Three-Dimensional Structure of the Argininosuccinate Lyase Frequently Complementing Allele Q286R[†,‡]

Liliana M. Sampaleanu,[§,||,⊥] François Vallée,[§,||,#] Gawen D. Thompson,[||,⊥] and P. Lynne Howell*[,||,⊥]

Structural Biology and Biochemistry, Hospital for Sick Children, 555 University Avenue, Toronto, M5G 1X8, Ontario, Canada, and Department of Biochemistry, University of Toronto, Toronto, M5S 1A8, Ontario, Canada

Received July 20, 2001; Revised Manuscript Received October 18, 2001

ABSTRACT: Argininosuccinate lyase (ASL) catalyzes the reversible breakdown of argininosuccinate to arginine and fumarate, a reaction involved in the biosynthesis of arginine in all species and in the production of urea in ureotelic species. In humans, mutations in the enzyme result in the autosomal recessive disorder *argininosuccinic aciduria*. Intragenic complementation has been demonstrated to occur at the ASL locus, with two distinct classes of ASL-deficient strains having been identified, the frequent and high-activity complementers. The frequent complementers participate in the majority of the complementation events observed and were found to be either homozygous or heterozygous for a glutamine to arginine mutation at residue 286. The three-dimensional structure of the frequently complementing allele Q286R has been determined at 2.65 Å resolution. This is the first high-resolution structure of human ASL. Comparison of this structure with the structures of wild-type and mutant duck δ1 and δ2 crystallins suggests that the Q286R mutation may sterically and/or electrostatically hinder a conformational change in the 280's loop (residues 270−290) and domain 3 that is thought to be necessary for catalysis to occur. The comparison also suggests that residues other than R33, F333, and D337 play a role in maintaining the structural integrity of domain 1 and reinforces the suggestion that residues 74−89 require a particular conformation for catalysis. The electron density has enabled the structure of residues 6−18 to be modeled for the first time. Residues 7−9 and 15−18 are in type IV β-turns and are connected by a loop. The conformation observed is stabilized, in part, by a salt bridge between the side chains of R12 and D18. Although the disease causing mutation R12Q would disrupt this salt bridge, it is unclear why this mutation has such a significant effect on the catalytic activity as residues 1−18 are disordered in all other δ-crystallin structures determined to date.

Argininosuccinate lyase catalyzes the reversible breakdown of argininosuccinate to arginine and fumarate, a reaction involved in the synthesis of arginine in all organisms and in the urea cycle in ureotelic species. In humans, mutations in ASL[1] result in the clinical condition *argininosuccinic aciduria* (*1*, *2*), an autosomal recessive disorder that displays considerable clinical and genetic heterogeneity. The clinical heterogeneity is manifested by variations in the age of onset and the severity of symptoms, with three distinct clinical phenotypes having been identified: neonatal, subacute, and late-onset. At present, the biochemical basis for this clinical variation is unclear as there is little correlation between the clinical phenotype and the residual enzymatic activity found in cultured fibroblasts or other tissues. The genetic heterogeneity was identified by complementation analysis using cultured fibroblasts from 28 unrelated patients (*3*). While all strains mapped to a single complementation group (i.e., affected a single locus), 12 distinct subgroups were identified, suggesting that extensive interallelic or intragenic complementation occurs (*4*). Intragenic complementation is a phenomenon that occurs when a multimeric protein is formed from subunits produced by two differently mutated alleles of the same gene. On complementation, a partially functional hybrid protein is produced from the two distinct types of mutant subunit, neither of which individually have appreciable enzymatic activity.

Analysis of the molecular basis of intragenic complementation at the ASL locus has identified the molecular defects in two classes of ASL-deficient strains. The frequent complementers, strains that participate in the majority of complementation events, were found to be either homozygous or heterozygous for a glutamine to arginine mutation at residue 286 (Q286R),[2] while the high-activity complementers, strains in which complementation is associated with the restoration of the greatest amounts of ASL activity, were

[†] Supported by grants from CIHR and NSERC to P.L.H. and by a Hospital for Sick Children graduate research scholarship to L.M.S.

[‡] The coordinates have been deposited in the Brookhaven Protein DataBank, accession number 1K62.

* Correspondence should be addressed to this author. Email: howell@sickkids.on.ca, Telephone: (416) 813-5378, Fax: (416) 813-5022.

[§] These authors contributed equally to the work described.

[||] Hospital for Sick Children.

[⊥] University of Toronto.

[#] Current address: Integrative Proteomics Inc., 100 College St., Toronto, M5G 1L5, Canada.

[1] Abbreviations: NCS, noncrystallographic symmetry; ASL, argininosuccinate lyase; Q286R ASL, glutamine to arginine mutation at residue 286 of human argininosuccinate lyase; tδc1, turkey δ1 crystallin; dδc1, duck δ1 crystallin; dδc2, duck δ2 crystallin; H91N dδc2, histidine to asparagine mutation at residue 91* of duck δ2 crystallin; H162N dδc2, histidine to asparagine mutation at residue 162* of duck δ2 crystallin.

10.1021/bi011525m CCC: $20.00
Published on Web 11/30/2001

found to be either homozygous or heterozygous for an aspartate to glycine mutation at residue 87 (D87G) (*5*). Kinetic analysis of both mutants individually in patient cell lines (*5*) or COS cell transfection assays (*6*) or kinetic analysis of purified protein from a recombinant *E. coli* expression system (*7*) shows little to no ASL activity. The complementation of these mutants and the recovery of ASL activity have been demonstrated in COS cells transfected with both the D87G and Q286R mutants (*5*, *6*), by their coexpression in *E. coli*, and also by mixing purified recombinant protein in vitro (*7*). In all cases, the heterotetrameric protein was found to exhibit approximately 30% wild-type ASL activity.

ASL belongs to a superfamily of homotetrameric enzymes that includes δ crystallin, class II fumarase, aspartase, adenylosuccinase lyase, and 3-carboxy-*cis*,*cis*-muconate lactonizing enzyme. These enzymes all have similar three-dimensional structures (*8*−*13*) and catalyze β-elimination reactions involving the cleavage of a carbon−oxygen or carbon−nitrogen bond with the formation of fumarate as one of the products. Although the overall sequence similarity between these proteins is low (only 15% sequence identity), there are three regions of highly conserved residues, denoted C1, C2, and C3 (Figure 1). These signature sequences are spatially distant from each other in the monomer but cluster together in the tetramer to form a multisubunit active site, in which regions C1, C2, and C3 are each contributed by a different monomer (Figure 2). ASL is directly related to δ-crystallin (64−71% sequence identity), the major structural component of avian and reptilian eye lenses. Modification of gene expression, overexpressed the ancestral ASL in the avian eye lens where subsequent gene duplication has resulted in two crystallin isoforms, δ1 and δ2 (*14*, *15*). While the role of both crystallin proteins is to maintain the physical properties of the lens (e.g., transparency and refractive index), the δ2 protein has also retained ASL activity and is the ASL ortholog in nonlens tissues. In contrast, δ1 crystallin has evolved and is no longer enzymatically active.

To understand how complementation of the D87G and Q286R ASL mutants leads to partial recovery of ASL activity in a heterotetrameric protein, it is first necessary to understand how the mutation affects the structure and function of the homotetrameric protein. The low-resolution 4 Å structure of the wild-type human ASL (*12*) and the high-resolution structures of duck δ-crystallin (*9*, *16*, *17*) have enabled the location of each mutation to be mapped. Q286 is part of superfamily signature sequence C3 and is highly conserved across the ASL/δ-crystallin family. Q286 is located on a highly flexible loop in the active site in close proximity to several residues (e.g., S281and K287) that have been shown to be important for catalysis. D87 is also located in the active site region, at the beginning of a short helix, α5, and close to H89, a residue important for substrate binding (*9*). The side chain of D87 also forms a cap on helix α5. In each of the four independent active sites, D87 and Q286 are contributed by different monomers. If both mutations are present in one active site, the symmetry of the molecule dictates that there will be at least one active site with no mutations. It is these "wild-type" active sites that are responsible for the recovery of ASL activity observed in the hybrid protein. We have previously hypothesized (*12*) that the conversion of the noncharged polar glutamine residue into a larger positively charged arginine residue might enable a salt bridge to be formed between Q286R and either D367 or E399 of a neighboring subunit. This additional interaction was proposed to affect the conformation of the loop to which Q286R is attached either by rigidifying it and/or by preventing a conformational change that may be necessary for catalysis to occur.

