

The Structure of the Cataract-Causing P23T Mutant of Human γ D-Crystallin Exhibits Distinctive Local Conformational and Dynamic Changes^{†,‡}

Jinwon Jung,[§] In-Ja L. Byeon,[§] Yongting Wang,^{||} Jonathan King,^{||} and Angela M. Gronenborn^{*,§}

Department of Structural Biology, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, and
Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received December 15, 2008; Revised Manuscript Received February 12, 2009

ABSTRACT: Crystallins are major proteins of the eye lens and essential for lens transparency. Mutations and aging of crystallins cause cataracts, the predominant cause of blindness in the world. In human γ D-crystallin, the P23T mutant is associated with congenital cataracts. Until now, no atomic structural information has been available for this variant. Biophysical analyses of this mutant protein have revealed dramatically reduced solubility compared to that of the wild-type protein due to self-association into higher-molecular weight clusters and aggregates that retain a nativelike conformation within the monomers [Pande, A., et al. (2005) *Biochemistry* 44, 2491–2500]. To elucidate the structure and local conformation around the mutation site, we have determined the solution structure and characterized the protein's dynamic behavior by NMR. Although the global structure is very similar to the X-ray structure of wild-type γ D-crystallin, pivotal local conformational and dynamic differences are caused by the threonine substitution. In particular, in the P23T mutant, the imidazole ring of His22 switches from the predominant N ϵ 2 tautomer in the wild-type protein to the N δ 1 tautomer, and an altered motional behavior of the associated region in the protein is observed. The data support structural changes that may initiate aggregation or polymerization by the mutant protein.

Structural proteins of the eye lens have to be maintained in a functional form over a lifetime. They comprise two protein superfamilies: the β/γ -crystallins and the small heat shock protein family, the α -crystallins (1–4). Knowledge of their structures, stability, and interactions is of fundamental importance for understanding the loss of lens transparency in cataract, the leading causes of vision impairment and blindness in the world with ~17 million cases worldwide per year (5). At present, the only available treatment is surgical intervention, and millions of cataract operations are performed each year (6), making alternative treatment strategies clearly desirable. Unfortunately, at present we have

no clear understanding of the molecular mechanisms responsible for cataract formation.

Cataracts are associated with a change in the refractive index of the protein-packed lens cells over distances comparable to the wavelength of light, resulting in light scattering. This is caused by a disruption in the order that exists in the densely packed protein phase within these cells, involving aggregation and precipitation, concomitant with the lens losing its transparency (3). In the mammalian eye, the crystallins constitute 90% of the soluble protein in lens fiber cells. High-resolution X-ray structures have been determined for numerous mammalian crystallins by Slingsby and co-workers (3, 4, 7, 8), and an NMR structure was determined for human γ S-crystallin (9). The stability, folding, unfolding, and association states have also been systematically studied, initially by Jaenicke and co-workers (3, 10). Several mutations in the genes for human γ C-crystallin (*CRYGC*) or γ D-crystallin (*CRYGD*) were shown to be associated with cataract formation. In human γ D (H γ D), the R14C mutation causes a progressive juvenile-onset cataract (11, 12) and R58H and R36S mutants (numbering according to Figure 1) of H γ D are believed to cause cataract through their spontaneous crystallization under physiological conditions (13, 14). The P23T mutation was shown to cause loss of solubility and is associated with lamellar cataract in Indian, Caucasian, and Chinese families (15–19).

To fulfill their optical function throughout life, most eye lens proteins have to be first and foremost soluble and stable: their highly concentrated solutions have to last for the entire lifespan of the organism, since cortical lens cells no longer

[†] This work was supported by a Korea Research Foundation Grant funded by the Korean Government (MOEHRD) to J.J. (KRF-2006-352-E00006) and by National Institutes of Health Grant GM 17980 and NEI Grant EY 015834 to J.K.

[‡] RCSB Protein Data Bank accession code rcsb101050, PDB entry 2kfb, and BMRB accession number 16173.

^{*} To whom correspondence should be addressed: Department of Structural Biology, University of Pittsburgh School of Medicine, 1051 Biomedical Science Tower 3, 3501 Fifth Ave., Pittsburgh, PA 15260. Phone: (412) 648-9959. Fax: (412) 648-9008. E-mail: amg100@pitt.edu.

[§] University of Pittsburgh.

^{||} Massachusetts Institute of Technology.

¹ Abbreviations: B γ D, bovine γ D-crystallin; CD, circular dichroism; CPMG, Carr–Purcell–Meiboom–Gill; DTT, dithiothreitol [(2S,3S)-1,4-bis(sulfany)butane-2,3-diol]; EDTA, 2-[2-[bis(carboxymethyl)-amino]ethyl(carboxymethyl)amino]acetic acid; FT-IR, Fourier transform infrared spectroscopy; H γ D, human γ D-crystallin; HMBC, heteronuclear multiple-bond correlation; HSQC, heteronuclear single-quantum correlation; IPAP, in-phase antiphase; IPTG, isopropyl τ -D-1-thiogalactopyranoside; MES, 2-(N-morpholino)ethanesulfonic acid; MWCO, molecular weight cutoff; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; TOCSY, total correlation spectroscopy.

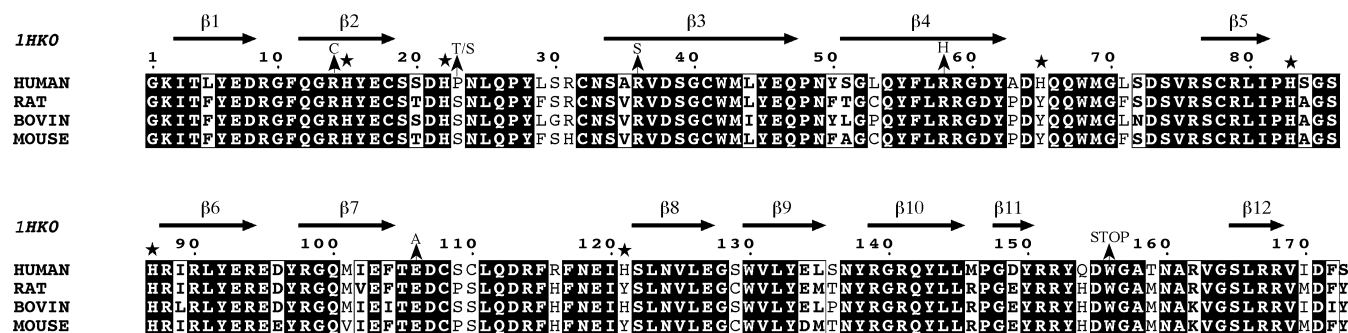


FIGURE 1: Amino acid sequence alignment of four mammalian γ D-crystallins. Identical and similar residues are denoted with black and white boxes, respectively. Secondary structure elements described in the crystal structure of human HgD (1HK0) are indicated at the top. The location and identity of known congenital cataract-causing mutations are marked by arrows and residue names. The positions of all histidines are denoted with stars. This figure was generated with ESPRIPT (63).

contain any organelles and thus are not capable of repair or apoptosis. For a number of HgD mutations, like the nonsense mutations at Trp156 (15), it is generally assumed that the associated truncated proteins are unstable and that the overall protein fold has been destroyed. Trp156 is located before the last β -strand of the C-terminal domain. Therefore, it is possible that the truncation of the polypeptide chain at this point results in an incompletely folded, open C-terminal domain that will be prone to aggregation. E106A (20) is believed to destabilize the protein and cause aggregation. Glu106 is located at the end of β -strand 7 and at the edge of the fourth Greek key motif. In the R14C mutant, introduction of a SH group could render the protein more susceptible to oxidative damage, including disulfide bond formation. A second cysteine cataract-causing mutant, G60C, was recently mapped in a Chinese family with congenital disease (21). In this case, the phenotype was described as “crystal clumps” within the lens. Here again, oxidative protein damage may be involved in causing this phenotype.

Unlike the above-described mutants, codon 24 mutations in *CRYGD* resulting in P23T substitutions are not easily explained (18, 22). Thorough biophysical analyses have been carried out with these mutant human proteins by Slingsby (23) and co-workers and Pande and co-worker (24). The mutant proteins exhibit significantly lower solubility than wild-type HgD in vitro. At 37 °C, they are soluble to <10 mg/mL in aqueous solution, substantially less soluble than what is observed for wild-type HgD (200 mg/mL). In addition, only very small effects on the stability of these mutant proteins were noted (23). In studies employing CD, fluorescence spectroscopy, and FT-IR methodologies, Pande et al. (24) and McManus et al. (25) detected no global structural changes of the P23T mutant compared to the wild type. Interestingly, though, while determining the phase diagram for the mutants, they noted that the solubility of these variants decreased with an increase in temperature, in sharp contrast to the behavior of the wild-type protein, which exhibited the normally observed increase in solubility with temperature; the liquid–liquid phase boundary and the collective diffusion coefficient for the mutants, however, were essentially unchanged (25). Pande et al. (24) concluded that the effect of the mutation was to enhance a self-association reaction into higher-molecular weight clusters and aggregates, without major perturbation of the subunit conformation.

Although high-resolution atomic structures are available for bovine and human HgD (7, 8), efforts to crystallize the P23T cataract mutants were unsuccessful (23, 24).

To shed light on the structural basis of the cataract-causing properties and possible mechanism for the lowered solubility of the P23T HgD mutant, we determined the high-resolution NMR structure of this protein in solution. Comparison of this mutant structure to the X-ray structure of wild-type HgD and its NMR solution parameters, with respect to both structural details and conformational and dynamic properties, reveals a crucial difference in the atomic details surrounding the site of mutation, as well as changes in dynamic behavior. The data suggest that the loss of solubility may depend on the P23T substitution affecting the behavior of the adjacent histidine residue, thereby enabling an aggregation reaction. The latter could result in reduced solubility and formation of high-molecular weight complexes that would be expected to scatter light in the visible region.

