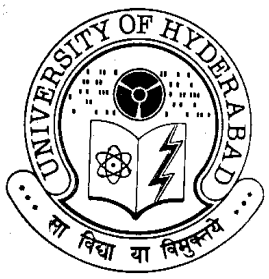
Molecular analysis of congenital cataract: Structure-function correlation of cataract-associated human γ-crystallins

Thesis submitted for the degree of

DOCTOR OF PHILOSOPHY

To

THE DEPARTMENT OF BIOCHEMISTRY
SCHOOL OF LIFE SCIENCES
UNIVERSITY OF HYDERABAD
HYDERABAD-500 046
INDIA



By

Venkata Pulla Rao Vendra

Hyderabad Eye Research Foundation
L. V. Prasad Eye Institute
Hyderabad – 500 034

February 2014

Enrolment No: 09LBPH22

Dedicated to beloved Balu sir and my parents

UNIVERSITY OF HYDERABAD
School of Life Sciences
Department of Biochemistry
Hyderabad - 500 046 (India)

DECLARATION

The research work embodied in this thesis entitled, "**Molecular analysis of congenital cataract: Structure-function correlation of cataract-associated human γ-crystallins**", has been carried out by me at the L. V. Prasad Eye Institute, Hyderabad, under the guidance of Prof. D. Balasubramanian. I hereby declare that this work is original and has not been submitted in part or full for any other degree or diploma of any other university.

Venkata Pulla Rao Vendra

Prof. D. Balasubramanian
Supervisor
Director, Research
L. V. Prasad Eye Institute
Hyderabad-500034

UNIVERSITY OF HYDERABAD
School of Life Sciences
Department of Biochemistry
Hyderabad - 500 046 (India)

CERTIFICATE

This is to certify that this thesis entitled, "**Molecular analysis of congenital cataract: Structure-function correlation of cataract-associated human γ-crystallins**", submitted by **Venkata Pulla Rao Vendra** for the degree of **Doctor of Philosophy** to the University of Hyderabad is based on the work carried out by him at the L. V. Prasad Eye Institute, Hyderabad, under my supervision. This work has not been submitted for any diploma or degree of any other University or Institution.

Prof. D. Balasubramanian
Supervisor
Director, Research
L. V. Prasad Eye Institute
Hyderabad-500034

ACKNOWLEDGEMENTS

I am very much indebted to my mentor **Prof. Dorairajan Balasubramanian** for guiding me for PhD. He has been a constant source of inspiration throughout my PhD and life. He gave me independence in work and made me think on my own which provoked the scientific instinct and more importantly self-confidence. I greatly admire his quick suggestions and solutions to problems. Every moment spent in his presence and every discussion we have had, though often just for few minutes, has been informative and a truly enriching experience. I owe him a lot for helping me grow both as a person and a researcher. I thank him for his patience, his confidence in me and for his encouragement, continued support and for making me to see the actual window of REAL SCIENCE.

I would like to thank **Prof. O.U.H.Setty** and **Dr. Manjula Sritharan,** University of Hyderabad, **Dr. Ghanshyam Swarup** and **Dr. G.R. Chandak** Centre for Cellular and Molecular Biology, **Dr. Chitra Kannabiran** and **Dr. Indumathi Mariappan,** L. V. Prasad Eye Institute, for reviewing my work constantly throughout my PhD

I am thankful to **Dr. G. N. Rao**, Chairman, L. V. Prasad Eye Institute and **Dr. CH. Mohan Rao**, Director, Centre for Cellular and Molecular Biology, for their encouragement, support and providing the necessary infrastructure.

I specially thank **Prof. N. Srinivasan, Dr. Garima Agarwal**, Indian Institute of Science and **Dr. Sushil Chandani**, Novarus Discoveries Pvt Ltd for helping me in doing the molecular modeling and dynamics of the proteins. I am thankful to **Dr. Yogendra Sharma** for his encouragement and valuable advice and suggestions and **Dr. Rajeev Raman** for his help in measurement of circular dichroism spectra. I also thank **Dr. Suman Thakur** for his help in running mass spectrophotometry.

I thank my senior, Boss, **Dr. Venu Talla** who taught me so many things in the lab and helped in establishing myself in the lab and I also thank my junior **Srinivasu K** for helping me in some experiments.

Thanks to my friends **Mangalangi UdayaKumar** and **C.Neeraja Sharma** for being with me to share the happiest and sorrowful moments throughout my PhD. Special thanks to Summer trainees **Shanmugapriya Vasudevan**, **Lauren Ostrenga**, **Anbukkarasi Muniyandil**, **Gowtham Sarvesh** and **Madhupreetha T** who helped me a lot in getting key findings.

