A Duck δ1 Crystallin Double Loop Mutant Provides Insight into Residues Important for Argininosuccinate Lyase Activity[†,‡]

May Tsai,[§,||] Liliana M. Sampaleanu,[§] Caroline Greene,[§] Louise Creagh,[⊥] Charles Haynes,[⊥] and P. Lynne Howell*[,§,||]

Structural Biology and Biochemistry, Research Institute, Hospital for Sick Children, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada, Department of Biochemistry, Faculty of Medicine, Medical Sciences Building, University of Toronto, Toronto, Ontario M5S 1A8, Canada, and Centre for Biological Calorimetry, Biotechnology Laboratory, University of British Columbia, 237-6174 University Boulevard, Vancouver, British Columbia V6T 1Z3, Canada

Received May 28, 2004; Revised Manuscript Received July 16, 2004

ABSTRACT: δ-Crystallin is directly related to argininosuccinate lyase (ASL), and catalyzes the reversible hydrolysis of argininosuccinate to arginine and fumarate. Two δ-crystallin isoforms exist in duck lenses, δ1 and δ2, which are 94% identical in amino acid sequence. Although the sequences of duck δ2-crystallin (dδc2) and duck δ1-crystallin (dδc1) are 69 and 71% identical to that of human ASL, respectively, only dδc2 has maintained ASL activity. Domain exchange experiments and comparisons of various δ-crystallin structures have suggested that the amino acid substitutions in the 20's (residues 22−31) and 70's (residues 74−89) loops of dδc1 are responsible for the loss of enzyme activity in this isoform. To test this hypothesis, a double loop mutant (DLM) of dδc1 was constructed in which all the residues that differ between the two isoforms in the 20's and 70's loops were mutated to those of dδc2. Contrary to expectations, kinetic analysis of the DLM found that it was enzymatically inactive. Furthermore, binding of argininosuccinate by the DLM, as well as the dδc1, could not be detected by isothermal titration calorimetry (ITC). To examine the conformation of the 20's and 70's loops in the DLM, and to understand why the DLM is unable to bind the substrate, its structure was determined to 2.5 Å resolution. Comparison of this structure with both wild-type dδc1 and dδc2 structures reveals that the conformations of the 20's and 70's loops in the DLM mutant are very similar to those of dδc2. This suggests that the five amino acid substitutions in domain 1 which lie outside of the two loop regions and which are different in the DLM, and dδc2, must be important enzymatically. The structure of the DLM in complex with sulfate was also determined to 2.2 Å resolution. This structure demonstrates that the conformational changes of the 280's loop and domain 3, previously observed in dδc1, also occur in the DLM upon sulfate binding, reinforcing the hypothesis that these events may occur in the active dδc2 protein during catalysis.

δ-Crystallin is the major soluble component of avian eye lenses and is directly related to the housekeeping enzyme, argininosuccinate lyase (ASL).[1] In all organisms, ASL is involved in arginine biosynthesis, catalyzing the reversible breakdown of argininosuccinate to arginine and fumarate. In duck lenses, gene recruitment (*1*) of ASL and subsequent gene duplication have resulted in two δ-crystallin isoforms, dδc1 and dδc2. These isoforms are 94% identical in amino acid sequence, and 69 and 71% identical to human ASL, respectively. While dδc2 has retained ASL activity and is the duck orthologue of ASL in non-lens tissues, dδc1 has evolved and is no longer enzymatically active. ASL and δ-crystallin also belong to a superfamily of enzymes which includes class II fumarase (*2*), adenylosuccinate lyase (*3*), L-aspartase (*4*), and 3-carboxy-*cis*,*cis*-muconate lactonizing enzyme (*5*). Despite a low level of overall amino acid sequence identity between the enzymes of approximately 15−30%, three regions of highly conserved amino acid residues (C1−C3) have been identified as consensus se-

[†] This work is supported by a grant from the Canadian Institutes for Health Research (CIHR) of Canada to P.L.H. M.T. and L.M.S. were supported, fully or in part, by the Ontario Student Opportunity Trust Fund, Hospital for Sick Children Foundation Student Scholarship Program. P.L.H. is the recipient of a CIHR Investigator award. Beam line X8C at the National Synchrotron Light Source was supported by Major Facilities Access grants from CIHR and from the Natural Sciences and Engineering Research Council (NSERC) of Canada.

[‡] The coordinates and structure factors have been deposited in the Protein Data Bank as entries 1U15 and 1U16 for the apo and sulfate-bound structures, respectively.

* To whom correspondence should be addressed. E-mail: howell@sickkids.ca. Telephone: (416) 813-5378. Fax: (416) 813-5379.

[§] Hospital for Sick Children.

[||] University of Toronto.

[⊥] University of British Columbia.

[1] Abbreviations: ASL, argininosuccinate lyase; dδc1, duck δ1-crystallin; dδc2, duck δ2-crystallin; CX, conserved amino acid region X, where X corresponds to 1, 2, or 3; 211-chimera, hybrid duck δ-crystallin protein, containing domain 1 from the active duck δ2-crystallin isoform and domains 2 and 3 from the inactive duck δ1-crystallin isoform; DLM, duck δ1-crystallin double loop mutant, where residues 22−31 and 74−89 have been mutated to those of duck δ2-crystallin; SLM X, duck δ1-crystallin single loop mutant, where X corresponds to either residues 22−31 or 74−89 which have been mutated to those of duck δ2-crystallin; ITC, isothermal titration calorimetry; ASU, asymmetric unit; rmsd, root-mean-square deviation. The amino acid numbering used throughout the text is that of dδc1. dδc2 has a two-residue insert at position 5, causing the amino acid numbering to be shifted by 2 relative to that of ASL and dδc1. In this paper, to avoid confusion when comparing structurally equivalent residues, we have used the dδc1 numbering for both the dδc1 and dδc2 isoforms.