EXPERIMENTAL PROCEDURES

Cloning and Protein Purification. The human γ D-crystallin (*CRYGD*) and mutant genes containing N-terminal hexa-His tags were inserted into the pQE.1 vector (Qiagen), and protein was produced in transformed *Escherichia coli* M15 cells, essentially as described previously (26). The equivalent coding sequences without the tag were cloned into pET14b vectors, and Rosetta2 *E. coli* BL21(DE3) cells (EMD Chemicals) were transformed and used for protein production. Cells were grown at 37 °C in LB broth or modified minimal medium. For protein expression, cells were grown at 37 °C and induced with 1 mM IPTG. Isotopic labeling of all proteins was carried out by growing the cultures in modified minimal medium containing 1 g/L $^{15}\text{NH}_4\text{Cl}$ and/or 2 g/L ^{13}C glucose as nitrogen and/or carbon sources, respectively.

Proteins were purified by HisTrap Ni^{2+} affinity chromatography (GE Healthcare), using a linear gradient of imidazole (from 20 to 500 mM) for elution, and further purified by gel filtration on Superdex75 26/60 column (GE Healthcare). Proteins without His tags were purified by ion exchange and gel filtration as follows. Clarified cell lysate was loaded on a HiTrap Q XL anion exchange column (GE Healthcare), equilibrated with 50 mM Tris, 1 mM EDTA, 1 mM DTT buffer (pH 8.2). The flow-through fraction was dialyzed overnight against 10 mM MES buffer (pH 6.0), 1 mM EDTA, 1 mM DTT, 2% glycerol, loaded on a HiTrap SP cation exchange column (GE Healthcare), and fractionated using a linear gradient of 1 M NaCl. Pooled fractions were subjected to gel filtration on a Superdex75 26/60

column (GE Healthcare) in 20 mM sodium phosphate buffer (pH 6.2), 5 mM DTT for the final purification.

The concentration of purified proteins was determined using an A_{280} extinction coefficient of $42.86 \text{ mM}^{-1} \text{ cm}^{-1}$.

NMR Spectroscopy. All NMR spectra for resonance assignments and NOE identification were recorded at 25 °C on either ^{15}N -labeled or ^{13}C - and ^{15}N -labeled samples. The temperature was calibrated with 100% methanol by using chemical shift differences of methyl and hydroxyl protons. Protein concentrations varied between 0.2 and 0.8 mM in 20 mM sodium phosphate buffer, 0.02% NaN_3 , and 5 mM DTT (pH 6.2) with 5% (v/v) D_2O . All spectra were recorded on Bruker AVANCE900, AVANCE800, AVANCE700, and AVANCE600 spectrometers equipped with 5 mm triple-resonance, three-axis gradient probes or a z-axis gradient cryoprobe. Backbone and side chain resonance assignments were carried out using HNCACB, CBCA(CO)NH, HBHA(CO)NH, and HCCH-TOCSY experiments (27, 28). Distance constraints were derived from three-dimensional (3D) simultaneous ^{13}C and ^{15}N NOESY experiments (mixing time of 120 ms) (29). ^1H – ^{15}N HMBC experiments with $1/4J$ values of 11, 16.66, 22, and 29.7 ms were collected for correlating $\text{N}\delta 1$, $\text{N}\epsilon 2$, $\text{H}\delta 2$, and $\text{H}\epsilon 1$ atoms of histidines, respectively (30). Residual HN and $\text{H}\alpha$ dipolar couplings were measured using in-phase/antiphase two-dimensional (2D) ^1H – ^{15}N HSQC and 3D coupled HN(CO)CA experiments, respectively, on protein samples in C12E5/hexanol mixtures (5%, w/v, molar ratio of 0.96) (31) and in isotropic medium using the IPAP method (32). Spectra were processed with NMRPipe (33) and analyzed using SPARKY3 [version 3.113 (34)]. Amide ^1H – ^{15}N combined chemical shift differences were calculated using the equation $\Delta\delta = [(0.125\Delta\delta_{\text{N}})^2 + \Delta\delta_{\text{H}}^2]^{1/2}$.

Structure Calculation. NOE cross-peak assignments and intensity-to-distance calibrations were automatically generated using CYANA 2.1 (35) with the chemical shift input table containing a total of 96.2% resonance assignments of the P23T mutant. The initial CYANA structure calculations were performed by including backbone torsion angle constraints (ϕ and ψ) from TALOS (36) and hydrogen bond constraints from H–D exchange and NOESY data analysis (1.8–2.0 Å for H–O and 2.7–3.0 Å for N–O). The CYANA calculations generated a total of 4018 interproton distance constraints. These structures were refined by simulated annealing with CNS (37) using an explicit water refinement protocol (38). At this stage, 118 HN and 144 $\text{H}\alpha$ residual dipolar coupling (RDC) constraints were added in the calculations. Three hundred structures were calculated, and the 20 lowest-energy structures were selected and analyzed using PROCHECK-NMR (39) and PALES (40). For the final 20 structures, 99.7% of all residues were in the most favored and additionally allowed regions of the Ramachandran plot. All structural figures were generated with MOLMOL (41) or PyMOL (DeLano Scientific). The atomic coordinates and NMR constraints have been deposited in the RCSB Protein Data Bank under accession code rcsb101050 (PDB entry 2kfb and BMRB accession number 16173).

Backbone ^{15}N Relaxation. NMR relaxation measurements were carried out to measure T_1 , T_2 , and heteronuclear $\{^1\text{H}\}$ – ^{15}N NOEs (42). Experiments were carried out at 25

°C on ^{15}N -labeled protein at three pH values (4.0, 6.2, and 7.4) in 20 mM sodium phosphate buffer, 5 mM DTT, and 0.02% NaN_3 at 600 MHz. Relaxation delays in the T_1 experiments were 60, 150, 300, 500, 800, 900, and 1200 ms. Spectra with 60 and 150 ms delays were collected twice for error estimation. The CPMG mixing times in the T_2 experiments were 16.96, 33.92, 50.88, 67.84, 84.80, 84, 101.76, 118.72, 135.68, and 152.64 ms using 1024×128 complex points. The spectra with delays of 16.96 and 33.92 ms were recorded twice for error estimation. Peak intensities were fitted using CURVEFIT (A. G. Palmer, III, Columbia University, New York, NY). The heteronuclear $\{^1\text{H}\}$ – ^{15}N NOE experiment was conducted with an interscan delay of 6 s. A presaturation pulse sequence was applied during the last 3 s of this interscan delay for a NOE spectrum, and no presaturation sequence was applied for a reference spectrum. Data were obtained for 162 of the 169 backbone amides at 600 MHz. Experimental values for the seven remaining residues could not be fitted satisfactorily due to either significant signal overlap or broadening due to chemical exchange.

Solubility Measurements. Unlabeled proteins with and without N-terminal hexa-His tags were purified as described above and eluted from a final size exclusion column in 0.1 M sodium phosphate buffer (pH 7.0), 5 mM DTT, and 0.02% NaN_3 . Wild-type and P23T mutant HgD proteins were concentrated with Centrprep (MWCO 10K, Millipore) and Microcon (MWCO 3.5K, Millipore) concentrators at 25 °C up to the detection of aggregate by visual inspection. The protein concentration in the supernatant at this point was measured spectrophotometrically after removal of the aggregate by centrifugation at 20000g for 30 min. To evaluate and confirm the reported temperature dependency of the solubility under our experimental conditions, the same procedure was performed at 4 °C.

Histidine Tautomeric States and pK_a Values. pH titrations were carried out on 0.2 mM protein samples in 20 mM sodium phosphate buffer, 5 mM DTT, and 0.02% NaN_3 . The pH of the samples was adjusted by adding small aliquots of 0.2 M HCl. Eleven and ten spectra were recorded for wild-type and P23T HgD, respectively. At each pH value, two-dimensional long-range ^1H – ^{15}N HMBC spectra to correlate the $\text{N}\delta 1$ and $\text{N}\epsilon 2$ ring atoms with the $\text{C}\epsilon 1\text{H}$ and $\text{C}\delta 2\text{H}$ ring protons of histidine were recorded with a 16.66 ms dephasing delay, during which the ^1H and ^{15}N signals become antiphase (43). Chemical shift changes at different pH values were fitted to the Henderson–Hasselbalch equation by nonlinear least-squares analysis using KaleidaGraph (version 4.0, Synergy software):

$$\delta_{\text{obs}} = \frac{\delta_{\text{AH}^+} + \delta_{\text{A}} \times 10^{\text{pH}-\text{p}K_a}}{1 + 10^{\text{pH}-\text{p}K_a}}$$

where δ_{obs} is the observed chemical shift of a resonance as a function of pH and δ_{AH^+} and δ_{A} represent the chemical shift values at the low and high extremes of pH, respectively.

RESULTS

Chemical Shift Comparison between Wild-Type and P23T HgD. ^1H – ^{15}N HSQC experiments were performed to assess whether threonine substitution for the Pro23 residue in the N-terminal HgD domain perturbed the structure of the

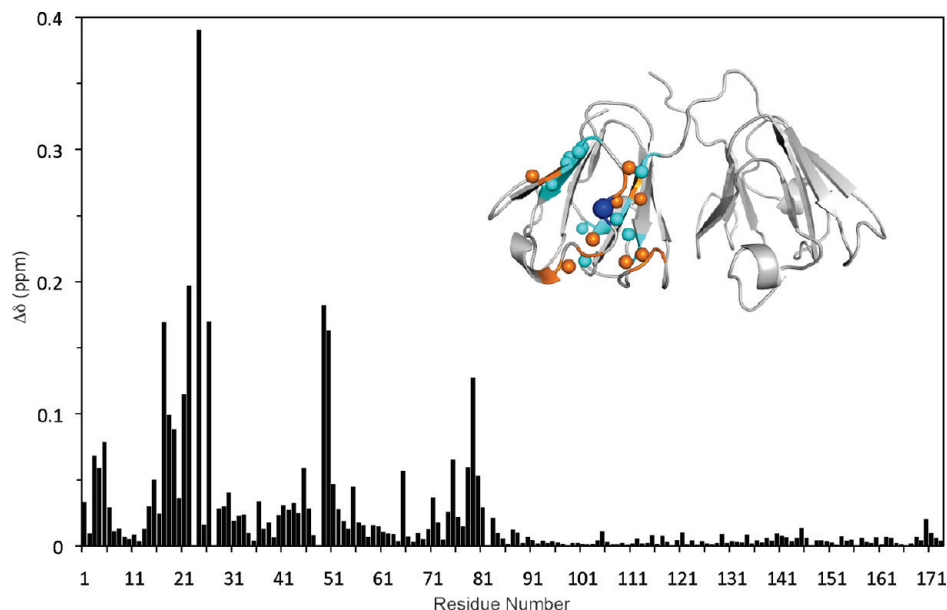


FIGURE 2: Chemical shift differences between wild-type and P23T mutant HgD. Combined amide ^1H – ^{15}N chemical shift differences are plotted vs residue number. Structural mapping of the chemical shift differences onto the X-ray structure of wild-type HgD (1HK0) is provided in the inset. The location of Pro23 is marked with a large blue sphere. Positions of residues whose amide resonances exhibit $\Delta\delta > 0.1$ ppm and $0.1 \text{ ppm} > \Delta\delta > 0.05$ ppm are labeled with small cyan and orange spheres, respectively. The chemical shift data refer to proteins without any His tags.

