical for substrate binding, as the His to Asn and human ASL His to Gln mutants have been shown to exhibit only 10% of the wild-type δ II crystallin activity (*1, 41*). This reduction in activity is the result of a 10-fold increase in K_m relative to that of the wild-type enzyme. Given the significant reduction of catalytic activity when His 91 is mutated (*1, 41*), it is perhaps surprising that this residue does not make a direct contact with the substrate but interacts via a water molecule. This aspect will be discussed further in the following sections.

Comparison between Domain 1 of the Four Monomers of the H162N Structure. A comparison of domain 1 of monomer D, which participates in substrate binding, and domain 1 of

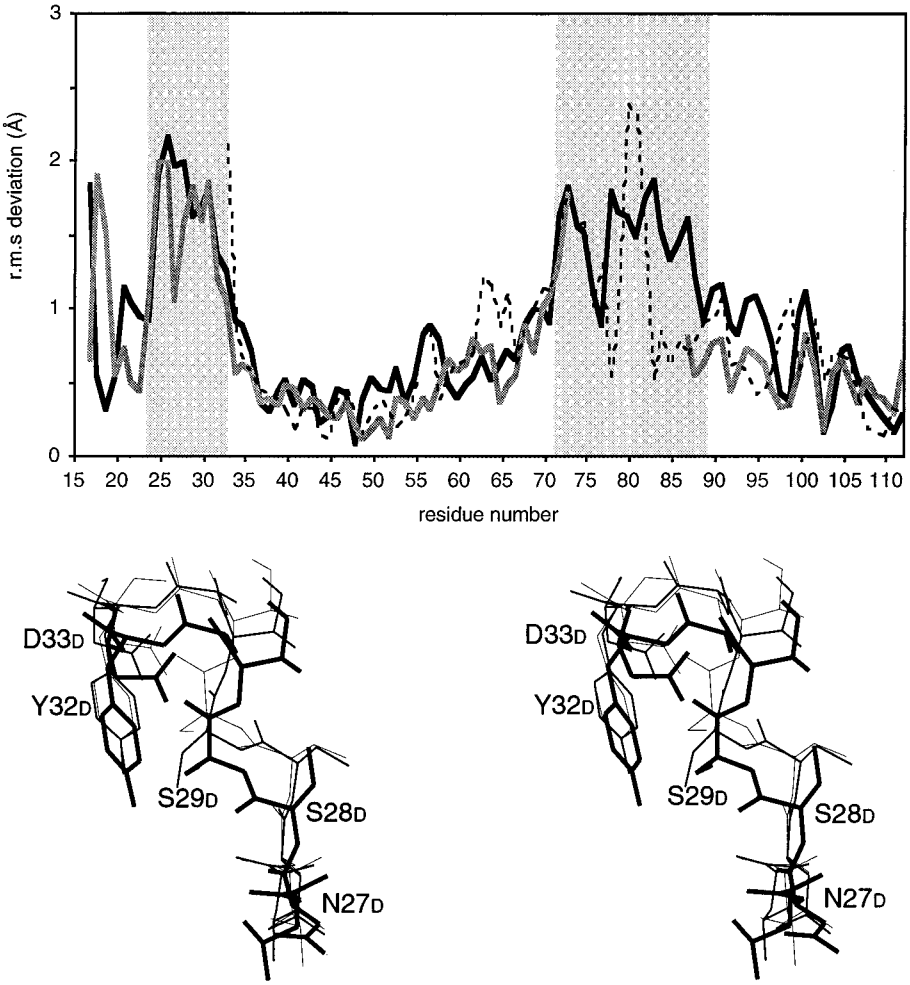


FIGURE 5: (a, top) Plot of the rms deviation between domain 1 of monomer D and monomers A, B, and C (small dashed black line, monomer A; solid gray line, monomer B; solid black line, monomer C). The squares colored in light gray indicate regions 25−34 and 76−91 where the rms deviations are found to be the most significant. (b, bottom) Stereoview of the superposition of residues 27−33 in domain 1 of monomer B (thin lines), monomer C (normal lines), and monomer D (thick lines) using the procedures described in Experimental Procedures. The homologous segment of monomer A has not been modeled in the H162N structure (see Results and Discussion) due to the poor quality of the electron density in this region. Monomer D is the only monomer involved in substrate binding.