The financial assistance from Department of Biotechnology, Department of Science and Technology and Hyderabad Eye Research Foundation is gratefully acknowledged.

My parents, my sister and brother-in-law have always been my source of strength. I thank them for showing me the way and hope to reach the heights that make them proud.

- Venkata Pulla Rao vendra

Abbreviations

μM	: Micromolar
μg	: Microgram
θ	: Elipticity
λ	: Wavelength of light , in nanometers
λ_{exc}	: Excitation wavelength
λ_{max}	: Emission maximum
ΔG°	: Gibbs free energy
C_p	: Specific heat
εM	: Molar extinction coefficient
φ	: Torsion angle
AQP 0	: Aquaporin 0
BSA	: Bovine serum albumin
bis-ANS	: 4,4'-dianilino-1,1'-binaphthyl-5,5' disulfonate
CD	: Circular Dichroism
cDNA	: Complementary DNA
cm	: Centimeter
C_m	: Equilibrium unfolding transition midpoints in units of M GuHCl
DMEM	: Dulbecco's Modified Eagle's Medium
DSC	: Differential Scanning Calorimetry
DTT	: Dithiothreitol
EDTA	: Ethylene diamine tetra acetic acid
EM	: Electron microscopy
ER	: Endoplasmic reticulum

ERGIC	: Endoplasmic reticulum Golgi intermediate compartment
FBS	: Fetal bovine serum
FITC	: Fluorescein-isothiocyanate
Da	: Dalton
GK	: Greek key
GuHCl	: Guanidine hydrochloride
h	: Hour
HGCC	: Human γC-crystallin
HGDC	: Human γD-crystallin
HGSC	: Human γS-crystallin
HLE	: Human lens epithelial cell line
I_f	: Fluorescent intensity
I_{em}	: Emission maximum
IPTG	: Isopropyl -β-D-thiogalactopyranoside
K	: Kelvin
kcal mol^{-1}	: Kilocalories per mole
KCl	: Potassium chloride
kDa	: Kilo Dalton
KI	: Potassium iodide
L	: Liter
LB	: Luria-Bertani broth
M	: Molar
MALDI	: Matrix Assisted Laser Desorption Ionization
mg	: Milligram
min	: Minutes
ml	: Milliliter

mm	: Millimeter
MRW	: Mean residual weight
Nile Red	: 9-diethylamino-5H-benzo [alpha] phenoxazin-5-one
nm	: Nanometer
NMR	: Nuclear Magnetic Resonance
ns	: Nanosecond
P	: Phenyl
PAGE	: Polyacrylamide gel electrophoresis
PBS	: Phosphate buffered saline
PCR	: Polymerase Chain Reaction
PI	: Propidium iodide
PMSF	: Phenyl methyl sulfonyl fluoride
PMT	: Photomultiplier tube
ps	: Picosecond
Q	: Quaternary ammonium
RT PCR	: Reverse Transcription Polymerase Chain Reaction
s	: Second
SDS	: Sodium dedecyl sulphate
SP	: Sulfopropyl
TEM	: Transmission Electron Microscopy
Tris-Cl	: Tris hydroxymethyl aminomethane chloride
UV	: Ultraviolet
V	: Volts
v/v	: Volume per volume
WT	: Wild type

CONTENTS

CHAPTER 1: INTRODUCTION

CHAPTER 2: MATERIALS AND METHODS

CHAPTER 3: RESULTS

Introduction

Cataract or eye lens opacification is the leading cause of blindness worldwide. It accounts for almost half of all cases globally (1). It occurs either because of mutations in the genes coding for the lens proteins (mainly in cases of congenital cataract) or due to environmental and metabolic changes that occur to the lens constituents that accumulate over a period of time, compromising transparency (age-related cataract). Congenital cataract is essentially a genetic disorder and is an important cause of childhood blindness affecting about 20,000-40,000 newborns worldwide yearly, afflicting over 1.4 million children in the world with a prevalence of 1 to 15/10,000 children (with birth prevalence of bilateral cataract being 1 to 3/10,000 births). Given a birth rate of 2% (that is, 20,000/million total population), approximately 4 children/million population/year will be born with bilateral cataract in industrialized countries, and the figure for developing countries is likely to be 10/million total population/year (2). A child cataract study in South India has shown that 25% cataracts were hereditary, 15% were due to congenital rubella and 51% due to undetermined causes (3).