protein. It is generally accepted that the ^1H – ^{15}N HSQC spectrum can be used as a “fingerprint” of the three-dimensional structure and therefore can be used to evaluate whether a protein undergoes substantial or small structural change(s) upon mutation (44). For wild-type and P23T mutant HgD, the overall appearance of the spectra was similar and all 169/170 possible ^1HN and ^{15}N resonances (174 residues minus 5/4 prolines) were assigned. The fact that complete backbone assignments were obtained indicates that the P23T substitution did not cause global unfolding and/or gross aggregation. Furthermore, the initial $\text{H}\alpha$, $\text{C}\alpha$, and $\text{C}\beta$ assignments on uniformly ^{13}C - and ^{15}N -labeled samples allowed us to readily identify the β -strands, indicating that the mutant HgD exhibited secondary structure elements essentially identical to those observed in the X-ray structure of wild-type HgD. Backbone torsion angle prediction based on chemical shifts using TALOS (36) resulted in good agreement with the angles of the wild-type X-ray structure for the majority of residues.

A comparison of the backbone amide chemical shifts for wild-type and P23T mutant HgD is provided in Figure 2, and a ribbon representation of the wild-type protein X-ray structure onto which the locations of residues that exhibit significant amide ^1H – ^{15}N chemical shift change are mapped (colored orange and cyan) is displayed in the inset. Essentially all shift changes are confined to the N-terminal domain. As expected, the largest differences are observed around the site of mutation, comprising Asp21–Leu25. In addition, the areas around residues 49 and 79 are also strongly affected, with Asn49 and Arg79 exhibiting $\Delta\delta$ values of 0.182 and 0.127 ppm, respectively. Eighteen of the 82 residues of the N-terminal domain exhibited chemical shift differences larger than 0.05 ppm. Essentially no changes are noted in the C-terminal domain (with an average value of 0.006 ppm, and no residue exhibiting any shift of >0.02 ppm). Similar trends were observed for the P23V and P23S mutants. However, the chemical shift changes for these

mutants are smaller than for P23T [$\text{P23T} \geq \text{P23V} \geq \text{P23S}$ (Figure S1A–C of the Supporting Information)].

For comparison, the shift changes that are caused by the presence of the N-terminal His tag are distributed throughout the entire protein, with an average shift value of 0.012 ppm and the largest values on the order of 0.18 ppm. These values are much smaller than chemical shift changes caused by mutation within a protein that can easily range from 0.5 to 0.8 ppm in our experience. A superposition of the ^1H – ^{15}N HSQC spectra of His-tagged and non-His-tagged mutant P23T HgD protein is provided as Supporting Information (Figure S2).

Structure of P23T HgD. The solution structure of the P23T mutant was determined using a final set of 4018 interproton distances, 106 hydrogen bond distances, 131 dihedral angles, and 262 RDC constraints (26 constraints/residue). Note, the final NOE data were collected on the His-tagged P23T HgD. However, the non-His-tagged protein exhibited essentially identical NOEs (see Figure S3 of the Supporting Information), although the low concentration limited the signal-to-noise ratio in these spectra. The structural ensemble is well-defined, satisfies all experimental constraints, displays excellent covalent geometry, and exhibits atomic root-mean-square deviations (rmsd) of 0.35 ± 0.05 and 0.90 ± 0.07 Å with respect to the mean coordinate positions for the backbone atoms (N, $\text{C}\alpha$, and C') and all heavy atoms, respectively. A 20-conformer ensemble is provided in Figure 3A, and all experimental constraints used in the structure calculations as well as pertinent structural statistics are summarized in Table 1. There is a slight variation in domain orientation as evidenced by lower backbone atomic rms deviations of 0.26 ± 0.04 and 0.30 ± 0.03 Å when best-fitting the individual N-terminal (residues 3–81) and C-terminal (residues 88–170) domains, respectively (Figure 3B,C). In addition, excellent agreement between measured and predicted RDC values was noted, with average rms deviations of 1.58 ± 0.14 Hz for

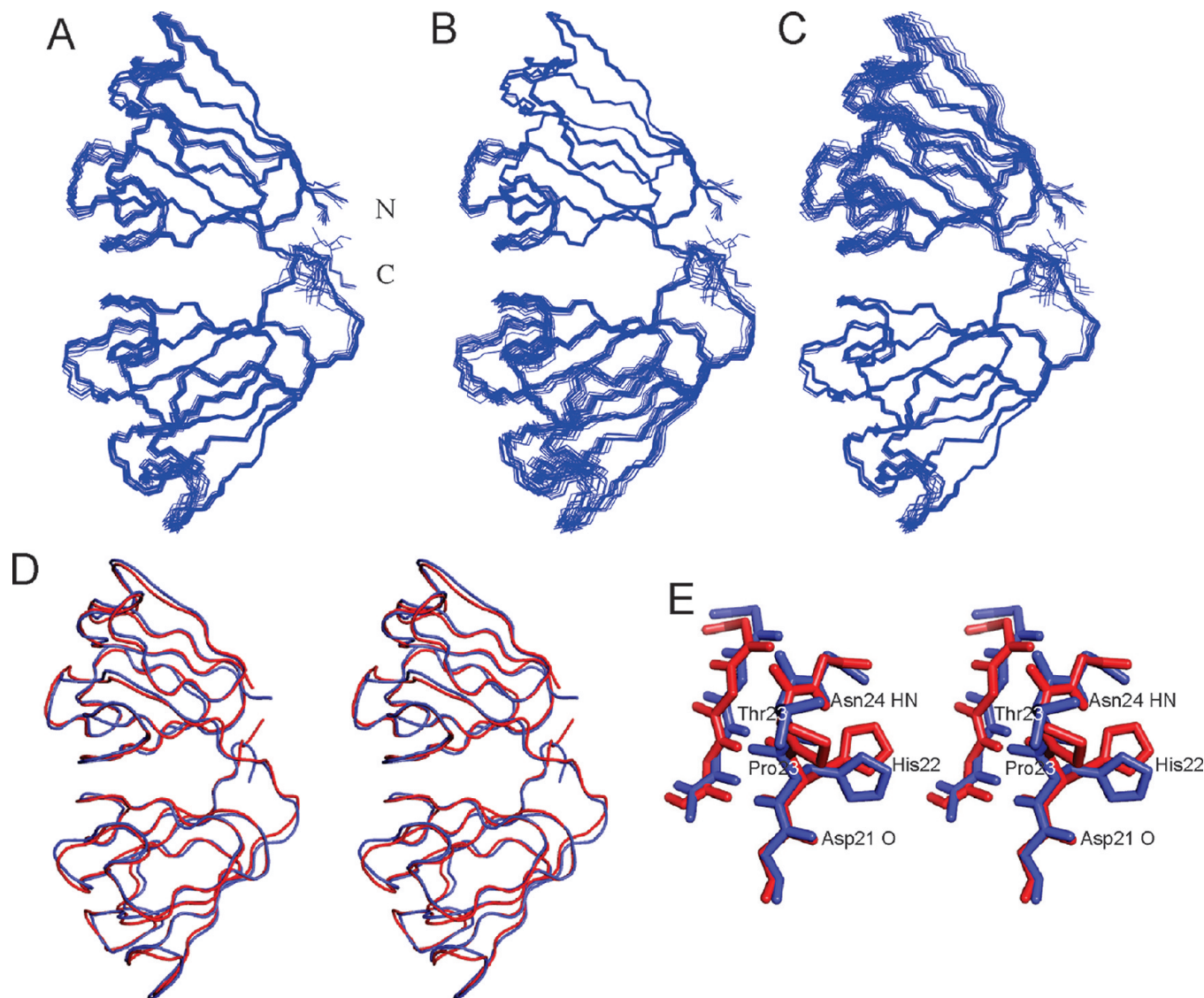


FIGURE 3: Solution structure of the P23T mutant HgD. (A) Backbone superposition of the final 20-conformer ensemble. (B and C) Superposition of the N-terminal (residues 3–81) (B) and C-terminal (residues 88–170) (C) domains only. (D) Backbone superposition of the wild-type X-ray (1HK0) (red) and current P23T HgD (blue, lowest-energy conformer) structures. (E) Stereoview of the superposition of 1HK0 (red) and the current P23T HgD structures (blue) at the mutation site. The P23T mutant HgD structure was calculated using distance and angular constraints measured for the His-tagged protein.

Table 1: Statistics for the Final 20-Conformer Ensemble

no. of distance constraints	
intraresidue ($i - j = 0$)	734
sequential ($i - j = 1$)	992
medium-range ($2 \leq i - j \leq 4$)	471
long-range ($i - j \geq 5$)	1821
total	4018
no. of hydrogen bond constraints	106×2
no. of dihedral angle constraints from TALOS	
ϕ, ψ	67, 64
average rmsd of atomic coordinates (Å)	
backbone heavy atoms	0.35 ± 0.05
all heavy atoms	0.9 ± 0.07
average target function (Å ²)	8.39 ± 0.53
no. of violations	
distances of >0.5 Å	3
dihedral angles of $>5^\circ$	none
Ramachandran plot analysis (%)	
most favorable region	60.0
additional allowed regions	39.7
generously allowed regions	0.3
disallowed regions	0.0

the H–N bonds (correlation coefficient of 0.993) and 2.61 ± 0.12 Hz for the H–C α bonds (correlation coefficient of 0.995).