monomers A, B, and C reveals rms deviations of 0.64, 0.64, and 0.69 Å, respectively. While most of the differences are the result of small changes in the backbone conformation, two regions differ by more than 1 Å (see Figure 5). These regions are adjacent in the structure and include residues 25−34 and 76−91. Two residues located in region 25−34 (Ser 29_D and Asp 33_D) interact directly or indirectly via a water molecule, with the $-NH_3^+$ group of the arginine moiety of the substrate. Only one residue (His 91_D) located in the second region 76−91 interacts indirectly via a water molecule with the arginine moiety of the substrate (see Table 2). While the structural variation occurring in region 25−34 is almost certainly the consequence of the substrate binding, the shift observed in the second region (data not shown) does not affect the position of the His 91 side chain. This conformation variation (76−91) is not affected by and should not affect substrate binding. Figure 5a clearly indicates that residues 35−75 are less conformationally variable.

Comparison between H91N and H162N Structures in the Vicinity of His 91. Abu-Abed et al. (*1*) proposed that His 91 was directly involved in binding. Our results show that His 91 and the substrate do interact but that the interaction is not direct but via a water molecule (Figure 4 and Table 2). This perhaps is not surprising given the low solvent accessibility of this residue (∼7 Å^2). The lack of significant conformational changes between the main chain and side chain of Asn 91 in the H91N structure and His 91 in the H162N structure shows that the position of the Asn residue in the H91N mutant is unchanged and therefore the increase in K_m observed cannot be due to a conformational change in this residue. Can we therefore explain the loss of 90% of activity in the H91N mutant?

A reorganization of the water molecule network in the vicinity of Asn 91 leading to a loss of substrate binding is improbable as there is no major side-chain conformational changes in the vicinity of His 91 in the absence of any bound substrate. The water molecule bridging His 91_D with the substrate in the H162N structure also interacts with residues Ser 29_D and Asp 33_D. A comparison between domain 1 of monomer D of H162N and the homologous monomer in the H91N structure reveals the same conformational differences that were observed between H162N monomers on substrate binding (i.e., movement of residues 25−34). There appears to be no structural reason why the conformational changes seen on substrate binding should not be possible in the H91N structure. Given the variation in the side-chain conformation for Arg 115 exhibited in the H162N structure, it may be tempting to suggest that the interaction between His 91 and

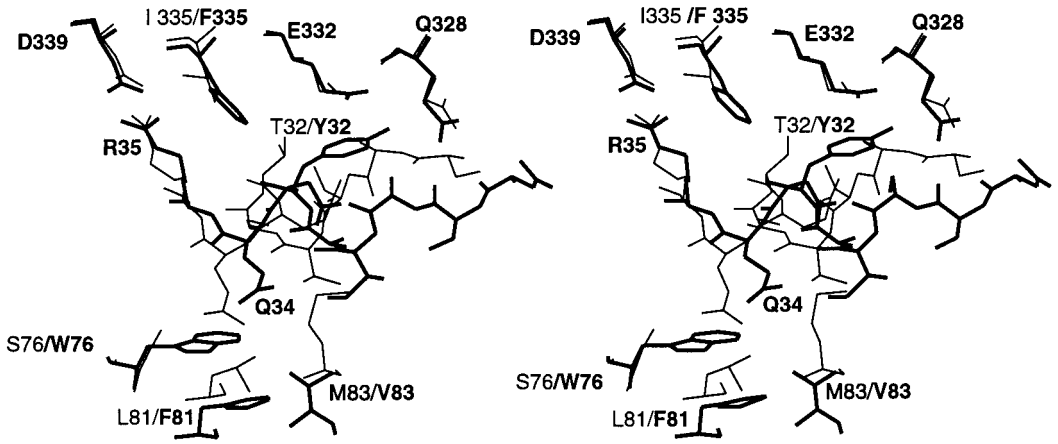


FIGURE 6: Comparison between domain 1 of TDCI (thin lines) and domain 1 of monomer D of H162N δ II crystallin (thick lines) in the vicinity of residues 25−34. This figure was prepared with the program TURBO−FRODO.

Arg 115 (van der Waals interaction between Cγ and Cβ of His 91 and Cδ and Nε of Arg 115) is not maintained in the H91N structure and that the increase in K_m was associated with an altered conformation for Arg 115, a residue that has been shown to be important in substrate binding (*38*). However, the van der Waals contact between Asn 91 Cβ and Arg 115 Cδ and Nε is maintained in the H91N structure, and there is in fact no evidence for any conformational variation of Arg 115 in the H91N structure. It therefore remains unclear exactly why the H91N mutant has lost 90% of its enzymatic activity.