1.1. Lens development

The lens is relatively a simple tissue in the sense that it is made up of cells from only one cell lineage; connective tissue, blood or nerve cells are not present. Lens cells exist in two distinct forms; fiber cells and epithelial cells. Fiber cells make up the bulk of the lens and a monolayer of epithelial cells covers the anterior surface of the fiber cells. Human lens induction starts at around 28 days' gestation. Thickening of surface ectoderm (cuboidal cells) near the optic vesicle

occurs and invaginates into the neural ectoderm, forms the lens pit and leaves the surface and become free by 33 days' gestation. Cuboidal cells remain unchanged in the anterior region but in the posterior region they elongate and form fiber cells to fill the lumen of that vesicle. These fiber cells are the first to be formed ones in the lens, are hence called as 'primary fiber cells' and this region is called the 'embryonic nucleus'. This entire process is completed by around 56 days' gestation. In the lens, depending upon the maturity of cells, two major zones can be seen, termed germinative zone and transitional zone. Epithelial cells in the germinative zone are metabolically active. Cells lying in the vicinity of the ciliary body are very active when compared to the cells which lie below the iris, and undergo mitosis to produce daughter cells that migrate to the transitional zone. Here they differentiate and elongate into the fiber cells. These fiber cells are formed later and called 'secondary fiber cells'. These differentiating fiber cells elongate into long, thin, ribbon-like structures that form the onion-like layers of the lens (Figure 1.1). During this process organelle degradation and high expression of a family of proteins called the crystallins takes place (4-6).

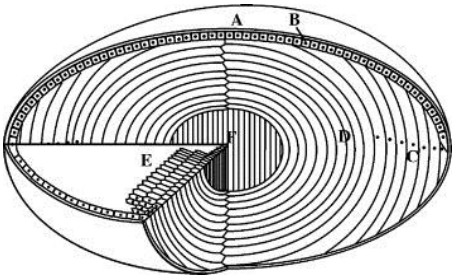


Figure 1.1: The human eye lens. A: capsule, B: epithelial cells, C: equatorial region, D: fiber cells removing their organelles denoted by the black dots, E: fiber cells with their characteristic hexagonal shape, and F: nuclear region (7).

The lens has developed numerous strategies to maintain its transparency and refraction by reducing or removing the scattering structures in it. Some of these are: (a) no vascular system, arteries and veins; (b) no organelles. Organell degradation takes place in fiber cells, losing the nucleus, endoplasmic reticulum (ER), Golgi and mitochondria, which are large enough in size as to cause light scattering (8); and (c) high expression of crystallins. They are expressed at high levels in such way that each fiber cell is filled with more than 90% of crystallins compared to the total protein content, and in the center of a human lens this concentration reaches a value of 450 mg/ml (9). Though these crystallins are expressed at such high concentrations, their compact structure, polydisperse nature and short range order allows them in keeping the lens plastic, elastic, transparent and to offer appropriate refractive gradients (10).

1.2. Lens and its constituents

1.2.1. α-Crystallin

Crystallins constitute 90% of the total lens protein content, 35% of the total lens mass and are comprised of two classes: α and βγ. α-Crystallin has two members αA and αB. In humans αA is a 173 amino acid residue protein and αB a 175 amino acid long protein. The sequence homology between αA and αB is about 57%, the molar ratio of these two proteins in mammalian lens is 3:1 and this ratio decreases gradually with age (11). α-Crystallins are expressed in lens epithelial cells as well as fiber cells whereas the expression of βγ-crystallins is restrticted largely to fiber cells (6). αA-crystallin is mainly found in the lens and

trace amounts of this protein expression are also seen in the spleen and thymus (12). αB-crystallin is an ubiquitously expressed protein present in a number of non-lenticular tissues including retina, heart, skeletal muscle, brain, spinal cord, skin, lungs (13, 14) and astrocytes in pathological conditions (15). Interestingly αB-crystallin has been associated with several neurological disorders. It accumulates in the Rosenthal fibers in Alexander disease (15), in Lewy bodies in diffuse Lewy body disease (16), in reactive glia in Creutzfeldt-Jakob disease (17) and its expression was also observed in scrapie infected hamster brain cells (18). The normal distribution of αB-crystallin in many cell types, its occurrence in neurological diseases and its increased expression under stress conditions suggest that it may play an important role in cell function. These two proteins together constitute 30% of the total lens crystallin content.

1.2.1.1. Structure of α-crystallin

The α-crystallins belong to the small heat shock protein family and associate to form large polydisperse multimers up to 60 units and form aggregates 800-1200 kDa in size, and can exchange subunits dynamically (19).