Figure 3D provides a backbone superposition of the wild-type X-ray (PDB entry 1HK0) and the present solution P23T mutant structures, demonstrating their high degree of similarity. The associated backbone (N, C α , and C') rms differences are 0.96, 0.74, and 0.79 Å for the full-length protein and the N- and C-terminal domains, respectively.

Local Structural Differences in P23T HgD. The pronounced chemical shift perturbation around the mutation site suggested that distinct local structural or dynamic differences were present in the mutant compared to wild-type HgD. Interestingly, the residue immediately preceding the proline in the amino acid sequence is His22, and in the X-ray structure (PDB entry 1HK0), this histidine is involved in an intricate interaction network (Figures 3E and 4A). In particular, the imidazole $\delta 1$ nitrogen accepts a hydrogen bond from the Asn24 backbone HN proton. In contrast, initial

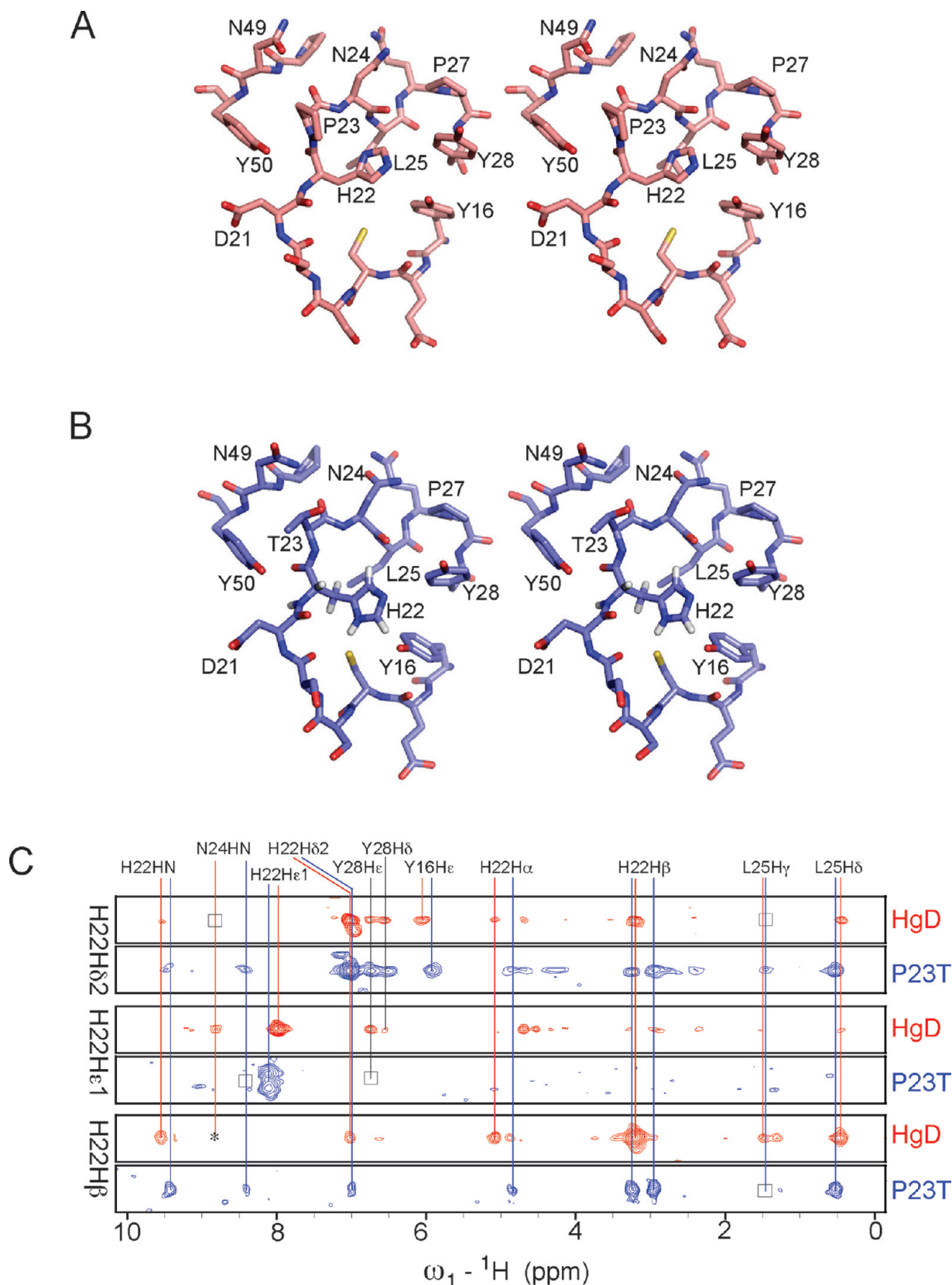


FIGURE 4: Detailed stereoviews of the local structure surrounding His22. (A) The R58H HgD mutant structure (1H4A) is used here for the purposes of comparison only (see the text). (B) Current P23T HgD structure. All protons have been omitted for the sake of clarity except for His22 to convey the predominant tautomer. Oxygen, nitrogen, and sulfur atoms are colored red, blue, and yellow, respectively. (C) Selected strips from the ^{13}C -edited NOESY spectra of both wild-type (red) and P23T HgD (blue), illustrating NOE cross-peaks involving His22. Absent NOEs in either spectrum are indicated by empty boxes, and those below the displayed contour levels are marked by asterisks.

structures for the P23T mutant showed this His22 ring structure is flipped with the His22 $\delta 1$ and $\epsilon 2$ nitrogens distant

from the Asn24 backbone HN proton; therefore, hydrogen bonding was no longer possible. To confirm such structural

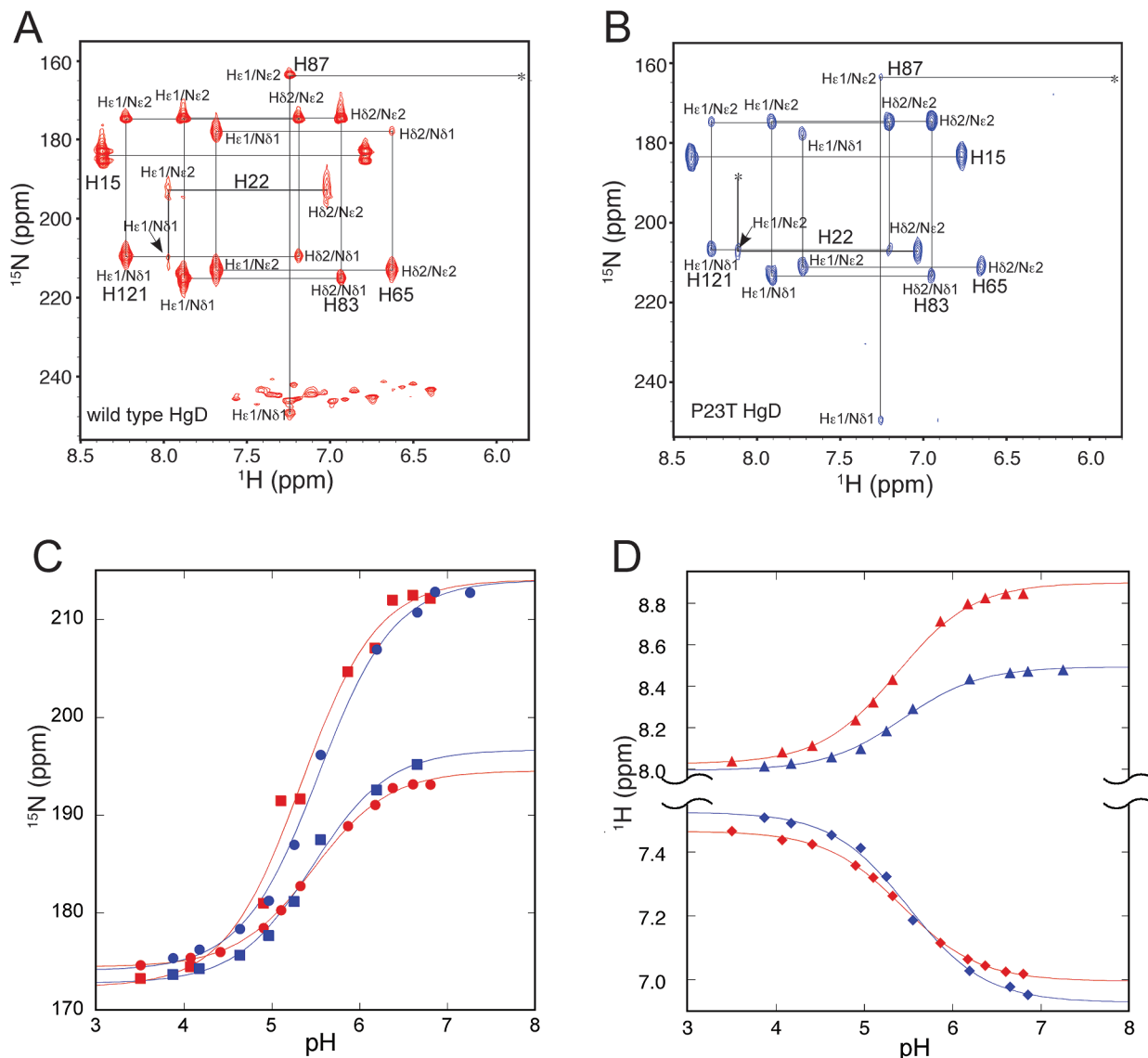


FIGURE 5: Preferred tautomers and pK_a values of histidines. ^1H - ^{15}N HMBC spectra of wild-type (A) and P23T HgD (B). Cross-peaks whose intensities are too small at the displayed contour levels but are clearly visible at lower levels are marked with asterisks, and resonances belonging to the same imidazole ring are connected and labeled by residue name and number and atom type. pH titration curves for N δ and N ϵ atoms (C) and HC δ 2 protons (D) in the His22 imidazole rings in wild-type (red) and P23T (blue) HgD. Resonance frequencies for the N δ 1 and N ϵ 2 atoms and the HC δ 2 protons are shown as squares, dots, and triangles, respectively. In panel D, the titration data for the backbone amide HN protons (diamonds) of Asn24 are displayed to illustrate their correlated behavior. These data were collected for proteins without the N-terminal His tag.

differences, we compared the 3D ^{13}C -edited NOESY spectra of wild-type and P23T mutant HgD in detail (Figure 4). Although very similar NOE patterns are observed, several clear differences were noted, especially in the NOEs involving the His22 protons. In the wild-type protein, an NOE is observed between the Asn24 backbone HN proton and the His22 H ϵ 1 proton, but not the H δ 2 proton, whereas in the P23T mutant protein, the opposite holds true. In addition, no NOE was seen between the His22 H δ 2 proton and the Leu25 H γ proton in the wild-type protein, while it clearly was present in the mutant. Conversely, an NOE was observed between the His22 H β proton and the Leu25 H γ proton in wild-type HgD, but not in the mutant. These NOE differences are reflected in the final wild-type HgD and P23T structures (panels A and B of Figure 4, respectively) that exhibit different local conformations around the His22 imidazole ring.