10.1021/bi0489006 CCC: $27.50
Published on Web 08/26/2004

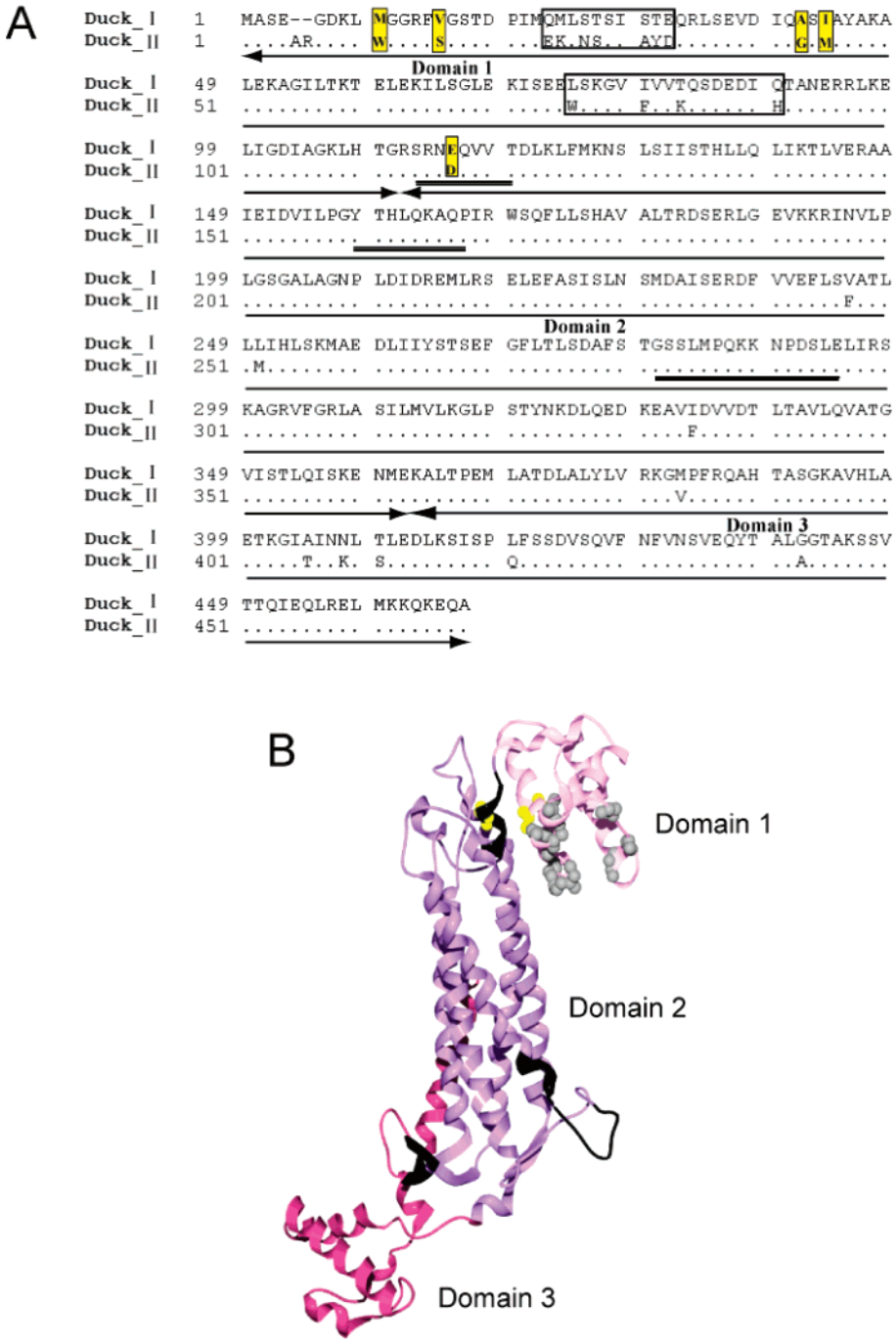

FIGURE 1: (A) Amino acid sequence alignment of dδc1 and dδc2. The end of each structural domain is represented by an arrowhead. The three regions of highly conserved amino acid residues across the superfamily (C1−C3) are underlined. C1 encompasses residues 112−119, C2 residues 157−166, and C3 residues 280−294. The 20's (22−31) and 70's (74−89) loops are boxed, and the remaining five amino acid residues which differ between the DLM and the 211-chimera are highlighted in yellow. (B) Schematic ribbon diagram of the DLM monomer. Each structural domain is indicated and colored in a different shade of pink. The regions of conserved amino acid residues (C1−C3) which constitute the active site in the tetramer are shown in black. The distribution of amino acid residues mutated in the 20's and 70's loops of the DLM, and those which still differ between the DLM and 211-chimera, are shown in gray and yellow CPK representation, respectively. M9 and V14 are not shown as the electron density for residues 1−16 was of insufficient quality to allow these residues to be modeled.

quences across the superfamily members. Amino acid residues in these conserved regions have been shown to be involved in the general acid−base catalytic mechanism of these enzymes (*6*−*10*).

Seventeen of the 27 amino acid differences between the active and inactive δ-crystallin isoforms reside in the N-terminal domain (domain 1) (Figure 1A). Domain exchange experiments have shown that ASL activity can be regenerated in the enzymatically inactive dδc1 by replacing its first structural domain with that of dδc2 (211-chimera) (*11*) (Figure 2). Comparisons of wild-type dδc1 and dδc2, dδc2 mutants, and turkey δ1-crystallin structures have also revealed both inter- and intraspecies conformational differences in two loop regions of domain 1, which encompass residues 22−31 (20's loop) and 74−89 (70's loop) (*12*). Given that the putative catalytic residues in ASL and dδc2 are conserved between the active and inactive δ-crystallin isoforms, the amino acid substitutions in the 20's and 70's

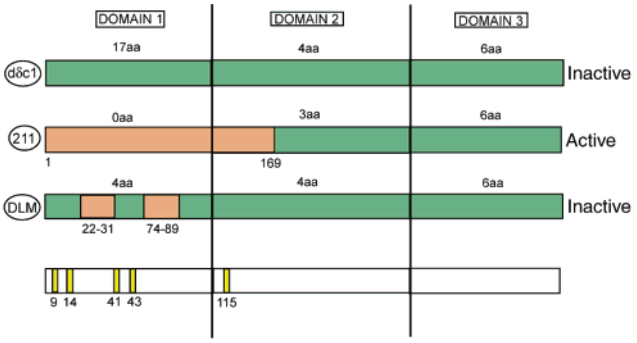

FIGURE 2: Schematic representation of dδc1, the 211-chimera, and the DLM. Residues corresponding to dδc2 are shown in orange, while those corresponding to dδc1 are shown in green. The number of residue differences between each domain of dδc1, the 211-chimera, or the DLM and dδc2 is indicated above each schematic representation. Because of the restriction enzyme used to create the 211-chimera, its domain 1 encompasses residues 1−169 (*11*). There is therefore an additional conservative amino acid substitution (D → E) between the DLM and 211-chimera at residue 115. Also shown (bottom) are the locations of the four other amino acid differences between the 211-chimera and DLM proteins.

loops of dδc1 and the resulting conformational variations have been hypothesized to prevent substrate binding, thus rendering dδc1 enzymatically inactive (*7*, *12*, *13*). During the determination of the structure of wild-type dδc1, a sulfate ion was located in the active site region of the protein. The presence of the ion resulted in a large conformational change of the 280's loop (residues 270−290) as well as a concerted rigid body movement of domain 3 (*12*). Superposition of dδc1 with the H160N− and S281A−dδc2 enzyme substrate complexes revealed that the sulfate ion was bound in a position similar to the fumarate moiety of argininosuccinate (*6*, *7*). As a result, the observed conformational movements of the 280's loop and domain 3 in dδc1 have been hypothesized to occur in dδc2 on substrate binding (*12*).

To investigate the hypothesis that the 20's and 70's loops are responsible for the loss of catalytic activity in dδc1, a series of mutants of dδc1 were generated in which all residues in the 20's and 70's loops, which differed between the active and inactive isoforms, were systematically replaced with those of dδc2. We found that all of these mutants, including the double loop mutant (DLM), were enzymatically inactive. As a result, the ability of dδc1 and the DLM to bind the substrate was investigated by isothermal titration calorimetry (ITC). These experiments confirmed that neither protein bound argininosuccinate at levels that could be detected by ITC. To examine the conformation of the 20's and 70's loops in the DLM and to understand why this mutant is unable to bind the substrate, despite the mutations introduced into the loop regions, its crystal structure was determined to 2.5 Å resolution. The structural comparison of the DLM with dδc1 and dδc2 found that the conformations of the 20's and 70's loops are similar to those observed in the active dδc2 enzyme. This and the kinetic results suggest that the remaining five residues in domain 1, residues 9, 14, 41, 43, and 115, which differ between the active 211-chimera and DLM (Figure 2) must be important for enzymatic activity. Since the amino acid sequence of the DLM is closer to dδc2 than dδc1, and to investigate whether this protein, like dδc1, was capable of large conformational changes, the structure of a DLM−sulfate complex was also determined to 2.2 Å resolution. Conformational changes of the 280's loop and domain 3 occur in the DLM upon sulfate binding, providing additional evidence that these movements may be relevant to catalysis in the active dδc2 isoform.

MATERIALS AND METHODS

Construction of dδc1 Mutants. dδc1 mutants targeting residues of the 20's and/or 70's loops (Table 1) were generated using the Unique Site Elimination Mutagenesis procedure (Pharmacia). Initially, the dδc1 gene was subcloned from pET-17b into the pUC19 vector at the *Xba* and *Hin*dIII sites (Pharmacia). The pUC19 vector containing the dδc1 gene served as a starting template for generation of all subsequent mutants. Initially, individual point and double mutants of dδc1 were constructed in which various residues of the 20's and 70's loops were mutated to those of dδc2. Subsequently, two single loop mutants (SLMs), SLM 22−31 and SLM 74−89, were constructed using the double mutants as starting constructs. In each case, the remaining residue differences between dδc1 and dδc2 in the 20's or 70's loop were mutated to those of dδc2 in two sequential steps. Using SLM 78−89 as the starting template, the DLM was generated by mutating the residues in the 20's loop to those of dδc2 in the same stepwise manner used to generate SLM 22−31 (Table 1). All mutants were subcloned back into the pET-17b vector, in frame with a 3′ six-histidine tag. In addition to the mutations at the desired codons, all oligonucleotide primers were designed with silent mutations that introduced a restriction site or inactivated an existing one to serve as a quick screening tool for identifying positive clones. DNA sequencing (The Center for Applied Genomics at The Hospital for Sick Children) confirmed the presence of all desired mutations.