We present here the structure determination and refinement of the Q286R ASL mutant and its comparison to previously determined δ-crystallin structures. At 2.65 Å, this represents the first high-resolution structure of human ASL. The results suggest that the Q286R mutation may sterically and/or electrostatically hinder a conformational change in the 280's loop (residues 270−290) and domain 3 that has been suggested to be important for catalysis (*17*), that residues other than R33, F333, and D337 play a role in maintaining the structural integrity of domain 1 and reinforce the suggestion that residues 74−89 require a particular conformation for catalysis to occur. The structure has also enabled the conformation of residues 6−18 to be modeled for the first time, providing insight into the effect that the disease causing mutation R12Q may have on the catalytic activity.

MATERIALS AND METHODS

Expression and Purification of the Q286R Human ASL Mutant. The Q286R mutant was generated using the wild-type human ASL cDNA in a pET-3c expression vector (Novagen) (gift of Dr. M. Hershfield) and the Q286R cDNA in a pESP-SVTEXP vector (gift of Dr. R. McInnes). These vectors, referred to as pET-3c-ASL and pESP-Q286R, respectively, were digested with the restriction enzymes *Spe*I and *Sfi*I (New England Biolabs). The R286 mutant fragment and the pET-3c-ASL vector with the fragment containing Q286 excised were extracted from agarose gels and dephosphorylated. T4 DNA ligase was subsequently used to ligate the purified fragments and generate the pET3c-Q286R vector. The presence of the Q to R mutation was confirmed by DNA sequencing. The Q286R mutant was expressed in *E. coli* and purified as described previously for the wild-type protein (*18*).

Kinetic Analysis. The enzymatic activity of Q286R and wild-type ASL proteins was determined spectrophotometrically. The kinetic assay was as described in Yu et al. (*19*). In brief, a stock solution of argininosuccinate (Sigma-Aldrich) was prepared in reaction buffer (20 mM potassium phosphate, pH 7.6, 1 mM EDTA) such that the final concentration of the substrate in the reaction could range from 0.02 to 2.0 mM. The reaction was initiated by adding 20−30 μL of the protein solution (10−20 μg of protein) to the reaction mixture for a total reaction volume of 500 μL. The reaction was monitored by measuring the increase of fumarate at 240 nm ($\epsilon = 2.44\ mM^{-1}\ cm^{-1}$). All assays were performed in triplicate. Initial velocities were used to determine the kinetic parameters (Table 1).

Circular Dichroism Spectroscopy. The CD spectra of the Q286R and wild-type proteins were measured on an AVIV circular dichroism spectrophotometer (model 62A DS) as

[2] Please note that the amino acid numbering throughout this paper corresponds to ASL/dδc1. dδc2 has a two-residue inset at amino acid 5. The numbering of the dδc2 mutant structures has been maintained in order to facilitate referencing to previously published papers.

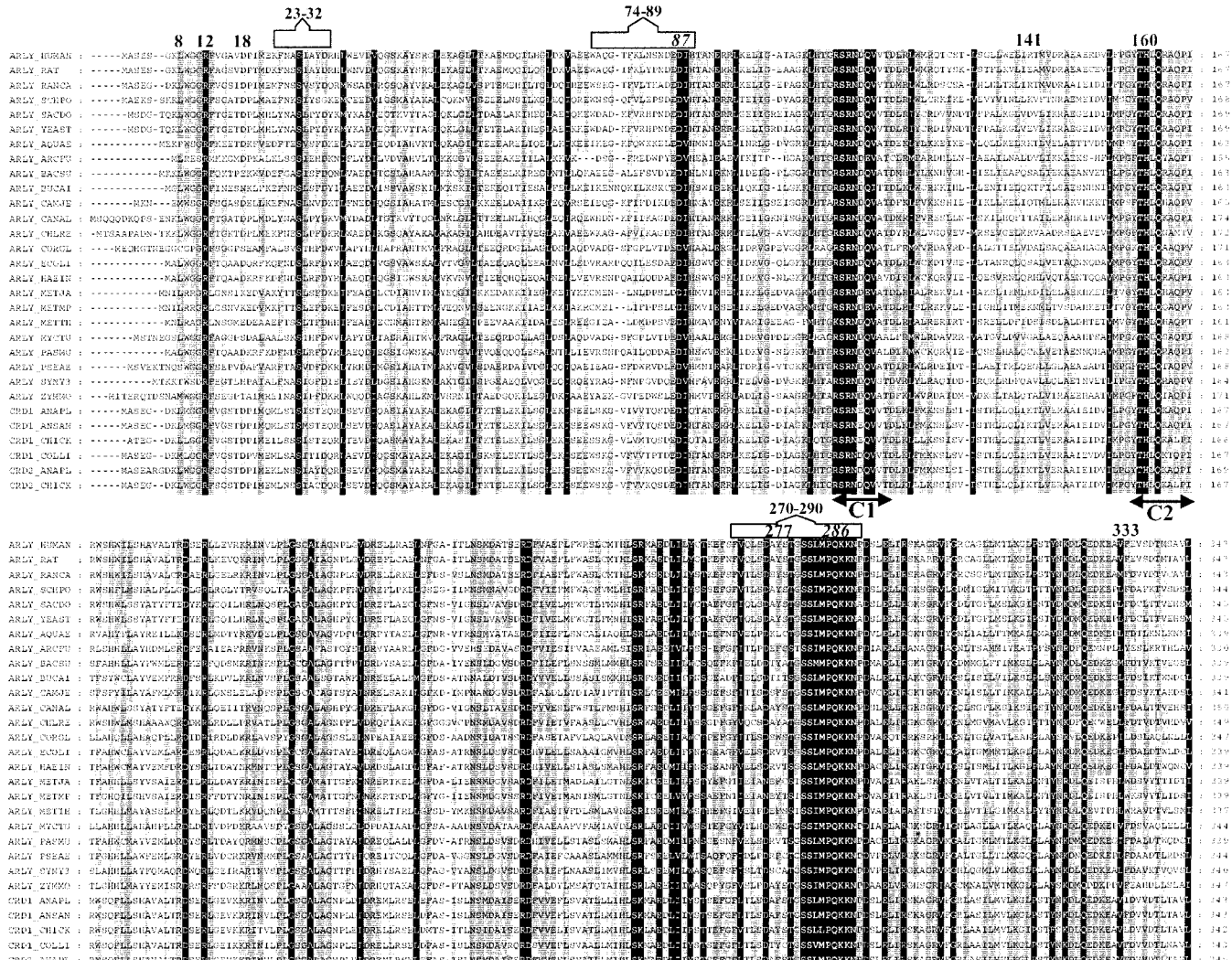

FIGURE 1: Amino acid sequence comparison of the ASL/δ-crystallin family. Amino acids residues that show 100%, 80% or greater, or 60% or greater conservation are shaded in black, dark gray, and light gray, respectively. The figure was prepared using the program GeneDoc (*36*). Abbreviations: ARLY, argininosuccinate lyase; CRD1, δ1 crystallin; CRD2, δ2 crystallin. The following letters represent the species names abbreviated according to the SWISS−PROT nomenclature (e.g., ANAPL corresponds to *Anas platyrhynchos* or the domestic duck). The sequence alignment was performed using Clustal X (*27*).

described in Yu et al. (*19*). The protein solution (0.15−0.2 mg/mL in 20 mM potassium phosphate, pH 7.6, 1 mM EDTA, and 1 mM DTT) was scanned from 200 to 260 nm in 1 nm increments. The thermal stability of each protein was also examined by measuring the loss of ellipticity at 222 nm as the protein sample was heat-denatured. The sample was heated from 25 to 101 °C in 2 °C increments with 1 min of equilibration before each reading.