Histidine Tautomeric States and pK_a Values. Given the structural difference in the His22 side chain conformation, we determined the tautomeric states of its imidazole ring as well as those of all histidine residues in wild-type and P23T HgD using ^1H - ^{15}N HMBC spectroscopy (Figure 5A,B). Histidine residues often serve critical roles in the structure or function of proteins, acting either as nucleophiles or as electrophiles. The intrinsic pK_a of a free imidazole ring is 6.3, and the pK_a for a His imidazole ring in a peptide is 6.54, with its ionization state being modulated by the environment (45–48). Establishing the tautomeric and protonation state of a histidine is impossible from crystallographic data, unless extremely high-resolution (<0.7 Å) maps can be obtained. In contrast, NMR is ideally suited for this purpose (43, 49, 50): unambiguous assignment of the N δ 1 and N ϵ 2 atoms as well as of the H δ 1 and H ϵ 2 protons were obtained from the characteristic cross-peak

Table 2: Histidine pK_a Values

	H15	H22	H65	H83	H87	H121
HgD	6.69	5.40	6.02	6.07	not determined	6.20
P23T	6.51	5.45	5.97	5.90	not determined	6.21

patterns and resonance intensities in ^1H – ^{15}N HMBC spectra. The three-bond $^3J_{\text{N}\delta 1-\text{H}\delta 2}$ coupling constant is always smaller than that of the two-bond $^2J_{\text{NH}}$ coupling in both protonated and neutral histidine, resulting in weak or absent cross-peaks for $^3J_{\text{N}\delta 1-\text{H}\delta 2}$ correlations (51). In a neutral histidine ring, the ^{15}N chemical shift of an unprotonated nitrogen typically resonates at 249.5 ppm, whereas a protonated nitrogen resonates at ~ 167.5 ppm, i.e., 82 ppm further upfield. In a positively charged ring, both protonated and partially charged nitrogens resonate at ~ 176.5 ppm. Intermediate values are indicative of fast exchange between two conformations, and the relative population of each tautomer can be deduced from the differences between the observed and theoretical chemical shift values. A histidine with an equal population of $^{15}\text{N}\epsilon$ and $^{15}\text{N}\delta$ neutral tautomers would exhibit a chemical shift of ~ 210 ppm.

The pK_a values and tautomeric states of all histidines, except His87, for wild-type and P23T HgD are listed in Table 2. Values for His87 could not be determined since its resonances barely moved during the titration between pH 8 and 6 and were too broad for detection below pH 5.86. Comparison of the data for wild-type and P23T HgD (Figure 5A,B and Table 2) reveals no differences in the pK_a values for all histidines, including His22. His22 exhibits the most acidic pK_a value (5.40 and 5.45 for wild-type and P23T HgD, respectively) compared to the other histidines which exhibit pK_a values between 5.9 and 6.7. The lowered pK_a value of His22 is caused by H-bonding that makes protonation more difficult. In addition, as is readily apparent from the different patterns in the ^1H – ^{15}N HMBC spectrum (Figure 5A,B), His22 exhibits an alternative tautomeric state in its neutral form. While all other histidines exist predominantly as the $\text{N}\epsilon 2$ tautomer, for His22 the $\text{N}\epsilon 2$ tautomer is the predominant neutral form in the wild-type protein, while the $\text{N}\delta 1$ tautomer is the predominant form in the P23T mutant. Alternative tautomeric states for His22 are also observed in the P23S and P23V mutants (Figure S4 of the Supporting Information). Furthermore, the ^{15}N chemical shifts at the high-pH end of the titration curves (neutral form) are 213.2 ppm for the $\text{N}\delta 1$ atom and 194.5 ppm for the $\text{N}\epsilon 2$ atom for the wild type and 196.7 and 213.2 ppm for P23T HgD, respectively. These chemical shifts are clearly different from the average values observed for a neutral histidine ring (~ 249.5 ppm for the unprotonated ^{15}N and ~ 167.5 ppm for the protonated ^{15}N). Therefore, both tautomers are present in the neutral state, with 5:3 and 3:5 ratios for wild-type and P23T HgD, respectively. The equilibrium between the two tautomers causes severe line broadening of the His22 imidazole ring nitrogen resonances in the wild-type protein, with less severe line broadening for the P23T HgD and the P23S and P23V mutants. Such line broadening indicates that the exchange time between the two tautomers approaches 48 μs at the crossover point ($\tau = 2/\Delta\omega$, where $\Delta\omega$ is the difference in chemical shift frequencies in units of radians per second), based on the chemical shift difference of 82 ppm

between the $\text{N}\delta 1$ and $\text{N}\epsilon 2$ atoms (equivalent to $\Delta\omega/2\pi = 6651$ Hz at 18.8 T) in wild-type HgD. The mutant proteins exhibit less severe line broadening, indicating faster exchange between the two tautomers.

The notable upfield shift of the imidazole ring $\delta 2$ proton resonance for His87 (~ 5.6 ppm) in both proteins is most likely caused by the proximity (ring current) of Trp130.

Dynamics of Wild-Type and P23T HgD. To assess the effect of the P23T mutation on the motional properties of the protein, R_1 , R_2 , and steady-state heteronuclear $\{^1\text{H}\}$ – ^{15}N NOEs were measured. Such data allow us to evaluate whether the observed chemical shift changes are associated with local unfolding. All experimental values are plotted versus the residue number in Figure 6. For wild-type HgD at pH 6.2, the average R_2 value was 15.73 ± 1.61 s $^{-1}$. Interestingly, several residues around Pro23 exhibited elevated values. For example, Asn24 showed the largest value of 19.39 s $^{-1}$, and Ser21 and His22 also displayed increased R_2 values of 17.52 and 18.26 s $^{-1}$, respectively (Figure 6). This feature, however, was not seen in the data of the P23T mutant, with Ser21, His22, and Asn24 exhibiting R_2 values of 15.50, 15.85, and 16.08 s $^{-1}$, respectively, well within the overall error of the average value of 15.14 ± 1.58 s $^{-1}$. Slightly elevated values were also observed for Gln67 and Val75 in wild-type HgD (18.15 and 18.53 s $^{-1}$, respectively) with corresponding values of 16.96 and 16.86 s $^{-1}$, respectively, in the P23T mutant. Interestingly, smaller than average values (10.73 and 10.00 s $^{-1}$, respectively) were observed for Phe117, both in the wild-type protein and in the mutant.

To investigate whether the exchange contribution seen in the R_2 data for wild-type HgD is associated with the tautomeric or protonation state of His22, we collected additional R_2 data for low-pH (pH 4.0) and high-pH (pH 7.4) samples. At pH 7.4, the exchange contribution to T_2 relaxation was increased, whereas at pH 4.0, essentially no exchange contribution was seen (Figure 7D). For example, the R_2 values of Ser21 and His22 are 30.03 and 24.30 s $^{-1}$, respectively, at pH 7.4, compared to the average value of 15.58 ± 2.30 s $^{-1}$. No relaxation data could be measured for Asn24, due to severe line broadening of the amide resonance. In contrast, at pH 4.0, no increased R_2 values are noted, and overall, a small decrease in R_2 values was observed for most residues. The relaxation data presented here, obtained at different pH values, in particular the noted increase in R_2 values for the residues around His22, can be explained by the conformational equilibrium involving the His22 ring. At high pH, such as 7.4, the neutral $\text{N}\epsilon 2$ tautomer is the major form of His22 with the $\text{N}\delta 1$ atom unprotonated and accepting a H-bond from the backbone amide proton of Asn24 in the wild-type protein. This tautomer undergoes exchange with the minor $\text{N}\delta 1$ tautomer where no H-bond is possible. At pH values below the pK_a , such as pH 4.0, His22 is fully positively charged and H-bonding is impossible. Therefore, no exchange between H-bonded and non-H bonded conformers occurs. Similarly, for the P23T mutant, the $\text{N}\delta 1$ tautomer is the major neutral H22 species for which H-bonding to the Asn24 backbone HN proton is impossible (Figure 4B). As a result, the R_2 values are less affected, even at higher pH [pH 6.2 (Figure 6A)].