Protein Expression, Purification, and Kinetic Analysis. The C-terminal His-tagged proteins were expressed and purified as described previously (*11*). The eluted protein (15 mL) was collected in one fraction and dialyzed overnight at 4 °C against 4 L of buffer containing 10 mM Tris-HCl (pH 7.5) and 1 mM EDTA (buffer A). All mutants were assayed for ASL activity by monitoring at 25 °C the production of fumarate at 240 nm ($\epsilon = 2.44$ mM^{-1} cm^{-1}) (*11*).

Isothermal Ttitration Calorimetry. Titrations of dδc1 and the DLM were carried out on a MCS-ITC instrument (MicroCal LLC, Northampton, MA). All titrations were performed in buffer A. Protein samples were centrifuged (4 min at 14000*g*) to remove any particulates prior to determination of the spectrophotometric concentration (ϵ_{280} = 13 370 M^{-1} cm^{-1} for dδc1 and 20 340 M^{-1} cm^{-1} for DLM). Buffer solutions were filtered (0.2 μm) before use to remove any particulates. For each titration, 26 consecutive 10 μL aliquots of 7 mM argininosuccinate in buffer A were injected into the sample cell (volume = 1.35 mL) containing either 92 μM dδc1 or 119 μM DLM protein. The time between each injection varied from 5 to 20 min depending on the kinetics of the binding reactions. Control experiments were performed by injecting 10 μL aliquots of argininosuccinate into protein-free buffer. Each ITC experiment was performed two or three times.

Crystallization, Data Collection, and Structure Determination of DLM. Crystals for both DLM structures were grown at room temperature using the hanging drop vapor diffusion

Table 1: Mutants Targeting Residues in the 20's and 70's Loops

Point mutants	Double mutants
L74W	Q22E/M23K
T82K	L74W/I79F
Q89H	

Multisite Mutants[a]

mutant	starting construct	step 1	step 2	step 3
SLM 23−32	(Q22E/M23K)	(Q22E/M23K/**S25N/T26S)**	[Q22E/M23K/S25N/T26S/ **S29A/T30Y/E31D**]	
SLM 74−89	(L74W/I79F)	(L74W/I79F/**Q89H)**	[L74W/I79F/**T82K**/Q89H]	
DLM	(L74W/I79F/T82K/Q89H)	**(Q22E/M23K**/L74W/ I79F/T82K/Q89H)	(Q22E/M23K/**S25N/T26S/** L74W/I79F/ T82K/Q89H)	[Q22E/M23K/S25N/ T26S/**S29A/T30Y/ E31D**/L74W/I79F/ T82K/Q89H]

[a] The new mutations introduced into previously generated double or multiple mutants are shown in bold. Starting or working constructs are shown in parentheses, while the final constructs are shown in square brackets.

Table 2: Crystallization, Data Processing, and Refinement

	DLM−sulfate	apo-DLM
Crystallization		
protein concentration (mg/mL)	8	8
crystallization conditions	100 mM MES (pH 6.5), 1.2 M $(NH_4)_2SO_4$, 10 mM $CoCl_2$	100 mM HEPES (pH 7.5), 8% PEG 2K MME, 300 mM $MgCl_2$
space group	$P42_12$	$P2_1$
unit cell dimensions	$a = b = 133.3$ Å, $c = 73.9$ Å, $\alpha = \beta = \gamma = 90°$	$a = 93.7$ Å, $b = 98.9$ Å, $c = 107.4$ Å, $\beta = 101.5°$
Data Processing Statistics		
no. of monomers/ASU	1	4
resolution limits (Å)	50−2.2	35−2.5
total no. of data	414249	249879
no. of unique data	34371	66502
mean redundancy	11.5	3.8
completeness (%)	99.9 (99.4)[b]	98.3 (97.2)[b]
average $I/\sigma(I)$	6.3	15.3
% reflections for which $I > 2\sigma(I)$	77.9 (56.5)[b]	84.5 (61.2)[b]
R_{sym}[a]	0.087 (0.320)[b]	0.078 (0.293)[b]
Refinement Statistics		
resolution range (Å)	50−2.2	35−2.5
R_{cryst} (%)[c]	21.2	22.0
R_{free} (%)[d]	26.5	28.0
no. of reflections in refinement	30867	58752
no. of reflections in test set	3420	6627
no. of non-hydrogen atoms (protein/sulfate/solvent)	3479/5/201	13817/−/625
mean *B*-factor (Å^2)		
protein	38	41
per monomer (A/B/C/D)	38	39/44/38/41
per domain (A/B/C/D)		
domain 1	45	45/50/51/59
domain 2	34	33/37/32/34
domain 3	42	48/57/41/42
sulfate	30	−
solvent	41	39
rms deviation from ideal		
bond lengths (Å)/angles (deg)	0.008/1.4	0.008/1.3
dihedral angles (deg)/improper angles (deg)	19.9/0.86	19.4/0.90

[a] $R_{sym} = \sum|I - \langle I\rangle|/\sum I$, where I is the measured intensity for symmetry-related reflections and $\langle I\rangle$ is the mean intensity for the reflection. [b] Last-resolution shell extends from 2.28 to 2.2 Å and from 2.59 to 2.5 Å, respectively. [c] $R_{crys} = \sum(|F_o| - |F_c|)/\sum(|F_o|)$. [d] $R_{free} = \sum(|F_{os}| - |F_{cs}|)/\sum(|F_{os}|)$, where s refers to a subset of data not used in the refinement, representing 10% of the total number of observations.

method. Two microliters of protein (8 mg/mL in buffer A) was mixed with an equal volume of reservoir buffer and suspended over 1 mL of reservoir buffer (Table 2). Three-dimensional crystals (∼0.25 mm × 0.25 mm × 0.20 mm) of the DLM−sulfate complex were obtained after two weeks, while two-dimensional plates (∼0.6 mm × 0.4 mm × 0.05 mm) of the apoprotein were obtained after one month. Data for the DLM−sulfate complex were collected to 2.2 Å resolution at Station X8C of the National Synchrotron Light Source (NSLS, Brookhaven, NY) ($\lambda = 0.96$ Å). Data for the apo structure were collected on our in-house data collection facilities at The Hospital for Sick Children to 2.5 Å resolution. In each case prior to flash-freezing, the crystals were immersed sequentially in solutions of mother liquor

containing 5, 10, 15, and 20% (v/v) glycerol (~5 s each). Both sets of data were processed using the d*TREK software package (*14*) (Table 2). Prior to structure determination and refinement, the DREAR program package (*15*) was used to apply Bayesian statistics to the reduced data to improve the weak data as well as to eliminate or correct negative intensities.

The structures of the apo- and sulfate-bound DLM were determined by the molecular replacement technique using a tetramer and monomer of the wild-type dδc2 protein (PDB entry 1HY1) as the starting model, respectively. Since the conformations of the 280's loop and domain 3 in the DLM−sulfate complex were expected to resemble those of dδc1, the dδc2 protein was used as a search model rather than dδc1. This ensured that bias in the conformation of these regions in the DLM−sulfate complex was kept to a minimum. As the apo-DLM structure was determined in the same space group as wild-type dδc2 ($P2_1$), difference Fouriers could have likely been used for structure determination. Molecular replacement was used for additional verification given the small variations in the unit cell dimensions and β angle. The cross rotation and translation functions were performed with data between 15 and 5 Å resolution. Correlation coefficients of 0.58 and 0.76 and R_{crys} values of 32 and 41% were obtained for the apo and sulfate-bound structures, respectively.

Structure Refinement and Model Building. Refinement of the structures was performed with CNS (*16*) with a maximum likelihood target function (*17*, *18*), a flat bulk solvent correction (*19*), and no low resolution or σ cutoff applied to the data. Ten percent of the reflections were randomly selected from the data and used to compute an R_{free} for cross validation of the model. Initially, rigid body refinement was used to refine the orientation and position of each subunit. All subsequent rounds of refinement consisted of simulated annealing and grouped or individual *B*-factor refinement followed by calculation of σ_A-weighted $2F_o - F_c$ and $F_o - F_c$ electron density maps. After the first round of simulated annealing and *B*-factor refinement, residues in the model were changed to reflect the correct sequence of the protein. Each round of refinement was alternated with manual rebuilding of the model using Xfit (*20*). Simulated annealing omit maps were used to remove bias and to verify the more variable regions of the model, especially the 20's, 70's, and 280's loops. During refinement of the apo-DLM structure, noncrystallographic symmetry restraints were applied and gradually released while monitoring the decrease in R_{free}.