Conformational Differences Seen in Domain 1 on Substrate Binding: Implications for δ I Inactivity. The replacement of His 91 with Gln in δ I crystallin could explain, in part, its loss of catalytic activity since this residue 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 complete loss of catalytic activity in δ I. The comparison of the H162N mutant structure with that of TDCI yields results similar to the comparison of the H91N mutant structure with TDCI (*1*). The two regions found previously to differ on substrate binding (residues 25−34 and 76−91) differ even more significantly when monomer D of H162N is compared to TDCI. Interestingly, with the exception of His 91, the residues in these regions which are involved (directly or indirectly via a water molecule) in substrate binding are conserved in TDCI. This indicates that the loss of affinity for the substrate in TDCI is most likely the consequence of the conformational changes seen in this region. The H91N structure had previously shown that the conformational changes in residues 76−91 were caused by the substitution of the small side chains in TDCI with larger and bulkier side chains in H91N (*1*). Interestingly, the number of substitutions between δ I and δ II crystallin found in the two structurally adjacent regions 25−34 and 76−91 represents approximatly 45% of the total number of differences found between the two proteins (see Figure 1c). In region 25−34, there are two nonconservative substitutions: Ile (δ I) to Lys 25 (δ II) (Met in duck δ I) and Thr (δ I) to Tyr 32 (δ II). The substitution (Met/Lys 25) is not in close proximity to the substrate binding site found in the H162N mutant structure, and therefore this substitution should not directly affect substrate binding. The substitution (Thr/Tyr 32) is more likely to be involved in the difference in binding affinity found between δ I and δ II. Figure 6 shows the conformational differences found in residues 25−32 of monomer D of H162N and TDCI and the contribution that other substitutions located either in region 76−91 of domain 1 or in domain 2 make to these conformational differences. The conservation of the interaction between Arg 35 and Asp 339 in both δ I and δ II crystallins means that residues 25−34 must accommodate the bulkier substitutions found in segment 76−91 [Ser (δ I) to Trp 76_D (δ II) (Ile in duck δ I) and Leu (δ I) to Phe 81_D (δ II)] and domain 2 [Ile (δ I) to Phe 333_D (δ II)]. These substitutions cause the large conformational differences that occur in this region. Furthermore, the side chains of Asn 328 and Ser 29 interact directly with the substrate in the H162N structure. In TDCI, however, the Nε2 atom of Asn 328 is hydrogen bonded to the Oγ atom of Ser 29. This hydrogen bond would prevent them from interacting with the substrate. While the substitutions do not alter significantly the position of residue 91 in either structure, in TDCI they change drastically the position of residues 25−32 and 76−91, which have been shown to participate in the stabilization of the arginine moiety of the substrate. We postulate that the conformation of residues 25−34 in TDCI may directly affect the substrate binding affinity and therefore suggest that the shift of residues 25−32 in δ I crystallin, in combination with the His → Gln mutation at position 91, is responsible for the loss of the catalytic activity.

ACKNOWLEDGMENT

We are grateful to Kiran Dole for technical assistance. The X-ray data were collected at the CCLRC, Daresbury Laboratory, U.K.

REFERENCES

1. Abu-Abed, M., Turner, M. A., Vallee, F., Simpson, A., Slingsby, C., and Howell, P. L. (1997) *Biochemistry 36*, 14012−14022.
2. Delaye, M., and Tardieu, A. (1983) *Nature 302*, 415−417.
3. Fernald, R. D., and Wright, S. E. (1983) *Nature 301*, 618−620.
4. Wistow, G. J., and Piatigorsky, J. (1988) *Annu. Rev. Biochem. 57*, 479−504.
5. Piatigorsky, J. (1993) *Dev. Dyn. 169*, 196−272.
6. Piatigorsky, J., and Wistow, G. J. (1991) *Science 252*, 1078−1079.
7. Piatigorsky, J., and Wistow, G. J. (1989) *Cell 57*, 197−199.