Protein Solubility. Our structural studies were initiated with proteins that contained N-terminal His tags, and surprisingly, we found that the His-tagged P23T mutant protein could be

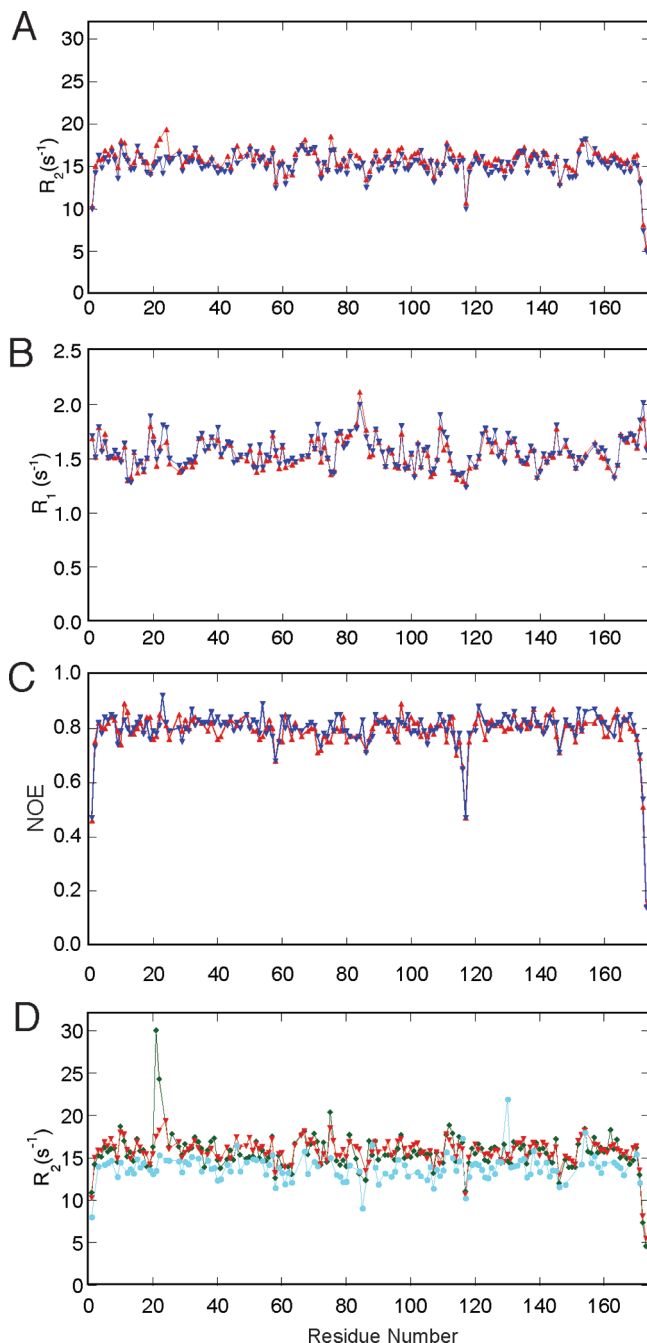


FIGURE 6: Backbone ^{15}N heteronuclear relaxation data for wild-type (red) and P23T (blue) HgD: (A) R_2 transverse rates, (B) R_1 longitudinal relaxation rates, and (C) ^1H – ^{15}N heteronuclear NOE values. In panel D, R_2 values for wild-type HgD at different pH values are displayed: pH 7.4 (green), pH 6.2 (red), and pH 4.0 (light blue). All data were collected at 600 MHz and 25 °C.

concentrated up to 90 mg/mL (no further concentration was attempted). As expected, wild-type, His-tagged HgD was also very soluble. A concentration of 90 mg/mL was easily reached in 0.1 M sodium phosphate buffer (pH 7) at room temperature (no further concentration was attempted). Since this behavior was clearly at odds with previous reports (23, 24), we carried out solubility studies for the non-His-tagged proteins. In this case, we were able to concentrate wild-type HgD up to 300 mg/mL, while the P23T mutant protein already exhibited visible aggregates when concentrated to 4 mg/mL. Decreasing the temperature from 25 to 4 °C allowed the P23T mutant to be concentrated somewhat further without

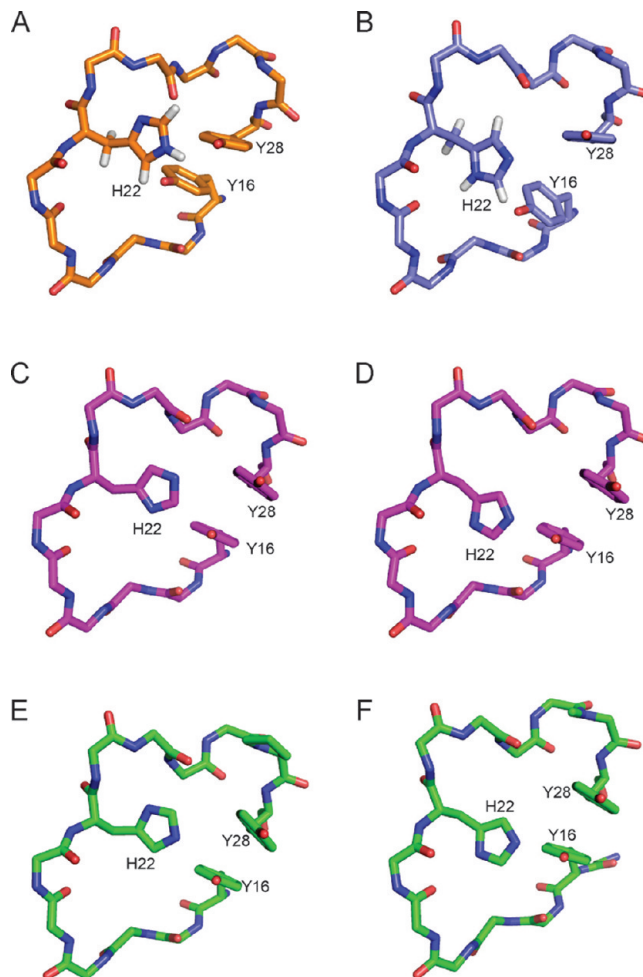


FIGURE 7: Histidine conformations in several γD -crystallin structures. NMR solution structure of the P23T mutant (A and B). X-ray structures of HgD R36S (PDB entry 2G98) (C and D) and bovine γD -crystallin (PDB entry 1ELP) (E and F). In panels A, C, and E the $\epsilon 2$ -tautomers and in panels B, D, and F the $\delta 1$ -tautomers are depicted. In the X-ray structures (C–F), different molecules in an asymmetric unit are shown.

aggregation (up to ~ 10 mg/mL), in agreement with previous reports. Therefore, adding a stretch of polar residues at the N-terminus significantly changed the aggregation properties of these proteins.

DISCUSSION

Structure. Previous studies using CD and FT-IR (23, 24) revealed that the overall structure of the P23T mutant is very similar to that of wild-type HgD. Similarly, characterization of the aggregates did not suggest a major change in the subunit structure (24). This work was aimed at investigating the structural properties of the cataract-causing codon 24 mutant of HgD at the atomic level, and we uncovered distinctive and possibly pivotal differences in a local area around the substitution.

Interestingly, chemical shift changes between mutant and wild-type HgDs (Figure 2) indicate that the two domains are essentially independent of each other and that structural perturbations in one domain are not communicated to the other. This is an unusual finding, given the domain interface and the fact that the linker between the domains is not flexible.

A best-fit superposition of our mutant structure with the wild-type X-ray structure (PDB entry 1HK0) is provided in Figure 3D and demonstrates their similarity, with rmsd values of 0.96 and 1.51 Å for backbone and side chain atoms, respectively. A comparison with the R58H mutant structure (PDB entry 1H4A) also yielded comparable overall rmsd values of 0.90 and 1.41 Å for the backbone and side chain atoms, respectively. Below, we used the 1H4A structure for comparing the local regions around His22 instead of the 1HK0 structure since in the latter structure the imidazole ring exhibited a poor fit to the electron density. It also did not match our NOE data for wild-type HgD in the area around His22. In contrast, the histidine side chain in 1H4A exhibited a much better fit to the density and the χ^2 angle of the imidazole ring agreed with our NOE data for the wild-type protein, making this X-ray model the better choice for comparing details around the mutation site.

As illustrated in Figure 3E, notable differences are observed between wild-type and mutant HgD: the C α atoms of His22 and Thr23 in P23T HgD are displaced by 0.6 and 0.5 Å from the positions in the wild-type protein, respectively, and substantial changes in the backbone dihedral angles are present. Stereoviews of the two structures around His22 are provided in Figure 4. Such conformational difference is not caused by the different methodologies that were used for structure determination, NMR for the P23T mutant structure and X-ray for the wild-type and R58H HgD structures. Our NOESY data recorded for the wild-type protein (see above) clearly demonstrate that the local structure around His22 for wild-type HgD in solution is close to that observed in the R58H HgD X-ray structure (PDB entry 1H4A). Indeed, the measured pattern of NOEs is clearly different between the wild type and mutant, and the wild-type pattern is compatible with only the 1H4A X-ray model.

Histidine 22 Tautomeric States. Comparison of the ^1H – ^{15}N HMBC spectra allowed us to accurately determine the tautomeric states and pK_a values of all histidines in wild-type and P23T HgD. A significant difference was found with respect to the preferred neutral tautomer of the His22 ring: in wild-type HgD, the H ϵ 2 tautomer is the dominant one, whereas in P23T HgD, the H δ 1 tautomer is preferred. Another interesting observation is related to the final nitrogen shifts for the unprotonated nitrogen in the neutral imidazole ring of His22. Typically, such nitrogens resonate at ~ 249.5 ppm, while the hydrogen-bearing nitrogen resonates at 167.5 ppm. We observe shifts of ~ 213 and ~ 195 ppm for the unprotonated and protonated nitrogens, respectively, for both wild-type and P23T proteins, somewhat intermediate values. This suggests a dynamic equilibrium between the two tautomeric forms, and we believe that this equilibrium involves a ring flip of the histidine ring; i.e., in one conformation, the N δ 1 atom is H-bonded with the Asn24 amide proton, and in the other conformation, the nitrogen is incapable of having such a H-bond. The fact that this dynamic equilibrium for the His22 side chain is present and that a different preferred tautomer is found in P23T HgD is also reflected in the consistently larger chemical shift changes throughout the N-terminal domain. The change in local charge is transmitted throughout the entire domain, indicating that charge effects can be transmitted over very long distances within proteins.