Peaks with proper hydrogen bonding coordination and with electron densities greater than 3.0σ on the σ_A-weighted $F_o - F_c$ map were picked automatically as water molecules. Additional water molecules with proper coordination and electron densities greater than 1σ and 2.5σ on the σ_A-weighted $2F_o - F_c$ and $F_o - F_c$ maps, respectively, were introduced manually. Strong tetrahedral density was observed in the active site region of the DLM−sulfate complex. A *cis*-peptide was modeled between S319 and T320 in all monomers. The quality of the models was assessed with PROCHECK (*21*), and the stereochemistry of the models was found to be very good. Greater than 90% of all residues in both models reside in the most favored region of the Ramachandran plot. No residues lie in the disallowed regions. Residue L204 is found in a type IV β-turn and is found in the generously allowed region. A summary of the refinement statistics is presented in Table 2.

Structural and Sequence Alignments. Comparisons of the various δ-crystallin structures were performed with PROFIT (version 6.0) written by G. D. Smith. Structurally equivalent residues in the five core helices of domain 2 were chosen and subjected to an iterative least-squares fitting procedure. The following residues were chosen: E115−I151, Q171−V196, V239−T265, L293−G316, and E330−Q354. For the purpose of all comparisons between δ-crystallin structures, monomer A of each protein was selected for the analysis. Comparison of different monomers belonging to the same protein (dδc2, dδc1, or apo-DLM) revealed overall rms deviations of 0.31−0.45 Å for 444−447 C_α atoms, suggesting no significant intrinsic conformational differences exist between monomers of the same protein. The Least Squares option in Xfit (*20*), as well as SWISS PDB Viewer (*22*), was employed for additional structural analyses.

RESULTS AND DISCUSSION

Role of Residues 22−31 (20's Loop) and 74−89 (70's Loop) in Substrate Binding. The 211-chimera, consisting of domain 1 from dδc2 and domains 2 and 3 from dδc1, exhibits ASL activity (*11*) (Figure 2). Comparisons of the active and inactive δ-crystallin structures also reveal that the largest conformational variability between the two isoforms is in the 20's and 70's loops (*12*). These structural differences are the consequence of seven and four intraspecies amino acid substitutions between dδc1 and dδc2 in the 20's and 70's loops, respectively (Figure 1A). Although the 20's and 70's loops are flexible, evidence suggests that a specific conformation of these loops is required for substrate binding and catalysis. The structures of inactive mutants of dδc2 with the substrate bound have identified a number of residues in these loops which interact either directly (S27) or via a water molecule (D31, D87, and H89) with the amino end of the arginine moiety of the substrate (*6*, *13*). Furthermore, hydrogen bonds are observed between N25 (Oδ1) and K323 (Nζ) as well as between E86 (Oϵ2) and R113 (Nη1). K323 and R113 are both involved in stabilizing the guanidine moiety of argininosuccinate. In addition, the main chain nitrogen atom of S27 establishes a hydrogen bond with the carbonyl oxygen of E86. Maintenance of interactions between residues in the 20's and 70's loops, both of which are conformationally flexible, has been suggested to be important for proper substrate binding and catalysis (*7*, *12*). Given the critical role of several of the residues in the 20's and 70's loops in substrate binding, any mutations and resulting conformational changes in these loops are likely to negatively impact substrate binding and catalysis. Indeed, mutagenesis of D31, D87, or H89 results in a catalytically impaired or inactive protein (*6*, *13*), and the D87G mutation in ASL has been implicated in the genetic disorder *argininosuccinic aciduria* (*23*). Since the putative catalytic residues are conserved between the dδc1 and dδc2 isoforms, the absence of enzymatic activity in dδc1 has been attributed to the amino acid substitutions in the 20's and 70's loops and the inability of the protein to bind the substrate (*7*, *12*, *13*).

Enzymatic Activity of the dδc1 Mutants. Our kinetic analysis reveals that all the dδc1 mutants targeting residues in the 20's and 70's loops (Table 1) were enzymatically

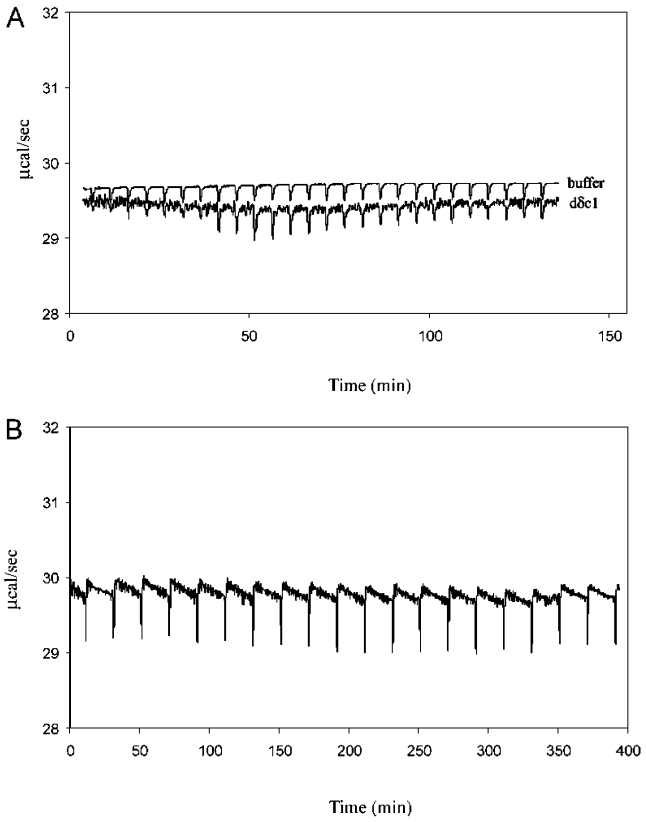

FIGURE 3: Raw ITC data for 10 μL injections of 7 mM argininosuccinate into (A) 92 μM dδc1 and buffer (heat of dilution) and (B) 119 μM DLM. Injections of substrate into the protein solution were performed at 5 and 20 min intervals for dδc1 and the DLM, respectively.

inactive. This was perhaps not unexpected for the point, double, or single loop mutants, as residues in both the 20's and 70's loops of dδc2 are required for correct anchoring of the arginine moiety of the substrate. Since the restoration of activity in the dδc1 isoform presumably requires that the protein adopt the dδc2 conformation in both the 20's and 70's loops, the absence of ASL activity in the DLM was therefore surprising and unexpected. Given the domain exchange experiments (*11*) and conformational differences observed in the 20's and 70's loops between dδc1 and dδc2 (*12*), a dδc1 mutant with residues in the 20's and 70's loops identical to those of dδc2 was expected to exhibit ASL activity.

Dδc1 and the DLM Do Not Bind Argininosuccinate. To determine whether dδc1 and the DLM are capable of binding argininosuccinate, ITC experiments were performed. Each 10 μL injection of the argininosuccinate solution into dδc1 resulted in very small exothermic heat signals (Figure 3A). These signals are very similar in magnitude to those obtained for the heat of dilution of the substrate (injection of the substrate solution into buffer alone), with only slight differences in observed integrated heats. The very small heat changes that were detected do not allow any precise quantitative conclusions to be drawn from the data. At the substrate and protein concentrations used in the experiment, ITC is able to detect binding constants of $\geq 5 \times 10^3$ M^{-1}. Our results therefore suggest that if dδc1 binds argininosuccinate, it does so with a binding constant of $< 5 \times 10^3$ M^{-1}.

For the DLM, each injection of substrate solution resulted in a fast exothermic signal followed by a very slow endothermic heat signal (Figure 3B). Such a heat pattern may be associated with two kinetic phases, such as a rapid contact phase followed by a slower conformational change, upon interaction of the protein with its substrate. However, the small invariant heats observed in this titration are strongly indicative of a weak or nonbinding system. We therefore conclude that the DLM does not bind argininosuccinate, or binds with a K_a below the detection limit of ITC at the substrate and protein concentrations that were used ($< 5 \times 10^3$ M^{-1}). The slow endothermic heat signal, as well as the increased noise in the baseline of the titration profile for the DLM system, can instead be attributed to the formation of precipitate during the experiments, which also contributes to the larger heat signals observed for the DLM system relative to those observed for dδc1 and the heat of dilution of the substrate. Any event leading to a gain or loss of heat from the system will be reflected as a change in temperature of the system, and will consequently be detected by the instrument.