8. Wistow, G., and Piatigorsky, J. (1987) *Science 236*, 1554−1556.
9. Chiou, S., Lee, H., Lai, T., and Chang, G. (1991) *Biochem. Int. 25*, 705−713.
10. Hendriks, W., Mulders, J. W., Bibby, M. A., Slingsby, C., Bloemendal, H., and de Jong, W. W. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 7114−7118.
11. Kim, R., and Wistow, G. J. (1993) *FASEB J. 7*, 464−469.
12. Nickerson, J. M., Wawrousek, E. F., Hawkins, J. W., Wakil, A. S., Wistow, G. J., Thomas, G., Norman, B. L., and Piatigorsky, J. (1985) *J. Biol. Chem. 260*, 9100−9105.
13. Nickerson, J. M., Wawrousek, E. F., Borras, T., Hawkins, J. W., Norman, B. L., Filupa, D. R., Nagle, J. W., Ally, A. H., and Piatigorski, J. (1986) *J. Biol. Chem. 261*, 552−557.
14. Wistow, G. (1993) *Trends Biol. Sci. 18*, 301−306.
15. Mori, M., Matsubasa, T., Amaya, Y., and Takiguchi, M. (1990) *Prog. Clin. Biol. Res. 344*, 683−699.
16. O'Brien, W. E., McInnes, R., Kalumuck, K., and Adcock, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 7211−7215.
17. Piatigorsky, J., O'Brien, W. E., Norman, B. L., Kalumuck, K., Wistow, G. J., Borras, T., Nickerson, J. M., and Wawrousek, E. F. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 3479−3483.
18. Garrard, L. J., Bui, Q. T. N., Nygaard, R., and Raushel, F. M. (1985) *J. Biol. Chem. 260*, 5548−5553.
19. Barbosa, P., Wistow, G. J., Cialkowski, M., Piatigorsky, J., and O'Brien, W. (1991) *J. Biol. Chem. 266*, 22319−22322.
20. Patejunas, G., Barbosa, P., Lacombe, M., and O'Brien, W. E. (1995) *J. Exp. Eye Res. 61*, 151−4.
21. Weaver, T., and Banaszak, L. (1996) *Biochemistry 35*, 13955−13965.
22. Weaver, T. M., Levitt, D. G., Donnelly, M. I., Wilkens Stevens, P. P., and Banaszak, L. J. (1995) *Nat. Struct. Biol. 2*, 654−662.
23. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276* (Part A), 307−326.
24. Navaza, J. (1994) *Acta Crystallogr. A50*, 157−163.
25. Adams, P. D., Pannu, N. S., Read, R. J., and Brunger, A. T. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 5018−5023.
26. Brünger, 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) *Acta Crystallogr. D54*, 905−921.
27. Rice, L. M., and Brunger, A. T. (1994) *Proteins 19*, 277−290.
28. Read, R. J. (1986) *Acta Crystallogr. A42*, 140−149.
29. Roussel, A., and Cambillau, C. (1991) in *Silicon Graphics geometry partners directory* (Graphics, S., Ed.) pp 86, Silicon Graphics, Mountain View, CA.
30. Kleywegt, G. J. (1995) *ESF/CCP 4 Newsl. 31*, 45−50.
31. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
32. Simpson, A., Bateman, O., Driessen, H., Lindley, P., Moss, D., Mylvaganan, S., Narebor, E., and Slingsby, C. (1994) *Nat. Struct. Biol. 1*, 724−733.
33. Turner, M. A., Simpson, A., McInnes, R. R., and Howell, P. L. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 9063−9068.
34. Shi, W., Dunbar, J., Jayasekera, M. M. K., Viola, R., and Farber, G. K. (1997) *Biochemistry 36*, 9136−9144.
35. Weaver, T., Lees, M., and Banaszak, L. (1997) *Protein Sci. 6*, 834−842.
36. Saribas, A. S., Schindler, J. F., and Viola, R. E. (1994) *J. Biol. Chem. 269*, 6313−6319.
37. Walker, D. C., Christodoulou, J., Craig, H. J., Simard, L. R., Ploder, L., Howell, P. L., and McInnes, R. R. (1997) *J. Biol. Chem. 272*, 6777−6783.
38. Chakraborty, A., Davidson, A., and Howell, P. L. (1999) *Biochemistry 38*, 2435−2443.
39. Weaver, T., Lees, M., Zaitsev, V., Zaitseva, I., Duke, E., Lindley, P., McSweeny, S., Svensson, A., Keruchenko, J., Keruchenko, I., Gladilin, K., and Banaszak, L. (1998) *J. Mol. Biol. 280*, 431−442.
40. Garrard, L. J., Bui, Q. T. N., Nygaard, R., and Raushel, F. M. (1985) *J. Biol. Chem. 260*, 5548−5553.
41. Barbosa, P., Cialkowski, M., and O'Brien, W. E. (1991) *J. Biol. Chem. 266*, 5286−5290.
42. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

BI982149H