Since it is clear from the ^1H – ^{15}N HMBC data that both His22 tautomers are present in significant amounts, we also calculated the P23T structure with the ϵ 2-tautomer for His22 (Figure 7A). Interestingly, two conformations for the His22 side chain are also observed in several X-ray structures. The structure of the R36S mutant of HgD (PDB entry 2G98) was determined from a crystal that was removed from the lens of a 5-year-old patient (13). In this structure, two different His22 conformations are observed for the two independent molecules in the asymmetric unit (Figure 7C,D). One conformation is very similar to the ϵ 2-tautomeric structure of P23T (Figure 7A), while the other resembles the δ 1-tautomer (Figure 7B). However, the assignment of the His22 imidazole ring nitrogens in the R36S mutant is different from that in P23T, most likely prompted by the limited resolution of this crystal structure, and not correct. Similarly, two conformations for His22 are also observed in the bovine γ D-crystallin (BgD) structure (PDB entry 1ELP) (7) (Figure 7E,F). Therefore, structural heterogeneity around the His22 ring may be a common feature in γ D-crystallin family members.

Effect of a His Tag. Another interesting observation that emerged from this work was the surprising, high solubility of His-tagged P23T HgD. In the X-ray structure of wild-type HgD (1HK0) and the R58H mutant (1H4A), dense packing is observed between the molecules in the lattice (solvent content of $\sim 40\%$) and the N-termini are close to neighboring molecules (4.4 or 8.6 Å). Therefore, the addition of additional charged residues could alter these interactions. Naturally, this will greatly affect the surface features of the protein, possibly affecting properties related to self-assembly. Indeed, the phase diagram of the related P23V HgD mutant revealed that for this protein, the solid-phase solubility is switched from normal (wild-type) to retrograde temperature dependence, without changing the liquid–liquid cloud points (25). Addition of several histidine residues at the N-terminus, in turn, may reverse this effect. Indeed, experimental determination of the isoelectric point of the proteins without and with the His tag revealed an increase by ~ 1 pH unit caused by the tag, changing from 7.4 to 8.3.

Motional Behavior. The conformational equilibrium detected in the His22 resonances is also observed in the relaxation parameters. It is most clearly observed in the R_2 values since they are sensitive to slower (micro- to millisecond time scale) motions. As one can note from the data provided in Figure 6, a significant difference is observed for wild-type and P23T HgD for the area encompassing residues 20–25. The values for the wild-type protein clearly exhibit the presence of an exchange contribution, while this appears to be absent from the mutant data. From our data on the His residues, we suspected that the protonation state of the imidazole ring may be involved in this exchange, and the T_2 data at several pH values support this notion. A clear increase in the exchange contribution was observed for the higher pH values, linking this backbone motion to the histidine ring flip and an accompanying H-bonding.

The increase in the level of local motion for the wild-type protein compared to the P23T HgD mutant seemed a surprising result. One may naively assume that a proline to threonine mutation would make the backbone more flexible, and one may observe the conformational equilibrium in the mutant. However, R_2 values are particularly sensitive to the

slower motions, and it may well be that a histidine ring flip is easily accomplished with an overall fast backbone motion while the presence of a proline slows the entire conformational motion, making it observable in the T_2 data.

Model for P23T HgD Aggregation. Currently, no structural information at the atomic level is available regarding the quaternary state of HgD in a cataract or the aggregated state. Inspecting crystal contacts may provide some information in that regard. In crystal lattice of 1HK0 and 1H4A, His22 is not involved any protein–protein contacts. However, Tyr28, a nearby residue, engages in contacts. This type of interaction is also observed in the 1ELP structure. In both cases, Tyr28 interacts with Ile102 of the C-terminal domain, in a “head-to-tail” fashion. In the case of 2G98, His22 makes direct contacts with Arg116 of the neighboring molecule. Therefore, His22 can be engaged in protein–protein interaction. Since P23T exhibits lower solubility than wild-type HgD, the δ 1-tautomer may be more prone to self-association. In the wild-type protein, the ϵ 2-tautomer is dominant and exchange between the conformations is slow; thus, aggregation may be thermodynamically and kinetically unfavorable in this case.

The best developed model in which a single amino acid substitution leads to a polymerized state, without a major change in the soluble species, pertains to sickle cell anemia. In sickle cell disease, a single point mutation at the b6 position of hemoglobin (Hb) (b6 Glu $\rightarrow$ Val) creates a hydrophobic patch on the surface of the molecule, which, under deoxygenating conditions, leads to the formation of long hemoglobin fibers within the red blood cell (52, 53). The formation of Hb S fibers is possible because the hydrophobic surface-exposed area around the valine can interact with the hydrophobic pocket formed at the EF corner of the b chain, when the hemoglobin is in the deoxygenated state (54).

Hb S polymerization proceeds through a nucleation and growth mechanism. The first step is nucleation, a thermodynamically unfavorable bimolecular process in which individual molecules associate, with one Hb S molecule acting as a donor and one as an acceptor. Eventually, a nucleus of critical size n is formed (55, 56). Pande et al. reported that P23T HgD aggregation also proceeds through a nucleation and growth mechanism (24).

At present, we cannot predict the site on HgD that is recognized by the site around P23T. This will require NMR structures of dimeric or higher-order intermediates in the aggregation pathway. We have previously been successful in determining structures of intermediates in the self-association of GB1 (57–59), and such methodology should be applicable to the crystallins as well.

Although our current data hold no indication of such reactions, we cannot rule out domain swapping (60, 61) or loop sheet insertion as a mechanism for polymerization. Interestingly, the polymerized form of a serpin, antithrombin, involved major conformational rearrangements, without any indication of these changes evident in the structure of the mutant precursor structures (62).

As for the question of why a γ D-crystallin with threonine or serine instead of proline at codon 24 works well in other organisms, like cow or mouse, one needs to consider all amino acid changes. Overall, there are 26 differences in sequence between BgD and HgD, some of which may be

more important than others. For example, Leu5 is substituted for Phe5 in BgD, and this larger, hydrophobic side chain is packed closely against His22, compensating for the change in local conformation induced by P23T. Therefore, in the context of the bovine sequence, no weakening of the H-bond would occur.

SUPPORTING INFORMATION AVAILABLE

Figures S1–S4 present NMR spectra that illustrate differences among wild-type, P23T, P23V, and P23S HgD and similarities between His-tagged and non-His-tagged mutant P23T HgD. This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES

- Wistow, G. J., and Piatigorsky, J. (1988) Lens crystallins: The evolution and expression of proteins for a highly specialized tissue. *Annu. Rev. Biochem.* 57, 479–504.
- Horwitz, J. (2003) α -Crystallin. *Exp. Eye Res.* 76, 145–153.
- Bloemendal, H., de Jong, W., Jaenicke, R., Lubsen, N. H., Slingsby, C., and Tardieu, A. (2004) Ageing and vision: Structure, stability and function of lens crystallins. *Prog. Biophys. Mol. Biol.* 86, 407–485.
- Jaenicke, R., and Slingsby, C. (2001) Lens crystallins and their microbial homologs: Structure, stability, and function. *Crit. Rev. Biochem. Mol. Biol.* 36, 435–499.
- Foster, A. (2001) Cataract and “Vision 2020—the right to sight” initiative. *Br. J. Ophthalmol.* 85, 635–637.
- Toh, T., Morton, J., Coxon, J., and Elder, M. J. (2007) Medical treatment of cataract. *Clin. Exp. Ophthalmol.* 35, 664–671.
- Chirgadze, Y. N., Driessen, H. P., Wright, G., Slingsby, C., Hay, R. E., and Lindley, P. F. (1996) Structure of bovine eye lens γ D (γ IIIb)-crystallin at 1.95 Å. *Acta Crystallogr. D* 52, 712–721.
- Basak, A. K., Bateman, O., Slingsby, C., Pande, A., Asherie, N., Ogun, O., Benedek, G. B., and Pande, J. (2003) High-resolution X-ray crystal structures of human γ D crystallin (1.25 Å) and the R58H mutant (1.15 Å) associated with aculeiform cataract. *J. Mol. Biol.* 328, 1137–1147.
- Wu, Z., Delaglio, F., Wyatt, K., Wistow, G., and Bax, A. (2005) Solution structure of γ S-crystallin by molecular fragment replacement NMR. *Protein Sci.* 14, 3101–3114.
- Rudolph, R., Siebendritt, R., Nessler, G., Sharma, A. K., and Jaenicke, R. (1990) Folding of an all- β protein: Independent domain folding in γ II-crystallin from calf eye lens. *Proc. Natl. Acad. Sci. U.S.A.* 87, 4625–4629.
- Stephan, D. A., Gillanders, E., Vanderveen, D., Freas-Lutz, D., Wistow, G., Baxevas, A. D., Robbins, C. M., VanAuken, A., Quesenberry, M. I., Bailey-Wilson, J., Juo, S. H., Trent, J. M., Smith, L., and Brownstein, M. J. (1999) Progressive juvenile-onset punctate cataracts caused by mutation of the γ D-crystallin gene. *Proc. Natl. Acad. Sci. U.S.A.* 96, 1008–1012.
- Pande, A., Pande, J., Asherie, N., Lomakin, A., Ogun, O., King, J. A., Lubsen, N. H., Walton, D., and Benedek, G. B. (2000) Molecular basis of a progressive juvenile-onset hereditary cataract. *Proc. Natl. Acad. Sci. U.S.A.* 97, 1993–1998.
- Kmoch, S., Brynda, J., Asfaw, B., Bezouska, K., Novák, P., Rezáčová, P., Ondrová, L., Filipce, M., Sedláček, J., and Elleder, M. (2000) Link between a novel human γ D-crystallin allele and a unique cataract phenotype explained by protein crystallography. *Hum. Mol. Genet.* 9, 1779–1786.
- Pande, A., Pande, J., Asherie, N., Lomakin, A., Ogun, O. O., King, J. A., Lubsen, N. H., Walton, D., and Benedek, G. B. (2000) Molecular basis of a progressive juvenile-onset hereditary cataract. *Proc. Natl. Acad. Sci. U.S.A.* 97, 1993–1998.
- Santhiya, S. T., Shyam Manohar, M., Rawley, D., Vijayalakshmi, P., Namperumalsamy, P., Gopinath, P. M., Löster, J., and Graw, J. (2002) Novel mutations in the γ -crystallin genes cause autosomal dominant congenital cataracts. *J. Med. Genet.* 39, 352–358.
- Nandrot, E., Slingsby, C., Basak, A. K., Cherif-Chefchaoui, M., Benazzou, B., Hajaji, Y., Boutayeb, S., Gribouval, O., Arbogast, L., Berraho, A., Abitbol, M., and Hilal, L. (2003) γ -D crystallin gene (CRYGD) mutation causes autosomal dominant congenital cerulean cataracts. *J. Med. Genet.* 40, 262–267.