Table 1: Kinetic Parameters of Wild-Type and Q286R Argininosuccinate Lyase

protein	V_{max} (μmol min^{-1} mg^{-1})	K_m (mM)	K_{cat} (s^{-1})	K_{cat}/K_m (M^{-1} s^{-1})	relative catalytic efficiency (%)
wild type	10.36 ± 0.90	0.12 ± 0.01	34.5 ± 3.0	$(2.98 \pm 0.19) \times 10^5$	100 ± 6.4
Q286R	0.18 ± 0.01	0.07 ± 0.01	0.60 ± 0.02	$(8.98 \pm 0.84) \times 10^3$	3.0 ± 0.3
R12Q	0.63 ± 0.03	0.08 ± 0.02	2.10 ± 0.11	$(2.93 \pm 0.11) \times 10^4$	9.82 ± 2.12

Table 2: Summary of Data Collection and Processing Statistics

space group	$P3_121$
cell dimensions	$a = b = 104.2$ Å, $c = 183.02$ Å, $\gamma = 120°$
no. of molecules/asymmetric unit	2
resolution limits (Å)	17−2.65
no. of overall reflections	79294
no. of unique reflections	33242
mean redundancy	2.4
completeness (%)	97.8 (99.2)[a]
average $I/\sigma(I)$	8.1
% reflections with $I > 2\sigma(I)$	72.6 (43.7)[a]
R_{sym}[b]	0.10 (0.46)[a]

[a] Last resolution shell 2.74−2.65 Å. [b] $R_{sym} = \sum|I - \langle I\rangle|/\sum I$ where I is the measured intensity for symmetry-related reflections and $\langle I\rangle$ is the mean intensity for the reflection.

Table 3: Summary of Refinement Statistics

resolution range (Å)	17−2.65
R_{factor} (%)[a]	17.5
R_{free} (%)[b]	23
no. of reflections used in refinement	31478
no. of reflections used to compute R_{free}	3208
no. of non-hydrogen atoms	
protein	7122
solvent	222
mean B-factor (Å^2)	
protein	44.7
per monomer: A/B	44.3/45.0
solvent	42.7
rms deviation from ideal values	
bond lengths (Å)	0.006
bond angles (deg)	1.1
dihedral angles (deg)	18.9
improper angles (deg)	0.73

[a] $R_{factor} = \sum(||F_o| - |F_c||)/\sum|F_o|$. [b] $R_{free} = \sum(||F_{os}| - |F_{cs}||)/\sum|F_{os}|$ where 's' refers to a subset of data not used in the refinement, representing ~10% of the total number of observations.

Crystallization, Data Collection, and Structure Solution. The mutant Q286R ASL was crystallized using the hanging-drop vapor diffusion method. The crystallization conditions were as described for wild-type human ASL (*12*). Diffraction data were measured at room temperature using a Mar345 imaging plate detector mounted on a Rigaku RF rotating-anode generator ($\lambda = 1.541$ Å) and processed using the DENZO/SCALEPACK software package (*20*) (Table 2). The structure of Q286R ASL was solved by molecular replacement (*21, 22*) using a monomer of H91N duck δ2 crystallin (*9*) as the search model. This model was used as at the time it represented the highest resolution structure of an active ASL/δ-crystallin available. Data between 8 and 4 Å resolution were used in all rotation and translational searches. Initial rotation searches produced two solutions, 7.3 and 6.2 standard deviations above the mean and 3.5σ higher than the next peak. The final molecular replacement solution for the two monomers in the asymmetric unit gave a correlation coefficient of 61% and an R_{factor} of 37.5%.

Structure Refinement. The structure was refined using CNS (*22*−*24*) with a maximum likelihood target function, a flat bulk solvent correction, and no low-resolution or sigma 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, σ_A-weighted $2|F_o|-|F_c|$ and $|F_o|-|F_c|$ electron density maps were computed. Corrections to the model were made between each round of refinement with the program TURBO−FRODO (*25*). Initially, the orientation and position of each monomer were refined. Each monomer was then further divided into four sections, consisting of residues 17−112, 113−361, 362−431, and 432−463 which correspond to the three structural domains of the protein and the C-terminal helix, respectively, and additional rounds of rigid-body refinement were performed. The structure was then refined by performing 5 cycles of torsion-angle refinement followed by 3 cycles of Cartesian refinement. 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} to 17.5% and 23.5%, respectively.

Noncrystallographic symmetry (NCS) restraints were applied separately to each section of the monomers (as defined above) with weights chosen to minimize the R_{free}. As the R_{factor} and R_{free} decreased, the NCS restraints were gradually relaxed. The initial round of refinement used the H91N dδc2 model without modification of the amino acid sequence. Prior to the second round of refinement, the residues of the initial model were annotated in accordance to the amino acid sequence of human ASL and the σ_A-weighted electron density maps. In the subsequent rounds of refinement, the N-terminal segment of monomer B (residues 6_B−16_B) and missing C-terminal residues were inserted. The peptide linkage between residues Ser 321 and Thr 322 was also modeled as a cis peptide. This peptide linkage was modeled as a trans peptide linkage in the low-resolution 4 Å structure of wild-type ASL (*12*). The conformation of this peptide bond does not appear to be related to the enzymatic activity as both the inactive dδc1 and the active dδc2 isoforms have a cis peptide at this position (*17*). A total of 222 ordered water molecules, obeying proper hydrogen bond geometry with electron densities greater than 1.0σ and 2.5σ on σ_A-weighted $2|F_o|-|F_c|$ and $|F_o|-|F_c|$ maps, respectively, were gradually included. The stereochemical quality of the model was analyzed by PROCHECK (*26*). The Ramachandran plot shows that 92% of the residues lie in the most favored region with no residues in the disallowed regions. The final refinement statistics are shown in Table 3.

Sequence Alignment and Structural Comparison. Sequence alignment of the ASL superfamily was performed with Clustal X (*27*). Reconstitution of the mutant Q286R ASL and dδc1 tetramers was carried out in TURBO−FRODO (*25*) using the properties of space groups $P3_121$ and $C222$, respectively. Structural alignment was carried out by super-

imposing the tδc1, dδc1, and dδc2 and the mutant H91N and H162N dδc2 structures on the Q286R ASL tetramer using the RIGID option in TURBO−FRODO. In the first step, a rigid-body alignment based on the coordinates of 10 structurally equivalent atoms located at the beginning and the end of each central helix in domain 2 was performed. From this initial superimposition, an iterative least-squares fitting procedure was carried out between all C_α atoms lying at a distance progressively decreasing from 10.0 to 0.1 Å. The root-mean-square deviation (rmsd) of all the C_α atoms was subsequently calculated in TURBO−FRODO.

RESULTS AND DISCUSSION

Kinetic Characterization of the Human Wild-Type and Q286R Mutant ASL Proteins. The effect of the Q286R mutation on the thermodynamic stability and enzymatic activity of ASL was determined from kinetic and spectroscopic analysis of the mutant and wild-type proteins. CD spectra of the wild-type and Q286R mutant proteins were measured from 200 to 260 nm. The spectra displayed minima at 208 and 222 nm, as is characteristic of proteins comprised of largely α-helical secondary structure [see Figure 2 of the following paper (*19*)]. Superposition of the spectra revealed that the Q286R mutation does not significantly alter the secondary structure of the protein. The unfolding of both proteins was monitored by the change in ellipticity at 222 nm as the sample has heat-denatured. The transition from folded to unfolded protein was found to be cooperative and irreversible, and the wild-type and Q286R mutant proteins were found to have comparable thermal stability [see Figure 3 of the following paper, Yu et al. (*19*)] with the midpoint of the transition at approximately 57 °C. The spectroscopic data indicate that the Q286R mutation does not affect either the overall structure or the stability of the ASL protein.