ITC was therefore unable to detect binding of argininosuccinate by either dδc1 or the DLM. These results provide the first experimental evidence that dδc1's loss of enzymatic activity is the consequence of its inability to bind the substrate and strongly suggest that the DLM mutant is inactive for the same reason. Given the kinetic and ITC results, and to better understand why the DLM mutant is catalytically inactive, its crystal structure was determined.

Structure Determination and Overall Fold of DLM Structures. The apo and sulfate-bound structures of the DLM were determined and refined to R_{crys} and R_{free} values of 21.2 and 22.0%, and 26.5 and 28.0%, respectively. The overall fold of the DLM structures is highly α-helical in nature and is similar to those of previously determined δ-crystallin structures (*6*, *7*, *12*, *13*, *24*). The proteins are homotetrameric, with each monomer comprised of three structural domains (Figure 1B). The structures of the apo and sulfate-bound DLM proteins were determined with a tetramer and monomer in the asymmetric unit (ASU), respectively. Application of appropriate symmetry elements to the contents of the DLM−sulfate ASU generates the tetramer. Electron density for the first 15−17 residues of each monomer is not visible in either the apo or sulfate-bound DLM structure, indicating that the N-terminal arm is flexible. Difficulty in modeling residues at the N-terminus of the protein has been reported previously (*6*, *7*, *12*, *13*, *24*). In addition, residues 464−466 at the C-terminus of each monomer have been omitted from the model as the electron density is also of poor quality in this region.

Conformation of 20's and 70's Loops. The structure of the DLM was determined to examine the conformation of the 20's and 70's loops in this protein and to determine whether any conformational variations exist in these regions, which may account for the absence of activity in the DLM. A structural comparison of the DLM and dδc2 reveals C_α rms deviations of ~1.1 and 1.0 Å in the 20's and 70's loops, respectively (Figures 4 and 5). Differences of ~2.6 and ~2.0 Å are seen for the 20's and 70's loops, respectively, when comparing the DLM and dδc1 structures (Figures 4 and 5). These results indicate that the two loops in the DLM adopt a conformation more closely resembling those of dδc2 than

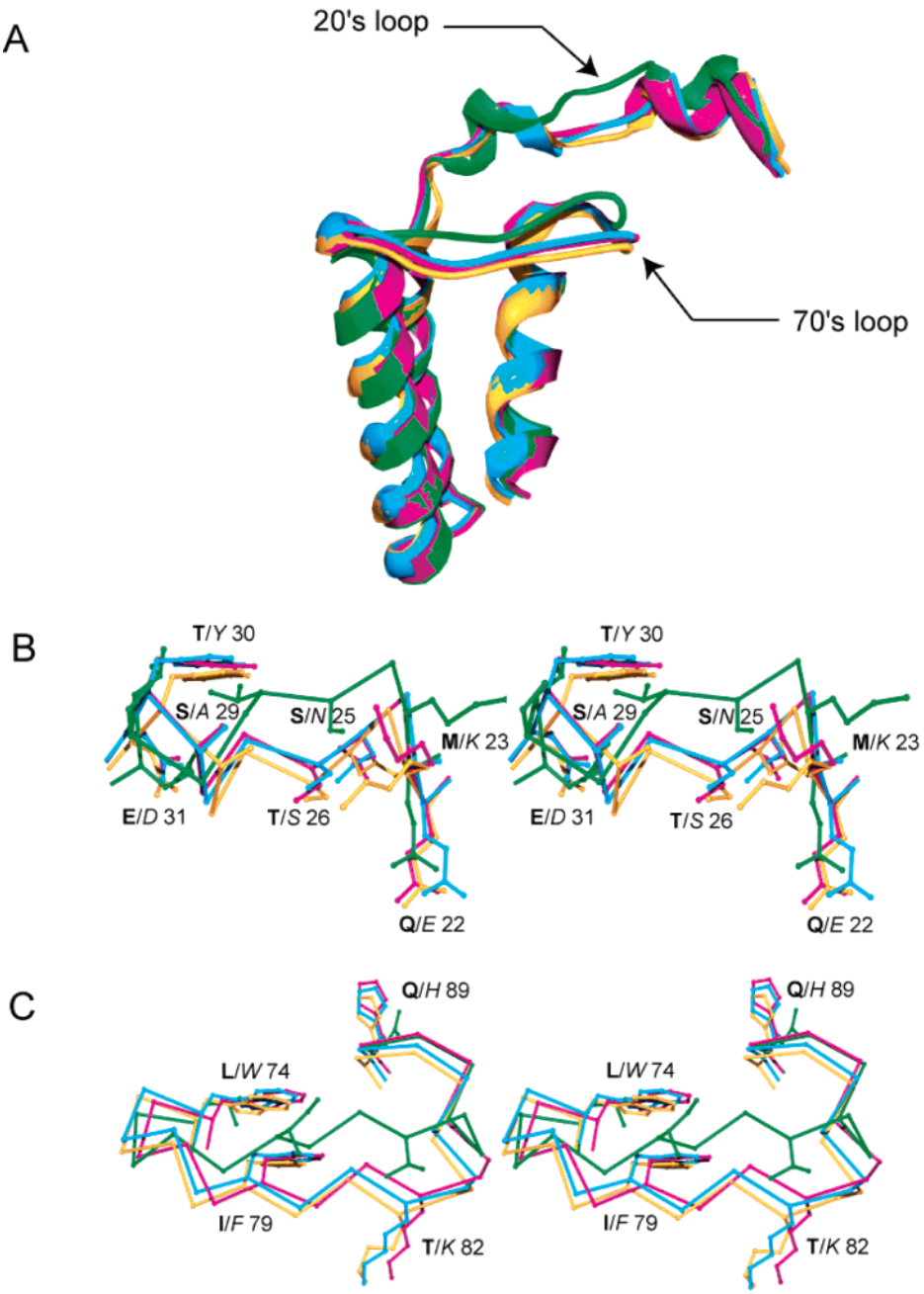

FIGURE 4: (A) Conformational differences observed in domain 1 (residues 18−100) of dδc1, dδc2, the apo-DLM, and sulfate-bound DLM structures. Close-up stereoview of the (B) 20's (residues 22−31) loop and (C) 70's (residues 74−89) loop. Residues targeted for mutagenesis in the DLM are shown in stick representation. For each amino acid, the one-letter code is shown in bold for dδc1 and in italics for dδc2/DLM protein. In each panel, the DLM−sulfate complex, apo-DLM protein, dδc1, and dδc2 are shown in pink, blue, green, and orange, respectively.

those of the inactive dδc1 protein. This was expected as the residues in these regions of the DLM are now identical to those found in dδc2. The variations that exist between the DLM and dδc2 in the 20's and 70's loops (Figure 4) most likely reflect the inherent flexibility of these loop regions. When monomers within wild-type dδc2 are compared, for example, C_α rms deviations of 0.33−0.93 and 0.35−1.59 Å are observed for the 20's and 70's loops, respectively. These types of variations are seen not only when comparing monomers within the same protein but also between monomers of different dδc2 mutants. These regions also have generally weaker electron density and result in higher than average *B*-factors. For example, average values of 52−61 and 46−66 Å^2 are observed for the 20's and 70's loops of the apo-DLM protein, respectively, compared to a mean *B*-factor of 41 Å^2 for the entire protein.

Although the conformations of the 20's and 70's loops in the DLM more closely resemble those of wild-type dδc2 rather than those of dδc1, the kinetic and ITC data indicate that the DLM mutant is enzymatically inactive and unable to bind substrate. Together, these findings suggest that although the identities of the residues in the 20's and 70's loops are important for substrate binding and catalysis, one or more of the remaining amino acid differences between the 211-chimera and DLM proteins (Figures 1 and 2) must play a crucial role in restoring enzymatic activity in the DLM.