17. Burdon, K. P., Wirth, M. G., Mackey, D. A., Russell-Eggitt, I. M., Craig, J. E., Elder, J. E., Dickinson, J. L., and Sale, M. M. (2004) Investigation of crystallin genes in familial cataract, and report of two disease associated mutations. *Br. J. Ophthalmol.* 88, 79–83.
18. Mackay, D. S., Andley, U. P., and Shiels, A. (2004) A missense mutation in the γ D-crystallin gene (CRYGD) associated with autosomal dominant “coral-like” cataract linked to chromosome 2q. *Mol. Vision* 10, 155–162.
19. Shentu, X., Yao, K., Xu, W., Zheng, S., Hu, S., and Gong, X. (2004) Special fasciculiform cataract caused by a mutation in the γ D-crystallin gene. *Mol. Vision* 10, 233–239.
20. Messina-Baas, O. M., Gonzalez-Huerta, L. M., and Cuevas-Covarrubias, S. A. (2006) Two affected siblings with nuclear cataract associated with a novel missense mutation in the CRYGD gene. *Mol. Vision* 12, 995–1000.
21. Li, F., Wang, S., Gao, C., Liu, S., Zhao, B., Zhang, M., Huang, S., Zhu, S., and Ma, X. (2008) Mutation G61C in the CRYGD gene causing autosomal dominant congenital coralliform cataracts. *Mol. Vision* 14, 378–386.
22. Plotnikova, O. V., Kondrashov, F. A., Vlasov, P. K., Grigorenko, A. P., Ginter, E. K., and Rogaev, E. I. (2007) Conversion and compensatory evolution of the γ -crystallin genes and identification of a cataractogenic mutation that reverses the sequence of the human CRYGD gene to an ancestral state. *Am. J. Hum. Genet.* 81, 32–43.
23. Evans, P., Wyatt, K., Wistow, G. J., Bateman, O. A., Wallace, B. A., and Slingsby, C. (2004) The P23T cataract mutation causes loss of solubility of folded γ D-crystallin. *J. Mol. Biol.* 343, 435–444.
24. Pande, A., Annunziata, O., Asherie, N., Ogun, O., Benedek, G. B., and Pande, J. (2005) Decrease in protein solubility and cataract formation caused by the Pro23 to Thr mutation in human γ D-crystallin. *Biochemistry* 44, 2491–2500.
25. McManus, J. J., Lomakin, A., Ogun, O., Pande, A., Basan, M., Pande, J., and Benedek, G. B. (2007) Altered phase diagram due to a single point mutation in human γ D-crystallin. *Proc. Natl. Acad. Sci. U.S.A.* 104, 16856–16861.
26. Kosinski-Collins, M. S., Flaugh, S. L., and King, J. (2004) Probing folding and fluorescence quenching in human γ D crystallin Greek key domains using triple tryptophan mutant proteins. *Protein Sci.* 13, 2223–2235.
27. Clore, G. M., and Gronenborn, A. M. (1998) Determining the structures of large proteins and protein complexes by NMR. *Trends Biotechnol.* 16, 22–34.
28. Bax, A., and Grzesiek, S. (1993) Methodological advances in protein NMR. *Acc. Chem. Res.* 26, 131–138.
29. Sattler, M., Maurer, M., Schleucher, J., and Griesinger, C. (1995) A Simultaneous ^{15}N , ^1H -HSQC and ^{13}C , ^1H -HSQC with sensitivity enhancement and a heteronuclear gradient-echo. *J. Biomol. NMR* 5, 97–102.
30. Day, R. M., Thalhauser, C. J., Sudmeier, J. L., Vincent, M. P., Torchilin, E. V., Sanford, D. G., Bachovchin, C. W., and Bachovchin, W. W. (2003) Tautomerism, acid-base equilibria, and H-bonding of the six histidines in subtilisin BPN' by NMR. *Protein Sci.* 12, 794–810.
31. Ruckert, M., and Otting, G. (2000) Alignment of biological macromolecules in novel nonionic liquid crystalline media for NMR experiments. *J. Am. Chem. Soc.* 122, 7793–7797.
32. Otting, M., Delaglio, F., and Bax, A. (1998) Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J. Magn. Reson.* 131, 373–378.
33. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* 6, 277–293.
34. Goddard, T. D., and Kneller, D. G. (2004) SPARKY 3, version 3.110, University of California, San Francisco.
35. Herrmann, T., Güntert, P., and Wüthrich, K. (2002) Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. *J. Mol. Biol.* 319, 209–227.
36. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR* 13, 289–302.
37. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D54* (Part 5), 905–921.
38. Linge, J. P., Williams, M. A., Spronk, C. A., Bonvin, A. M., and Nilges, M. (2003) Refinement of protein structures in explicit solvent. *Proteins* 50, 496–506.
39. Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR* 8, 477–486.
40. Zweckstetter, M., and Bax, A. (2000) Prediction of sterically induced alignment in a dilute liquid crystalline phase: Aid to protein structure determination by NMR. *J. Am. Chem. Soc.* 122, 3791–3792.
41. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graphics* 14, 51–55, 29–32.
42. Farrow, N. A., Muhandiram, R., Singer, A. U., Pascal, S. M., Kay, C. M., Gish, G., Shoelson, S. E., Pawson, T., Forman-Kay, J. D., and Kay, L. E. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ^{15}N NMR relaxation. *Biochemistry* 33, 5984–6003.
43. Pelton, J. G., Torchia, D. A., Meadow, N. D., and Roseman, S. (1993) Tautomeric states of the active-site histidines of phosphorylated and unphosphorylated IIIGlc, a signal-transducing protein from *Escherichia coli*, using two-dimensional heteronuclear NMR techniques. *Protein Sci.* 2, 543–558.
44. Forman-Kay, J. D., Clore, G. M., Stahl, S. J., and Gronenborn, A. M. (1992) ^1H and ^{15}N resonance assignments and secondary structure of the human thioredoxin C62A, C69A, C73A mutant. *J. Biomol. NMR* 2, 431–445.
45. Nozaki, Y., and Tanford, C. (1967) Examination of titration behavior. *Methods Enzymol.* 11, 715–734.
46. Bachovchin, W. W. (1986) ^{15}N NMR spectroscopy of hydrogen-bonding interactions in the active site of serine proteases: Evidence for a moving histidine mechanism. *Biochemistry* 25, 7751–7759.
47. Song, J., Laskowski, M., Jr., Qasim, M. A., and Markley, J. L. (2003) NMR determination of pKa values for Asp, Glu, His, and Lys mutants at each variable contiguous enzyme-inhibitor contact position of the turkey ovomucoid third domain. *Biochemistry* 42, 2847–2856.
48. Bachovchin, W. W. (2001) Contributions of NMR spectroscopy to the study of hydrogen bonds in serine protease active sites. *Magn. Reson. Chem.* 39, S199–S213.
49. Van Dijk, A. A., Scheek, R. M., Dijkstra, K., Wolters, G. K., and Robillard, G. T. (1992) Characterization of the protonation and hydrogen bonding state of the histidine residues in IIAMtl, a domain of the phosphoenolpyruvate-dependent mannitol-specific transport protein. *Biochemistry* 31, 9063–9072.
50. Bhattacharya, S., Sukits, S. F., MacLaughlin, K. L., and Lecomte, J. T. (1997) The tautomeric state of histidines in myoglobin. *Biophys. J.* 73, 3230–3240.
51. Blomberg, F., Maurer, W., and Ruterjans, H. (1977) Nuclear magnetic resonance investigation of ^{15}N -labeled histidine in aqueous solution. *J. Am. Chem. Soc.* 99, 8149–8159.
52. Eaton, W. A., and Hofrichter, J. (1987) Hemoglobin S gelation and sickle cell disease. *Blood* 70, 1245–1266.
53. Eaton, W. A., and Hofrichter, J. (1990) Sick cell hemoglobin polymerization. *Adv. Protein Chem.* 40, 63–279.
54. Ferrone, F. A. (1993) The polymerization of sickle hemoglobin in solutions and cells. *Experientia* 15, 110–117.
55. Hofrichter, J., Ross, P. D., and Eaton, W. A. (1974) Kinetics and mechanism of deoxyhemoglobin S gelation: A new approach to understanding sickle cell disease. *Proc. Natl. Acad. Sci. U.S.A.* 71, 4864–4868.
56. Ferrone, F. A., Hofrichter, J., and Eaton, W. A. (1985) Kinetics of sickle hemoglobin polymerization. II. A double nucleation mechanism. *J. Mol. Biol.* 183, 611–631.
57. Byeon, I. J., Louis, J. M., and Gronenborn, A. M. (2003) A protein conformationist: Core mutations of GB1 that induce dimerization and domain swapping. *J. Mol. Biol.* 333, 141–152.
58. Byeon, I. J., Louis, J. M., and Gronenborn, A. M. (2004) A captured folding intermediate involved in dimerization and domain-swapping of GB1. *J. Mol. Biol.* 340, 615–625.
59. Louis, J. M., Byeon, I. J., Baxa, U., and Gronenborn, A. M. (2005) The GB1 amyloid fibril: Recruitment of the peripheral β -strands of the domain swapped dimer into the polymeric interface. *J. Mol. Biol.* 348, 687–698.

60. Schlunegger, M. P., Bennett, M. J., and Eisenberg, D. (1997) Oligomer formation by 3D domain swapping: A model for protein assembly and misassembly. *Adv. Protein Chem.* 50, 61–122.
61. Gronenborn, A. M. (2009) *Curr. Opin. Struct. Biol.* (in press).
62. Yamasaki, M., Li, W., Johnson, D. J., and Huntington, J. A. (2008) Crystal structure of a stable dimer reveals the molecular basis of serpin polymerization. *Nature* 455, 1255–1258.
63. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) ESPript: Analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15, 305–308.

BI802292Q