The wild-type ASL and the Q286R mutant both exhibit Michealis−Menten kinetics with no evidence of negative cooperativity. The results of the kinetic analysis show that the catalytic efficiency of the Q286R mutant is only 3% that of the wild-type protein (Table 1). An approximate 2-fold decrease in K_m and a 60-fold reduction in K_{cat} are observed, suggesting that the mutation has less affect on substrate binding than it does on catalysis. The loss of activity observed in the Q286R recombinant ASL protein correlates well with previous results on patient cell lines, as less than 3% ASL activity was observed in a neonatal onset patient homozygous for the Q286R mutation (*28*). In contrast, less than 0.05% ASL activity was observed for the Q286R mutant in COS cell assays (*5*). Although all results indicate a significant reduction in the catalytic activity, there is some discrepancy regarding the extent of this loss. The small variations observed in the quantification of the activity of the mutant protein are most likely due to the indirect methods used to determine protein concentration in the COS cell assays.

Overall Fold. Q286R ASL has the same overall fold as wild-type ASL (*12*), wild-type and mutant duck δ2 crystallins (*9, 16, 17*), and duck and turkey δ1 crystallin (*8, 17*). The ASL/δ-crystallin family also has the same overall topology as *E. coli* and yeast fumarase (*10, 29, 30*), *E. coli* aspartase (*11*), and *Thermotoga maritima* adenylosuccinate lyase (*13*). Each of the monomers present in the asymmetric unit of the

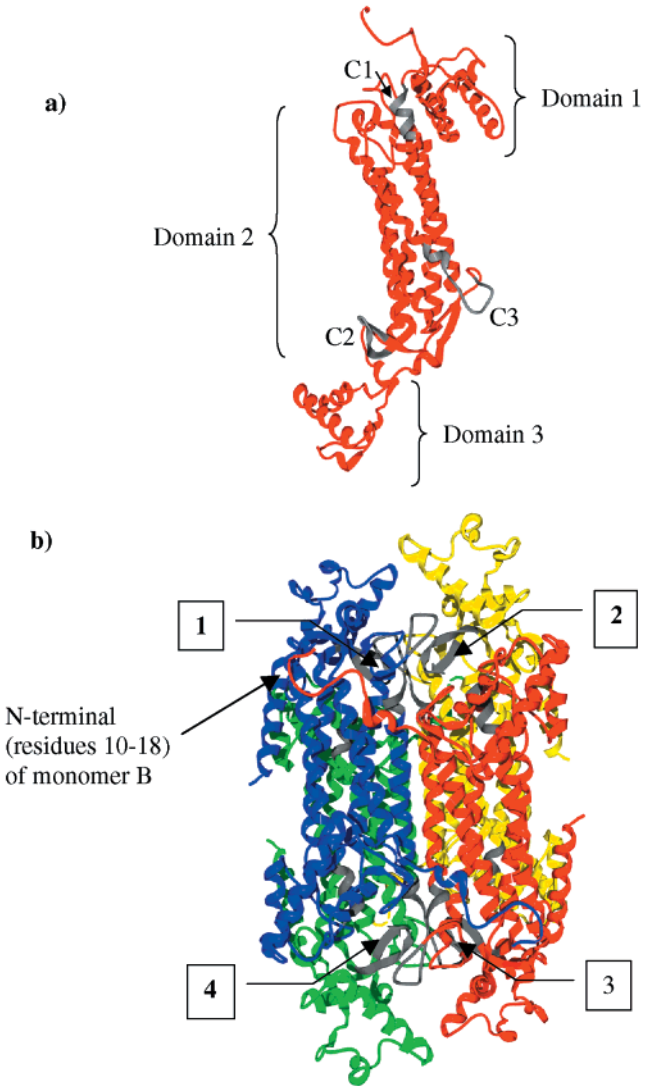

FIGURE 2: Structure of ASL. (a) Schematic diagram of the three-dimensional topology of the ASL monomer. The three structural domains are indicated, and the conserved amino acid regions C1, C2, and C3 are shown in gray. (b) Schematic diagram of the tertiary structure of the tetramer of Q286R ASL. The numbers denote the location of the four active sites. The location of residues 6−18 of monomer B is also indicated. Monomers A, B, C, and D are colored in green, red, blue, and yellow, respectively. Monomers A and B are present in the asymmetric unit.

Q286R ASL structure consist of 452 amino acids, corresponding to residues 15−464. Eight additional N-terminal residues (residues 6−14) are defined in monomer B. The N-terminal residues (1−14) of monomer A could not be modeled owing to the poor quality of the electron density in this region of the structure. The secondary structure of each monomer is predominantly α-helical with each monomer consisting of three distinct structural domains (Figure 2a). Domains 1 and 3 have a similar overall topology and consist of two helix−turn−helix motifs arranged mutually perpendicular to each other. Domain 2 consists of nine helical segments, five of which form a central five-helix bundle with an up−down−up−down−up topology. ASL is functional as a tetramer, and while only two monomers (A and B) are present in the crystallographic asymmetric unit, the symmetry of the space group can be used to generate the active tetramer. The tetramer is formed by two dimers of closely interacting monomers (A/C and B/D). Three of the five

helices of the central five-helix bundle participate in the association of two monomers. Two pairs of closely associated monomers interact to form the tetramer. Each monomer contributes one helix to form a four-helix bundle at the core of the protein. The overall structure of the tetramer exhibits *222* noncrystallographic symmetry (Figure 2b). The structural comparisons between the Q286R ASL and wild-type ASL, and wild-type dδc2 and tδc1 show rms deviations between Cα positions of 0.56, 0.87, and 0.98 Å, respectively, suggesting that there are no major structural differences between these models. There is, however, an rms deviation between Cα positions in Q286R ASL and dδc1 of 1.7 Å. This large difference is related to the conformational changes in domain 3 and the 280's loop seen in the duck δ1 crystallin structure that have been suggested to be important for catalysis (see below) (*17*).

The Active Site and the Q286R Mutation. The active site region was initially identified in the structure of tδc1 (*8*) and subsequently confirmed by structures of the H162N (*16*) and S283A dδc2 (*31*) mutants with bound argininosuccinate as well as *E. coli* fumarase C with bound inhibitors and substrates (*10, 29, 30*). The conserved superfamily sequences C1, C2, and C3 (Figure 1) define the four independent active site clefts located at each 'corner' of the tetramer (Figure 2b). Residue 286 belongs to the signature sequence C3. This signature sequence forms part of a highly flexible loop, called the 280's loop, which consists of residues 270−290. Residue 286 is located close to the tip of this loop. Residues in this loop are typically difficult to model, as the quality of electron density is often poor in this region and variable numbers of residues in this loop are missing from both the H162N dδc2 (*16*) and the *T. maritima* adenylosuccinate lyase structures (*13*). In the Q286R ASL structure, while the position of the main chain atoms of this loop in both monomers in the asymmetric unit can be clearly seen in the electron density map (Figure 3a), little to no electron density exists for the side chains of residue R286 in monomers A and B, or S281 in monomer B. R286 was therefore modeled as a glycine although the nucleotide sequencing data clearly indicate that the glutamine residue has been substituted with an arginine. The superposition of Q286R ASL with those δ-crystallin structures in which the 280's loop has been modeled (*8, 9, 17*) reveals that the loop is in a very similar conformation in both monomers in the asymmetric unit and in all structures examined except the dδ1c structure (Figure 3b). This finding suggests that no significant conformational changes result from the Q286R substitution. The conformation of the 280's loop in the dδ1c structure deviates considerably from all other structures determined to date and is the result of a sulfate anion bound in the active site region of this protein (see below).

To examine further how the Q286R mutation might affect catalysis, the glycine at position 286 was modeled as an arginine residue, and the different preferred rotamer conformations of the arginine side chain were examined. Our analysis suggests that in addition to the previously proposed interactions of R286 with E399 or D367 of domain 3 of a neighboring monomer (*12*), R286 may also potentially interact with S365 of the neighboring monomer and with residues Q272 and P285 on the 280's loop of the same monomer. Our inability to model the side chain of R286 in the Q286R ASL structure, however, provides no evidence that any of these potential interactions are formed.