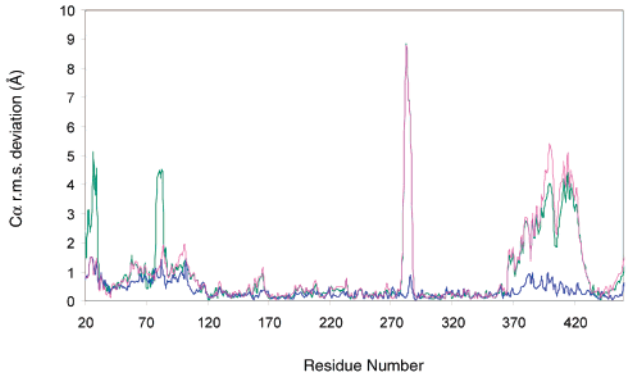

FIGURE 5: Plot of C_α rms deviations between dδc2 and monomers A of dδc1 (green), the DLM−sulfate complex (pink), and apo-DLM (blue). Residues 18−463 of each protein were used for the structural alignments as described in Materials and Methods.

Sulfate Ion Interactions in DLM. Given that the DLM more closely resembles dδc2 than dδc1 in amino acid sequence, its structure was determined in complex with sulfate to determine whether the 280's loop and domain 3 in this mutant undergo conformational changes similar to those observed previously in dδc1 (*12*). Strong tetrahedral density in the active site of the DLM allowed a sulfate ion to be modeled (Figure 6A). Each of the four active sites in the protein is found at the interface of three different monomers (e.g., monomers A, B, and D), with each monomer contributing residues from a different conserved region to the active site. The structure determination of various ASL/fumarase C superfamily members in complex with a substrate or inhibitor has confirmed the location of the active site (*2*, *6*, *7*, *9*, *25*).

The interactions observed between the protein and sulfate ion in the DLM−sulfate complex are identical to those observed in the dδc1 structure (*12*) (Figure 6B). The anion makes direct interactions via its oxygen atoms with T159 (Oγ1) and H160 (Nε2), residues in conserved region C2 of the protein, as well as with S281 (Oγ), K287 (Nζ), and N289 (Nδ2), from conserved region C3 of the protein. N114 from C1 interacts both directly (Nδ2) and indirectly (Oδ1) via Wat27 with the sulfate ion. S112, a residue also found in C1, interacts only indirectly (Oγ) via Wat218 with the O_4 atom of the anion. All of the residues observed to interact with the sulfate ion have been proposed to be involved in either substrate binding or catalysis (*6*, *7*, *10*, *13*, *26*). The spatial conservation of residues and water molecules involved in binding the sulfate ion in dδc1 and the DLM demonstrates the highly specific manner in which the anion is bound. The sulfate ions in both dδc1 and the DLM superimpose onto the fumarate moiety of argininosuccinate in the H160N (*7*) and S281A dδc2 (*6*) structures, suggesting that the sulfate ion mimics in some way this portion of the substrate.

Sulfate-Induced Conformational Changes. A structural comparison of the DLM−sulfate complex and substrate free dδc2 protein reveals a C_α rms deviation of up to ∼8 Å in the 280's loop. The conformation of this loop is very similar to that observed for the dδc1−sulfate complex, as indicated by the C_α rms deviation of only 0.15 Å when this region in both proteins is compared (Figures 5 and 7A). The conformational movement of the 280's loop that occurs on sulfate binding in the DLM and dδc1 results in closure of the loop over the active site, an event that has been suggested to sequester the substrate from the solvent during catalysis (*12*). The interactions between the sulfate anion and the residues of the 280's loop assist in rigidifying the loop. This is supported by the lower *B*-factors for the residues of this loop in the complexed DLM structure (∼34 Å^2), when compared to the value for residues in the 280's loop of the apo structure (∼42 Å^2). The observed loop movement is unlikely the result of crystal packing, as the structures of dδc1 and the DLM were determined in space groups $P32_1$ and $P42_12$, respectively, and residues in the 280's loop do not make any contacts with symmetry-related monomers in either unit cell.

Domain 3 is the most flexible domain. C_α rms deviations of ∼1.5 Å for this region are observed when comparing different monomers within each of the dδc2, dδc1, or apo-DLM structures. However, when a comparison of wild-type dδc1, dδc2, and DLM proteins is performed in the context of the entire monomer, a C_α rms deviation of up to ∼4.0 Å is observed in domain 3 of the sulfate-bound DLM protein (Figures 5 and 7B). This deviation is similar to that observed for domain 3 of the dδc1−sulfate complex (Figure 5). Comparison of domain 3 regions from various δ-crystallin structures, independent of the rest of the protein, results in an average C_α rms deviation of ∼0.4 Å for residues 363−464. This suggests that the rigid body movement of domain 3 in the DLM (and dδc1) is a direct consequence of sulfate binding as the deviation for domain 3 of the DLM is greatest when the structural comparison between different δ-crystallin proteins is performed in the context of the entire monomer. The maintenance of hydrogen bonds between S281 and H388, as well as between S282 and R385 in dδc1 upon closure of the 280's loop, results in the rigid body movement of domain 3 (*12*). These interactions are also conserved in the DLM−sulfate complex. The conformational movements of the 280's loop and domain 3 in the DLM−sulfate complex result in a more compact structure, as indicated by the buried surface area (*16*) of 29 510 Å^2, compared to the area of 27 467 Å^2 for the apo-DLM structure.

In the absence of sulfate, the 280's loop of the DLM protein is in the "open" conformation and there is no rigid body movement of domain 3 (Figure 7A). The conformation of the 280's loop and domain 3 in the apo DLM protein is similar to that of wild-type dδc2 as revealed by C_α rms deviations of 0.30 and 0.38 Å, respectively, when the A monomers of both proteins are compared (Figure 5). These values are similar to those observed for these regions when different monomers within wild-type dδc2, the dδc1−sulfate complex, or either of the DLM proteins are compared. The DLM resembles dδc2 more closely than dδc1 in amino acid sequence. Therefore, although the DLM is enzymatically inactive, the observed rigid body movement of domain 3 and closure of the 280's loop upon sulfate binding (Figure 7) in this mutant reinforce the idea that these conformational movements may be relevant to catalysis in the active dδc2 protein. Formation of the carbanion or aci-carboxylate intermediates has been hypothesized to trigger closure of the 280's loop in the active dδc2 protein (*27*). The charge distribution of the sulfate ion resembles the arrangement of negative charge on the carboxylate groups of the aci-carboxylate intermediate. Determining the structure of dδc2 with a substrate analogue or inhibitor with the same geometry and charge distribution as the aci-carboxylate intermediate would allow this hypothesis to be tested. Nitro analogues of