The effect of the Q286R mutation was therefore examined in light of the conformational changes observed in the structure of dδc1 (*17*). In this structure, a sulfate anion was found in the active site region in a location similar to that of the fumarate moiety of the argininosuccinate substrate observed in the H162N dδc2 structure (*16*). Sulfate is an inhibitor of dδc2/ASL (Sampaleanu and Howell, unpublished results). Two conformational changes are seen in the dδc1 structure. The major structural change observed is an approximately 8 Å shift of the 280's loop. The shift of this loop is coupled with a concerted rigid body movement of domain 3 of a neighboring monomer. To determine what effects these conformational changes may have on the catalytic activity of the Q286R mutant, we modeled the argininosuccinate substrate into the active site of the Q286R mutant in the same position as that observed in H162N dδc2 mutant. The Q286R ASL and dδc1 structures were then superimposed. The van der Waals surface of the two proteins clearly shows that the conformational changes observed in the dδc1 structure would sequester the substrate in the Q286R structure from the solvent (Figure 4). Although these conformational changes have currently only been observed in the inactive dδc1 structure, they are changes that could potentially occur during catalysis, as sequestration of substrate from solvent is a common phenomenon observed in numerous catalytic mechanisms. Replacement of Q286 with a bulkier positively charged arginine side chain in the dδc1 structure would result in possible interactions not only with residues M284, P285, K287, and K288 in the 280's loop but also with residues D367, A370, and T371 in domain 3 of the adjacent monomer. This suggests that the Q286R substitution might sterically and/or electrostatically interfere with the conformational change observed for the 280's loop and domain 3. The substitution could therefore affect catalysis by preventing the conformational change from occurring. This hypothesis is consistent with the results of the kinetic analysis, which show that binding of the substrate is less affected ($\sim$2-fold reduction in K_m) than the rate of catalysis (60-fold reduction in K_{cat}), and with the structure of the argininosuccinate−S283A dδc2 mutant complex (*31*), in which the 280's loop is in the open, substrate-unsequestered conformation. Residues in the 280's loop do not appear to be involved in substrate binding but have been implicated in catalysis. S281[2] has been suggested to be the putative catalytic acid (*17*) while K287 is thought to help stabilize the carbanion intermediate (*32*).

The N-Terminal Residues. Residues in the N-terminal segment of the protein appear to be important for catalysis as an arginine to glutamine mutation at residue 12 has been identified in a patient with the late-onset phenotype of the disease, *argininosuccinic aciduria* (*6*). The R12Q mutant exhibits less than 50% ASL activity in a qualitative biochemical assay that used a crude enzyme extract and monitored the formation of [^{14}C]argininosuccinate (*6*). Amino acid sequence comparison shows that R12 is strictly conserved across all known ASL/δ-crystallin sequences and that L8 and G11 are more than 80% conserved (Figure 1). The N-terminal segment of the protein (residues 1−17) is, however, highly flexible as indicated by the absence of electron density in this region in all δ-crystallin and ASL

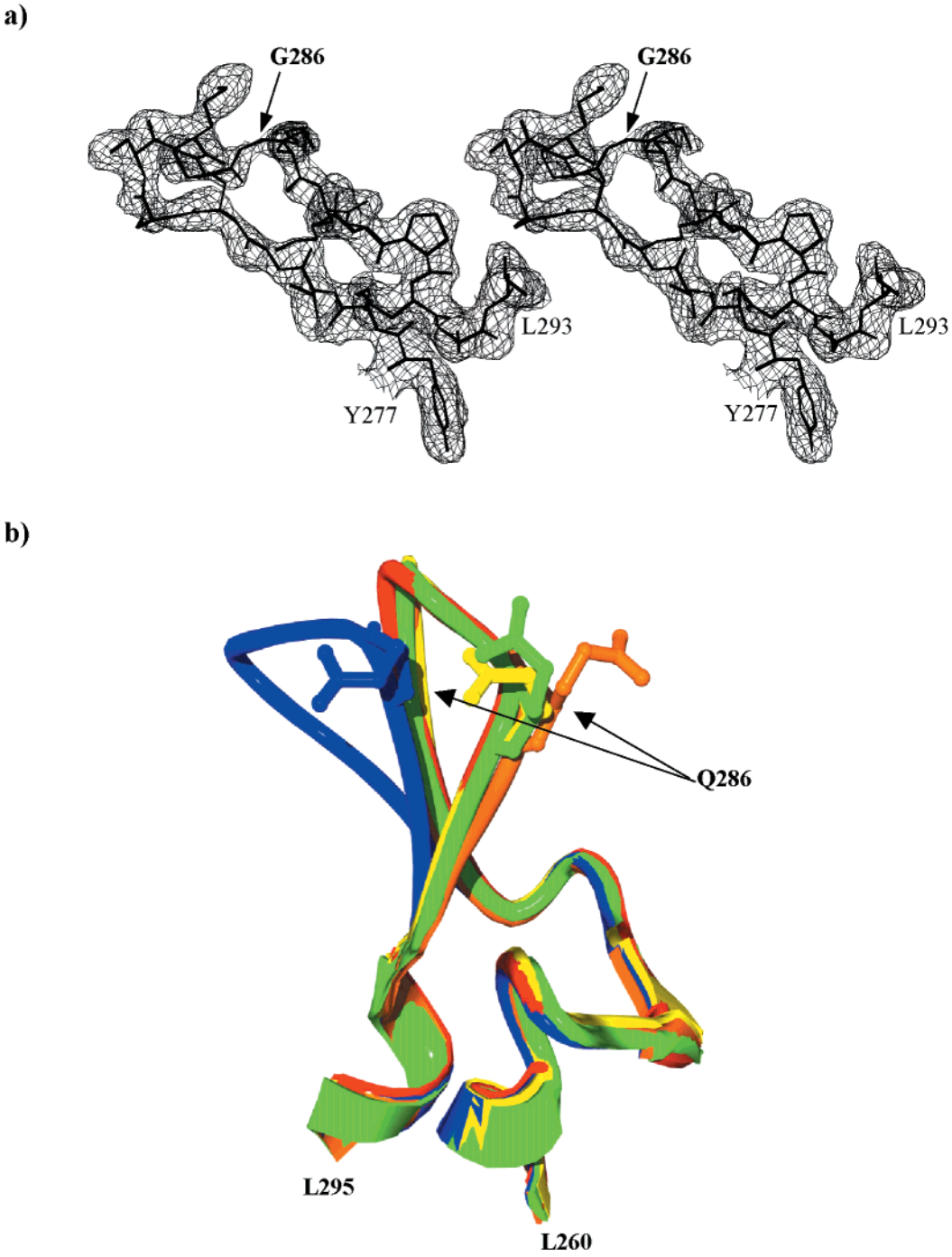

FIGURE 3: The 280's loop. (a) Stereoview of the σ_A-weighted $2|F_o|-|F_c|$ final electron density contoured at 1σ for residues 277−293 of monomer A. (b) Superposition of the 280's loop (the 260−295 region is shown) from Q286R ASL (red), wild-type dδc2 (orange), dδc1 (blue), tδc1 (green), and the mutant H91N dδc2 (yellow) structures.

structures determined to date. In the Q286R ASL structure, the electron density in this region was strong enough to allow residues 6−14 to be modeled in one of the two monomers in the asymmetric unit (see Figure 5a for an example of the quality of electron density for residues 11−18 of monomer B). The *B*-factors for these residues (72.5 Å^2) are, however, significantly higher than the average overall *B*-factor for the structure (45.2 Å^2). The N-terminal arm of monomer B interacts with another monomer (monomer C) in the Q286R ASL tetramer (Figures 2b and 5b). The lack of contacts between this N-terminal segment and other symmetry-related molecules suggests that the conformation of these residues is not dependent on the crystal packing. Residues 7−9 and 15−18 are in type IV β-turns while residues 10−16 form a loop and are within hydrogen bond distance of residues R141, T142, Y277, Q344, and S351 of the neighboring monomer, monomer C (Figure 5b,c). R12 appears to be particularly important in maintaining the conformation of the loop as its Nϵ and Nη2 atoms interact with the Oδ1 and Oδ2 of D18 of the same monomer (Figure 5c). D18 is more than 60% conserved in all ASL family members (Figure 1). The Nη1, amide nitrogen, and carbonyl oxygen atoms of R12 also interact with Oϵ1 of Q344, Oγ of S351, and Oη of Y277 of monomer C, respectively (Figure 5c). Y277, Q344, and S351 are not conserved across the ASL family. For example, Y277 is frequently replaced by phenylalanine while S351 is

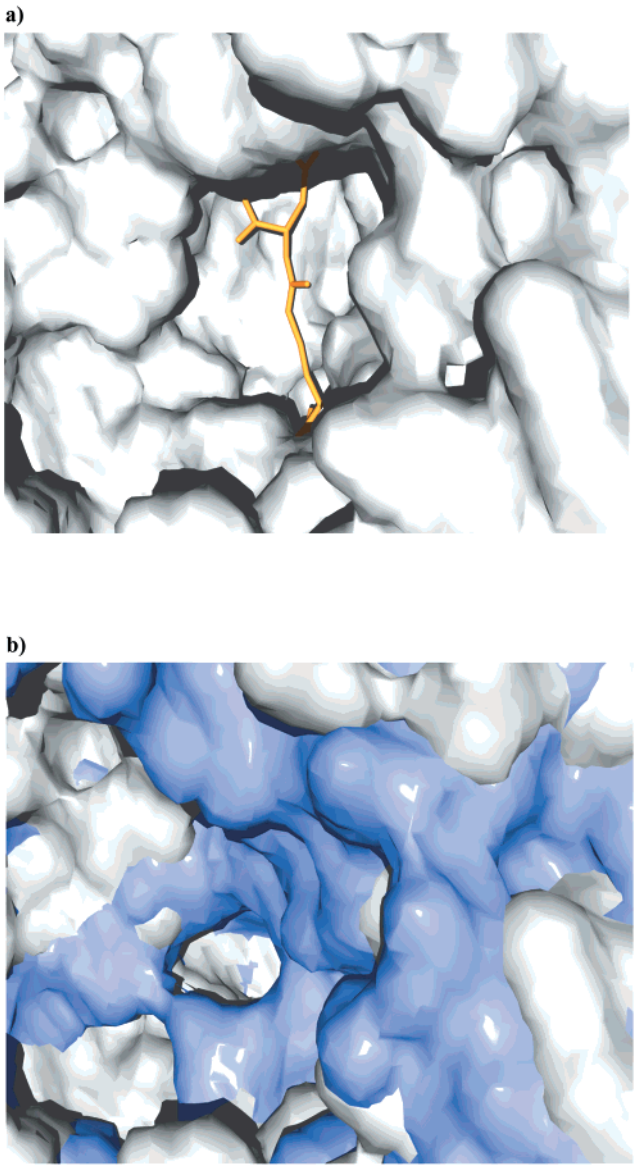

FIGURE 4: Surface representation of the active site region of (a) Q286R ASL (gray) with modeled argininosuccinate substrate (orange). (b) Surface representation of the active site region of Q286R ASL (gray) superimposed with the surface representation for wild-type dδc1 (purple). The conformational change observed in dδc1 sequesters the substrate from the solvent, which is present in this panel but cannot be seen.

replaced in several species with either aspartate or lysine (Figure 1). Residues G10 and G11 are stabilized by the interaction of their amide nitrogen atoms with the Nη2 of R141, and Nη1 of R141 and Oγ1 of T142, respectively. The aromatic side chain of W9 also packs against the side chain of R146 while Nϵ1 of W9 makes a hydrogen bond with the carbonyl oxygen of S351 (Figure 5c).

The role of the N-terminal arm in catalysis is unknown, but it is interesting to note that this region of the protein is in close proximity to the active site (Figure 2b), although in its current conformation R12 would not interact with the substrate. The intrinsic flexibility of residues 1−15 in all other ASL/δ-crystallin structures and the lack of strict sequence conservation of residues involved in stabilizing residues 9−12 in the Q286R ASL structure suggest that the conformation of this segment is unlikely to be conserved across the ASL family. While the substitution of R12 with a glutamine residue would undoubtedly disrupt the salt bridge between R12 and D18, the consequence that this change has on catalysis is hard to predict. One potential consequence of the mutation may be that in the absence of the R12−D18 interaction, the N-terminal arm adopts an alternative conformation that in some way obstructs the active site, hence interfering with substrate binding and catalysis. Alternatively, the N-terminal segment might adopt a conformation in which it interacts with the 280's loop and residues in domain 3, therefore affecting the mobility of these regions that have been hypothesized to be important for catalysis (*17*). The latter suggestion is in agreement with the kinetic analysis of the R12Q ASL mutant that indicates the mutation produces a 18-fold decrease in K_{cat} and a more modest, 2-fold decrease in K_m (Table 1). These results suggest that the mutation affects mainly catalysis, not substrate binding.

Conformational Changes in Domain 1 (23−32 and 74−89 Regions). Superposition of the Q286R monomer A with the equivalent monomers of the available δ-crystallin structures (Figure 6a) reveals significant structural differences in two regions of domain 1, residues 23−32 and 74−89. Domain swapping experiments have shown that domain 1 (residues 1−110) is crucial for enzymatic activity (*33*). ASL activity was recovered in the enzymatically inactive dδc1 when domain 1 was replaced with the equivalent residues from dδc2 (*33*). Previous structural comparisons have also revealed the importance of residues 23−32 for substrate binding and the effect that conformational changes in the 74−89 loop have on neighboring regions involved in substrate binding (*9*, *16*, *17*). The Q286R ASL structure reinforces the importance of the 23−32 loop for substrate binding with its Cα backbone clustering with that of the dδc2 structures (Figure 6). There are several amino acid substitutions in this region (Figure 1). For example, K23 and Y30 in dδc2 are replaced by methionine and threonine, respectively, in dδc1 and by isoleucine and threonine, respectively, in tδc1. In contrast, most of the amino acid differences between dδc2 and ASL in this region are conservative changes (Figure 1).

We have previously suggested that residues R33, F333, and D337 play an important role in maintaining the structural integrity of domain 1 and the overall structure of the δ-crystallin monomer (*17*). Although an ∼4 Å shift in the backbone is observed between residues 23−32 in the δc2 and δc1 structures, the salt bridge between the side chains of R33 and D337 is maintained in all δ-crystallin structures. Additional van der Waals interactions also occur between R33 and residue 333 of domain 2 of the same δ-crystallin monomer (Figure 6b). Residue 333 is isoleucine and leucine in dδc1 and tδc1, respectively, and a phenylalanine in dδc2 and ASL. The same ∼4 Å shift is observed between Q286R ASL and δc1 structures; however, in human ASL, R33 is replaced by histidine. This substitution and the maintenance of same relative conformation of residues 23−35 and 326−337 suggest that other residues contribute to the stabilization the ASL monomer in this region. For example, the H33 side chain establishes favorable van der Waals interactions with the side chains of F333 and R32 (Q32 in dδc2) (Figure 6b). Additionally, a number of bulkier aromatic substitutions occur in ASL (e.g., W35 is a serine in dδc2) leading to favorable packing interactions with other neighboring hydrophobic or aromatic residues (e.g., W35, W74, and F79;