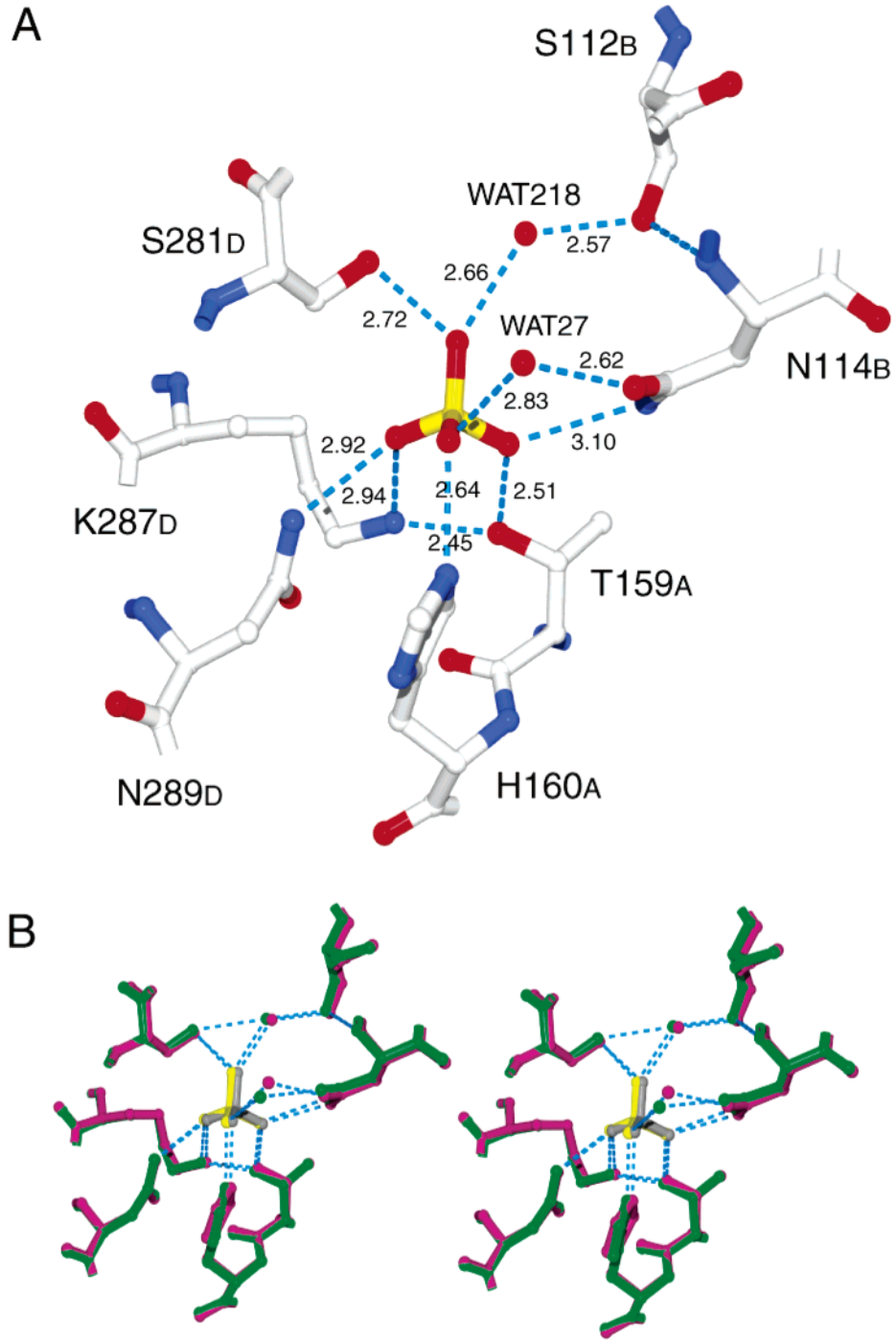

FIGURE 6: (A) Sulfate ion interactions with active site residues and water molecules in the DLM−sulfate complex. The monomer to which each residue belongs is indicated by the subscripted letter following the residue number. (B) Stereoview of the structural superposition of dδc1 (green) and the DLM−sulfate complex (pink), demonstrating the similarity of the sulfate ion interactions in both proteins. The orientation of the figure and labeling for the active site residues is as shown in panel A. In panel B, the sulfate ions bound to dδc1 and the DLM are shown in yellow and gray, respectively. Hydrogen bonds are shown as blue dashed lines with the indicated distances in angstroms.

argininosuccinate in the nitronate form have been shown to be strong competitive inhibitors for bovine ASL *(28)*.

Remaining Amino Acid Substitutions. The structure determination of the DLM has revealed that the 20's and 70's loops in the protein adopt conformations closely resembling those of dδc2. The absence of enzymatic activity in the DLM is puzzling and unexpected, as mutagenesis of the residues that differed between dδc1 and dδc2 in the 20's and 70's loops was expected to restore activity in the inactive isoform. However, the DLM differs from the active 211-chimera in five additional amino acid substitutions. These substitutions are at positions 9, 14, 41, 43, and 115 (Figures 1 and 2). These residues are M, V, A, I, and E in the DLM and dδc1, respectively, and W, S, G, M, and D in dδc2 and the 211-chimera, respectively. Dδc2 also has a two-amino acid insert at residue 5 (Figure 1); this insertion is not found in any other ASL−δ2-crystallin sequence and is therefore not believed to be important enzymatically. The conservative nature of the substitutions at positions 41, 43, and 115 of dδc1 would lead us to postulate that these residues will not

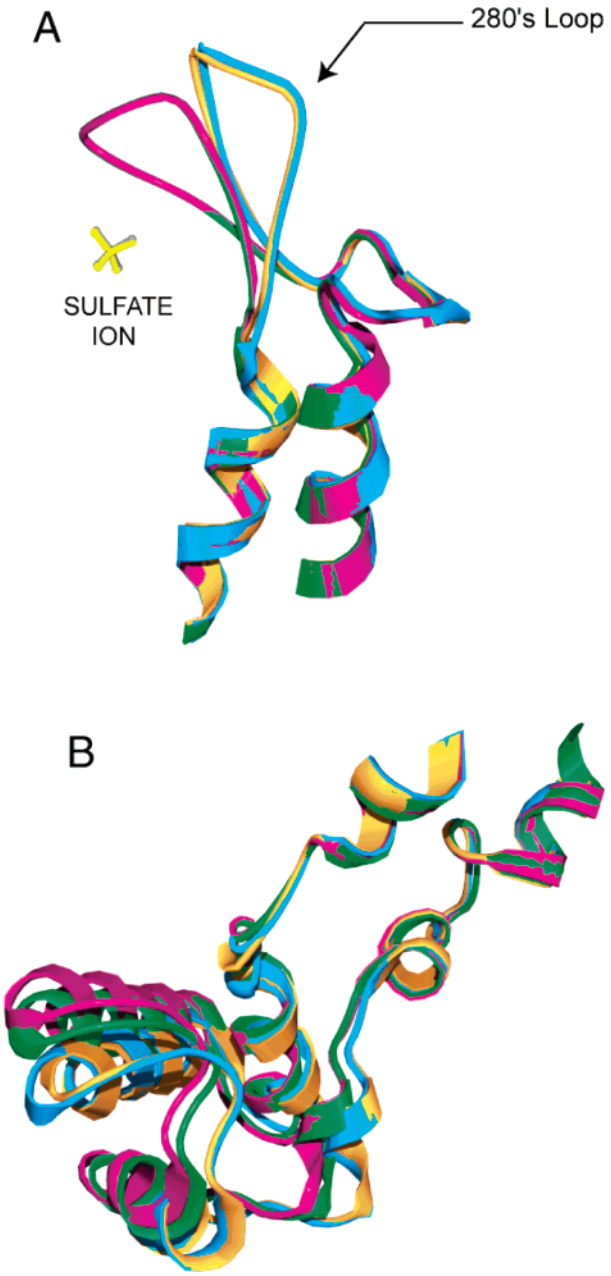

FIGURE 7: Conformational differences observed in (A) the 280's loop (residues 255−300 shown) and (B) domain 3 (residues 357−450 shown) among dδc1 (green), the DLM−sulfate complex (pink), apo-DLM (blue), and dδc2 (orange). In panel A, the sulfate ions bound to dδc1 and the DLM are shown in gray and yellow, respectively.

be major determinants in conferring enzymatic activity to the DLM.

From a structural perspective, the importance of the residues in the N-terminal arm is difficult to assess, as there is no electron density for the first 15−17 residues in any of the δ-crystallin structures determined to date (*6*, *7*, *12*, *13*, *24*). Nevertheless, the N-terminal arm has been shown to be important enzymatically, as the R12Q mutation in ASL results in the genetic disease *argininosuccinic aciduria* (*29*). The R12Q ASL mutant exhibits only 10% of the wild-type activity *in vitro* (*30*). Residues 6−17 of the protein were observed in the Q286R ASL structure in one of the four monomers (*30*). Residues 10−16 form a loop and are within hydrogen bonding distance of R141, T142, Y277, Q344, and S351 from a neighboring monomer. van der Waals interactions are also observed between W9 and R146. Small perturbations to residues involved either directly or indirectly in substrate binding have been shown to have a significant impact on catalysis (*6*), suggesting that the amino acids substitutions at positions 9 and 14 may have a negative impact on catalysis. Of the two substitutions, we hypothesize that the M9W substitution will have the largest effect on catalysis, as human ASL, like dδc1 and the DLM, also has a valine at residue 14. An alternative hypothesis is that the absence of favorable interactions between residues in the N-terminal arm and those found in the neighboring monomer may force the N-terminal arm to adopt a different conformation (*30*). This conformation could be one that would interfere with movement of the 280's loop and domain 3 and, as a result, prevent catalysis from occurring.

Despite the fact that W9, S14, A41, M43, and D115 have not been previously implicated in substrate binding or catalysis in dδc2, our data clearly suggest an enzymatic role for some or all of these residues. Also of interest is the observation that W9, G41, and D115 are highly conserved across the 89 available ASL−δ2-crystallin sequences, with 68, 89, and 100% conservation, respectively. Site-directed mutagenesis of the five remaining amino acid substitutions in the DLM to those of the dδc2 protein will establish which residue(s) is required for activity, as well as the precise role of each in either substrate binding or catalysis.