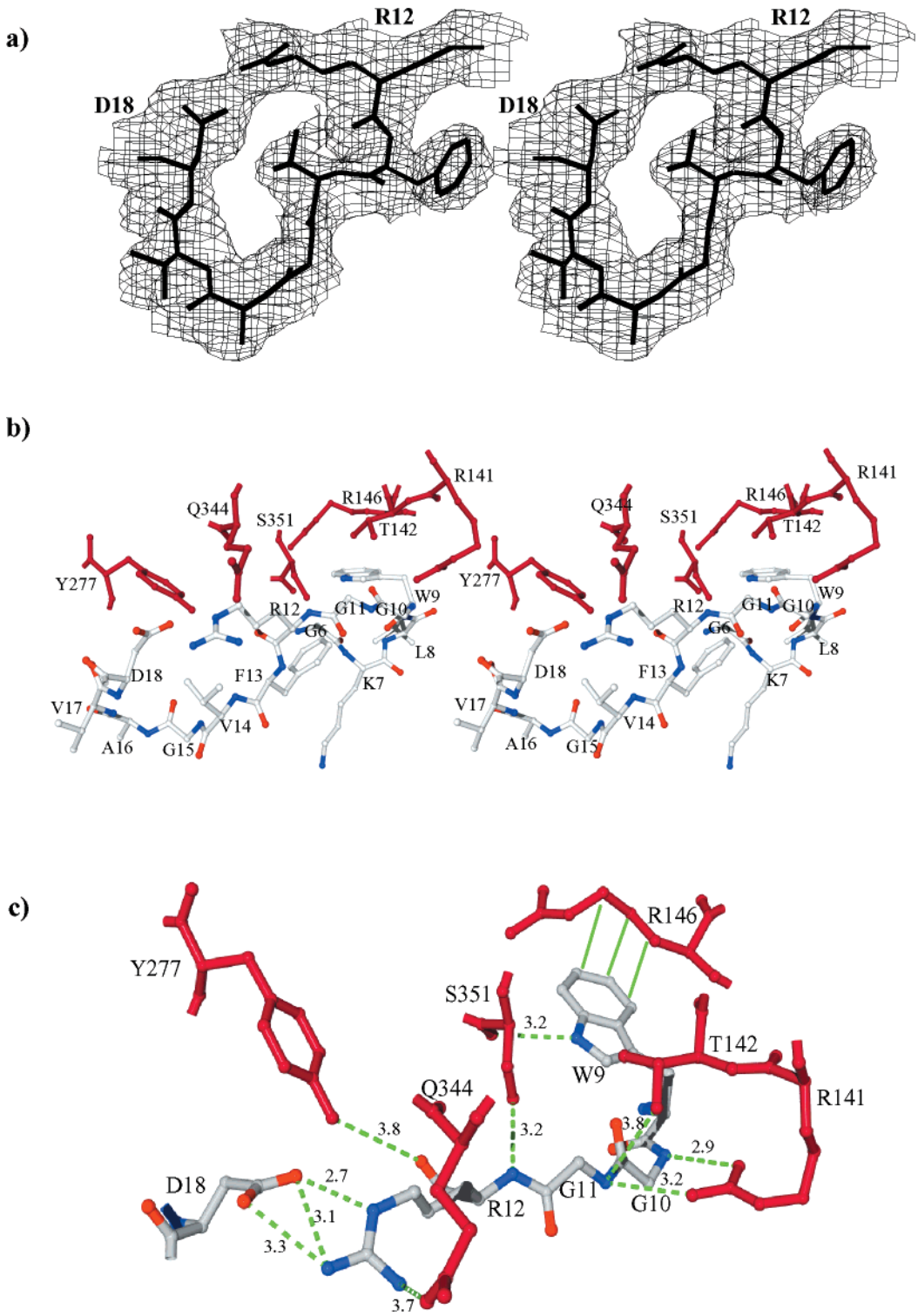

FIGURE 5: N-terminal residues. (a) The σ_A-weighted $|2F_o|-|F_c|$ omit map of N-terminal residues 11−18. The map is contoured at $+1.0\sigma$. (b) Stereoview, in the same orientation as in panel a, of the N-terminal segment of monomer B. Depicted are residues 6−18 of monomer B and residues from monomer C (red) involved in stabilization of the N-terminal segment. (c) Hydrogen bond (dashed green) and van der Waals interactions (solid green) involving the N-terminal residues of monomer B. Distances are given in angstroms.

see Figure 6b). These interactions help stabilize the 23−32 and 326−337 regions of the monomer, regions of the protein that have been shown to be important for substrate binding (*16*, *31*).

The conformation of residues in the 74−89 loop appears to be highly variable with structural differences occurring not only between dδc1 and tδc1 but also between both δc1 structures and the dδc2 and Q286R ASL structures (Figure 6a). The conformation of this loop in Q286R ASL is close to that seen in dδc2 rather than either δc1 structures, which suggests that although the 74−89 loop is flexible, a specific conformation in this region may be required for enzyme catalysis to occur. Several other observations support this hypothesis. The mutation of D87 to glycine in ASL causes the disease *argininosuccinic aciduria* (*28*), while mutation of H89 to asparagine in dδc2 results in a protein with only

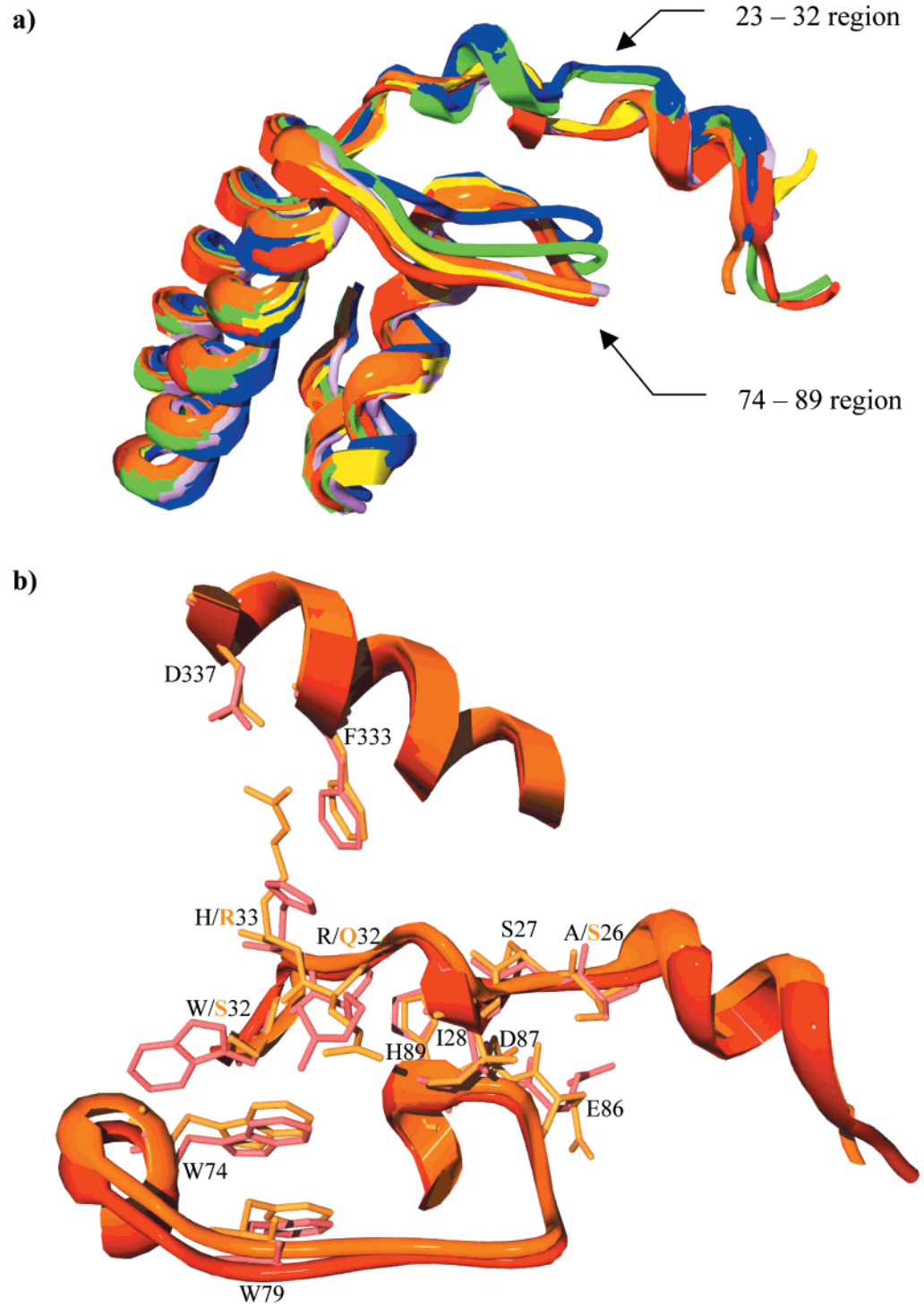

FIGURE 6: Structural comparison of domain 1. (a) Comparison of domain 1 of Q286R ASL (red), wild-type dδc2 (orange), dδc1 (blue), tδc1 (green), and the mutant H91N (yellow) and H162N (purple) dδc2 structures. Arrows indicate the largest conformational differences that occur between the models, the 23−32 and 74−89 loops. (b) Close-up of the 23−32 and 74−89 loops and the neighboring region including residues 326−337. The side chains of key residues are shown to illustrate the contacts that are involved in stabilizing these regions of the monomer.