CONCLUSIONS

The DLM was found to be enzymatically inactive via kinetic analysis. This result was unexpected as domain exchange experiments have previously shown that ASL activity could be restored in dδc1 by replacing its first structural domain with that of dδc2, and that the 20's and 70's loops exhibited the largest conformational differences when comparing the active and inactive isoforms. Substrate binding by the DLM and dδc1 could not be detected by ITC, providing strong evidence that these proteins are catalytically inactive because they are unable to bind argininosuccinate. The determination of the structure of the DLM has established that the 20's and 70's loops in this protein adopt conformations similar to those found in wild-type dδc2. Clearly, one or more of the remaining amino acid substitutions in domain 1 of the DLM play a critical role in substrate binding or catalysis. Site-directed mutagenesis of these residues in both dδc2 and the DLM is currently in progress and will help to establish which residue(s) is absolutely required for activity and the importance of the N-terminal arm. The structure of a DLM−sulfate complex revealed that conformational changes in the 280's loop and domain 3, similar to those observed in dδc1, occur in the DLM upon sulfate binding. Given that the amino acid sequence of the DLM is more similar to that of dδc2 than to that of dδc1, this result provides additional evidence that these events may occur in the active form of the protein on substrate binding or catalysis.

REFERENCES

1. Piatigorsky, J., O'Brien, W. E., Norman, B. L., Kalumuck, K., Wistow, G. J., Borras, T., Nickerson, J. M., and Wawrousek, E. F. (1988) Gene sharing by δ-crystallin and argininosuccinate lyase, *Proc. Natl. Acad. Sci. U.S.A. 85*, 3479−3483.

2. Weaver, T. M., Levitt, D. G., Donnelly, M. I., Stevens, P. P., and Banaszak, L. J. (1995) The multisubunit active site of fumarase C from *Escherichia coli*, *Nat. Struct. Biol. 2*, 654−662.
3. Stone, R. L., Zalkin, H., and Dixon, J. E. (1993) Expression, purification, and kinetic characterization of recombinant human adenylosuccinate lyase, *J. Biol. Chem. 268*, 19710−19716.
4. Shi, W., Dunbar, J., Jayasekera, M. M., Viola, R. E., and Farber, G. K. (1997) The structure of L-aspartate ammonia-lyase from *Escherichia coli*, *Biochemistry 36*, 9136−9144.
5. Williams, S. E., Woolridge, E. M., Ransom, S. C., Landro, J. A., Babbitt, P. C., and Kozarich, J. W. (1992) 3-Carboxy-cis,cis-muconate lactonizing enzyme from *Pseudomonas putida* is homologous to the class II fumarase family: a new reaction in the evolution of a mechanistic motif, *Biochemistry 31*, 9768−9776.
6. Sampaleanu, L. M., Yu, B., and Howell, P. L. (2002) Mutational analysis of duck δ2 crystallin and the structure of an inactive mutant with bound substrate provide insight into the enzymatic mechanism of argininosuccinate lyase, *J. Biol. Chem. 277*, 4166−4175.
7. Vallee, F., Turner, M. A., Lindley, P. L., and Howell, P. L. (1999) Crystal structure of an inactive duck delta II crystallin mutant with bound argininosuccinate, *Biochemistry 38*, 2425−2434.
8. Lee, T. T., Worby, C., Bao, Z. Q., Dixon, J. E., and Colman, R. F. (1999) His68 and His141 are critical contributors to the intersubunit catalytic site of adenylosuccinate lyase of *Bacillus subtilis*, *Biochemistry 38*, 22−32.
9. Weaver, T., Lees, M., and Banaszak, L. (1997) Mutations of fumarase that distinguish between the active site and a nearby dicarboxylic acid binding site, *Protein Sci. 6*, 834−842.
10. Saribas, A. S., Schindler, J. F., and Viola, R. E. (1994) Mutagenic investigation of conserved functional amino acids in *Escherichia coli* L-aspartase, *J. Biol. Chem. 269*, 6313−6319.
11. Sampaleanu, L. M., Davidson, A. R., Graham, C., Wistow, G. J., and Howell, P. L. (1999) Domain exchange experiments in duck δ-crystallins: functional and evolutionary implications, *Protein Sci. 8*, 529−537.
12. Sampaleanu, L. M., Vallee, F., Slingsby, C., and Howell, P. L. (2001) Structural studies of duck δ1 and δ2 crystallin suggest conformational changes occur during catalysis, *Biochemistry 40*, 2732−2742.
13. Abu-Abed, M., Turner, M. A., Vallee, F., Simpson, A., Slingsby, C., and Howell, P. L. (1997) Structural comparison of the enzymatically active and inactive forms of δ crystallin and the role of histidine 91, *Biochemistry 36*, 14012−14022.
14. Pflugrath, J. W. (1999) The finer things in X-ray diffraction data collection, *Acta Crystallogr. D55* (Part 10), 1718−1725.
15. Blessing, R. H. (1987) Data reduction and error analysis for accurate single-crystal diffraction intensities, *Cryst. Rev. 1*, 3−58.
16. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination, *Acta Crystallogr. D54* (Part 5), 905−921.
17. Pannu, N. S., Murshudov, G. N., Dodson, E. J., and Read, R. J. (1998) Incorporation of prior phase information strengthens maximum-likelihood structure refinement, *Acta Crystallogr. D54*, 1285−1294.
18. Adams, P. D., Pannu, N. S., Read, R. J., and Brunger, A. T. (1997) Cross-validated maximum likelihood enhances crystallographic simulated annealing refinement, *Proc. Natl. Acad. Sci. U.S.A. 94*, 5018−5023.
19. Wang, B. C. (1985) Resolution of phase ambiguity in macromolecular crystallography, *Methods Enzymol. 115*, 90−112.
20. McRee, D. E. (1999) XtalView/Xfit: A versatile program for manipulating atomic coordinates and electron density, *J. Struct. Biol. 125*, 156−165.
21. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures, *J. Appl. Crystallogr. 26*, 283−291.
22. Kaplan, W., and Littlejohn, T. G. (2001) Swiss-PDB Viewer (Deep View), *Briefings Bioinf. 2*, 195−197.
23. Walker, D. C., McCloskey, D. A., Simard, L. R., and McInnes, R. R. (1990) Molecular analysis of human argininosuccinate lyase: mutant characterization and alternative splicing of the coding region, *Proc. Natl. Acad. Sci. U.S.A. 87*, 9625−9629.
24. Simpson, A., Bateman, O., Driessen, H., Lindley, P., Moss, D., Mylvaganam, S., Narebor, E., and Slingsby, C. (1994) The structure of avian eye lens δ-crystallin reveals a new fold for a superfamily of oligomeric enzymes, *Nat. Struct. Biol. 1*, 724−734.
25. Weaver, T., and Banaszak, L. (1996) Crystallographic studies of the catalytic and a second site in fumarase C from *Escherichia coli*, *Biochemistry 35*, 13955−13965.
26. Chakraborty, A. R., Davidson, A., and Howell, P. L. (1999) Mutational analysis of amino acid residues involved in argininosuccinate lyase activity in duck δII crystallin, *Biochemistry 38*, 2435−2443.
27. Sampaleanu, L. M., Codding, P. W., Lobsanov, Y. D., Tsai, M., Smith, G. D., Horvartin, C., and Howell, P. L. (2004) Structural studies of duck δ2 crystallin mutants provide insight into the role of T161 and the 280's loop in catalysis, *Biochem. J.*, in press.
28. Raushel, F. M. (1984) Nitro analogs of substrates for argininosuccinate synthetase and argininosuccinate lyase, *Arch. Biochem. Biophys. 232*, 520−525.
29. Craig, H. J. (1994) Molecular analysis of intragenic complementation at the human argininosuccinate lyase locus, Ph.D. Thesis, University of Toronto, Toronto, ON.
30. Sampaleanu, L. M., Vallee, F., Thompson, G. D., and Howell, P. L. (2001) Three-dimensional structure of the argininosuccinate lyase frequently complementing allele Q286R, *Biochemistry 40*, 15570−15580.

BI0489006