10% activity (*9*). Both D87 and H89 interact with the substrate (via water molecules) as shown in the H162N and S283A dδc2 structures with bound argininosuccinate (*16, 31*). Mutations and/or conformational changes that affect these side chains might therefore interfere with substrate binding. Additionally, any large conformational variation in the 74−89 loop will affect residues in the neighboring regions, specifically residues 23−32 of domain 1 (Figure 6) and the conserved sequence C1 of domain 2 (Figure 1). Both of these regions contain residues that have been shown to be important for substrate binding (*16, 31, 34*).

Intragenic Complementation. The structural, kinetic, and thermodynamic data presented in this and the following paper by Yu et al. (*19*) reveal that the Q286R mutation does not affect the overall structure and/or stability of the protein. Our current hypothesis is that the Q286R mutation does not affect substrate binding but rather hinders a conformational change in the 280's loop and domain 3 that we believe is

necessary for catalysis to occur (*17*). The structural results presented here therefore reinforce the hypothesis that complementation between the Q286R and D87G mutant alleles occurs through a regeneration of native "wild-type" active sites (*35*). As can be seen in the following paper (*19*), while this hypothesis can account for the complementation and partial recovery of activity between two stable active site mutants, it does not explain how complementation can occur when one of the mutations lies outside the active site region. In the following paper (*19*), we show that stable active mutants such as Q286R can complement with destabilizing, nonactive site mutants such as M360T and A398D by increasing the stability of the heterotetrameric protein. The recovery of activity seen in this type of complementation was significantly less than that observed for the Q286R:D87G complementation event. Between 6 and 15% ASL activity was recovered versus ~35% for the Q286R:D87G complementation event; however, even this level of activity will probably ensure a milder phenotype than would otherwise result when no complementation could occur.

The results presented in this and the following paper (*19*) suggest that examination of genotype−phenotype relationships in genetic diseases involving multimeric proteins will not be straightforward. Correlation will require careful characterization of the effects of individual mutations in order to gain a thorough understanding of how the disease manifests itself and what the consequences to the patient will be.

ACKNOWLEDGMENT

We thank Dr. M. Hershfield and Dr. R. McInnes for the wild-type ASL and Q286R vectors, respectively, B. Yu for performing the kinetic and spectroscopic assays, Dr. A. Davidson for use of his CD spectrophotometer and helpful discussions, and Patrick Yip for technical assistance.

REFERENCES

1. Allan, J. D., Cusworth, D. C., Dent, C. E., and Wilson, V. K. (1958) *Lancet 1*, 182−187.
2. Levy, H. L., Coulombe, J. T., and Shih, V. E. (1980) in *Neonatal Screening for Inborn Errors of Metabolism* (Bickel, H., Guthrie, R., and Hammersen, G., Eds.) pp 89−103, Springer-Verlag, New York.
3. McInnes, R. R., Shih, V., and Chilton, S. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 4480−4484.
4. Simard, L., O'Brien, W. E., and McInnes, R. R. (1986) *Am. J. Hum. Genet. 39*, 38−51.
5. 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.
6. Craig, H. J. (1994) in *Molecular and Medical Genetics*, p 90, University of Toronto, Toronto.
7. Thompson, G. D. (1998) in *Biochemistry*, p 122, University of Toronto, Toronto.
8. Simpson, A., Bateman, O., Driessen, H., Lindley, P., Moss, D., Mylvaganam, S., Narebor, E., and Slingsby, C. (1994) *Nat. Struct. Biol. 1*, 724−734.
9. Abu-Abed, M., Turner, M. A., Vallée, F., Simpson, A., Slingsby, C., and Howell, P. L. (1997) *Biochemistry 36*, 14012−14022.
10. Weaver, T. M., Levitt, D. G., Donnelly, M. I., Stevens, P. P., and Banaszak, L. J. (1995) *Nat. Struct. Biol. 2*, 654−662.
11. Shi, W., Dunbar, J., Jayasekera, M. M., Viola, R. E., and Farber, G. K. (1997) *Biochemistry 36*, 9136−9144.
12. Turner, M. A., Simpson, A., McInnes, R. R., and Howell, P. L. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 9063−9068.
13. Toth, E. A., and Yeates, T. O. (2000) *Struct. Fold. Des. 8*, 163−174.
14. Mori, M., Matsubasa, T., Amaya, Y., and Takiguchi, M. (1990) *Prog. Clin. Biol. Res. 344*, 683−699.
15. 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.
16. Vallée, F., Turner, M. A., Lindley, P. L., and Howell, P. L. (1999) *Biochemistry 38*, 2425−2434.
17. Sampaleanu, L. M., Vallee, F., Slingsby, C., and Howell, P. L. (2001) *Biochemistry 40*, 2732−2742.
18. Turner, M. A., Achyuthan, A. M., Hershfield, M. S., McInnes, R. R., and Howell, P. L. (1994) *J. Mol. Biol. 239*, 336−338.
19. Yu, B., Thompson, G. D., Yip, P., Howell, P. L., and Davidson, A. R. (2001) *Biochemistry 41*, 15581−15590.
20. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276, Part A*, 307−326.
21. Adams, P. D., Pannu, N. S., Read, R. J., and Brunger, A. T. (1999) *Acta Crystallogr., Sect. D: Biol. Crystallogr. 55*, 181−190.
22. 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) *Acta Crystallogr., Sect. D: Biol. Crystallogr. 54*, 905−921.
23. Adams, P. D., Pannu, N. S., Read, R. J., and Brunger, A. T. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 5018−5023.
24. Rice, L. M., and Brunger, A. T. (1994) *Proteins: Struct., Funct., Genet. 19*, 277−290.
25. Roussel, A., and Cambillau, C. (1991) *Turbo-Frodo*, Silicon Graphics, Mountain View, CA.
26. Laskowski, R. A., MacArthour, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallog.*, 283−291.
27. Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G., and Gibson, T. J. (1998) *Trends Biochem. Sci. 23*, 403−405.
28. Walker, D. C., McCloskey, D. A., Simard, L. R., and McInnes, R. R. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 9625−9629.
29. Weaver, T., and Banaszak, L. (1996) *Biochemistry 35*, 13955−13965.
30. 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.
31. Sampaleanu, L. M., Yu, B., and Howell, P. L. (2002) *J. Biol. Chem*. (in press).
32. Saribas, A. S., Schindler, J. F., and Viola, R. E. (1994) *J. Biol. Chem. 269*, 6313−6319.
33. Sampaleanu, L. M., Davidson, A. R., Graham, C., Wistow, G. J., and Howell, P. L. (1999) *Protein Sci. 8*, 529−537.
34. Chakraborty, A. R., Davidson, A., and Howell, P. L. (1999) *Biochemistry 38*, 2435−2443.
35. Howell, P. L., Turner, M. A., Christodoulou, J., Walker, D. C., Craig, H. J., Simard, L. R., Ploder, L., and McInnes, R. R. (1998) *J. Inherit. Metab. Dis. 21*, 72−85.
36. Nicholas, K. B., Nicholas, H. B., Jr., and Deerfield, D. W. (1997) *EMBNEW. NEWS*, 4−14.

BI011525M