The heterogeneous nature of this molecule has made difficult to be crystallized. An alternative to X-ray crystallography, cryo-electron microscopy (cryo-EM) has been used to solve its three dimensional structure. Even though the resolution (4 nm) of this technology is not as good as X-ray crystallography, useful information has been obtained. Of the two components of α-crystallin, recombinant αB-crystallin was used for cryo-EM study, because of its less

heterogeneous nature than αA-crystallin and native α-crystallin. The cryo-EM results reveals αB-crystallin to have a variable quaternary structure with a central cavity, and these assemblies are approximately 11 nm in diameter, having a 3 nm thick protein shell, surrounding a central cavity (20, 21). A model of αB-crystallin based on cryo-EM is shown in Figure 1.2

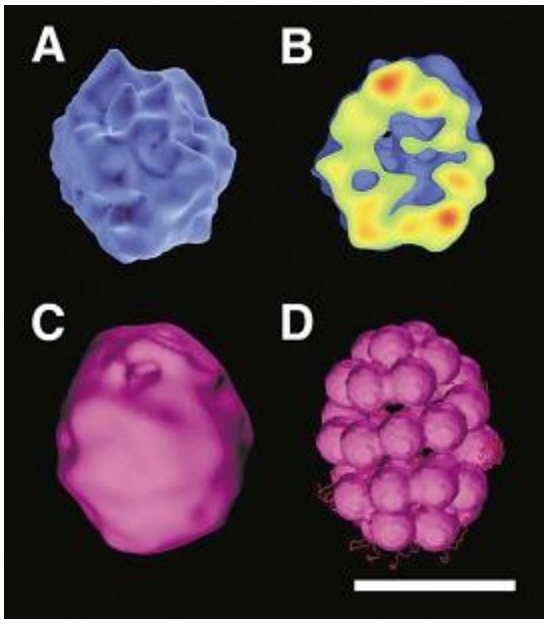


Figure 1.2: Cryo-EM reconstructions and models of recombinant αB-crystallin (A&B) and αB-crystallin with bound α-lactalbumin (C&D) as a model target protein. (A) A cryo-EM reconstruction of αB-crystallin showing the irregular surface of the protein assembly. (B) A cropped view of the αB-crystallin reconstruction showing the internal cavity. The crop plane is shown with red representing the strongest density and green the weakest. The surface is shown in blue. (C) A cryo-EM reconstruction of αB-crystallin with bound α-lactalbumin. (D) A model of αB-crystallin (magenta) with bound α-lactalbumin (red) (Scale bar 100 A˚) (20).

The recent cryo-EM study by Peschek (22) revealed human αB-crystallin oligomer to be roughly in spherical shape with a diameter of 13.5 nm (Figure 1.3A) and a central cavity of 8.5 nm diameter which is surrounded by a

symmetrical protein shell with a mean thickness varying from 2.5 to 4 nm (Figure 1.3B). The overall quaternary structure of human recombinant αB-crystallin shows similarities to the structures of Hsp20.2 from *Archaeoglobus fulgidus* (23) and Hsp16.5 from *Methanococcus Jannaschii* (24), which both form spherical, 24-meric complexes with octahedral symmetry.

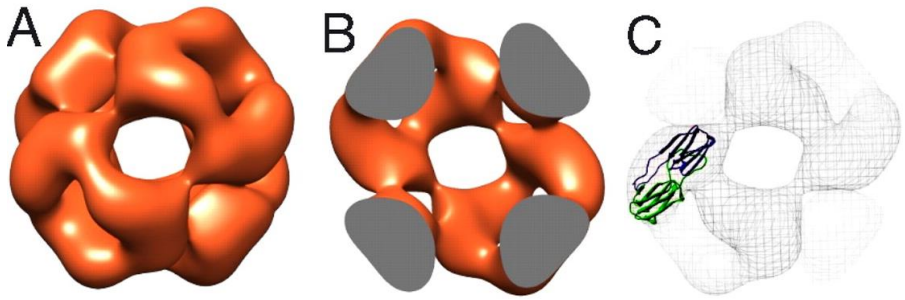


Figure 1.3: 3D reconstruction of human recombinant αB-crystallin using EM images. (A) Surface representations of the 3D model of human recombinant αB-crystallin. (B) Density cross sections through the 3D model of αB-crystallin. (C) Superposition of the ribbon representation of the dimeric αB-crystallin domain of Hsp16.5 from *M. Jannaschii* (25) (Scale bar, 10 nm) (22).

The crystal structure (2.9 A°) of excised α-crystallin domain (67-157) from human αB-crystallin has been solved and reveals that it form homodimers. Each monomer is formed with seven β-strands; one side of this beta sandwich consists of three β-strands that form an antiparallel beta sheet interaction at the dimer interface (see Figure 1.3C and Figure 1.4). These homodimers share a common groove at the interface and analyses of residues within the groove of the αB-crystallin interface shows that it has a high density of positive charges and the groove is the likely binding site for the N-terminal extension (25). In this crystal structure of this excised molecule, a crystallographic 2-fold axis is positioned between β6+β7 strands such that they form an extended antiparallel βsheet

linking two domains, similar to the proposed solution dimer of αB-crystallin based on NMR spectroscopy (26) and β2/β7 edge of the sandwich is opened out.

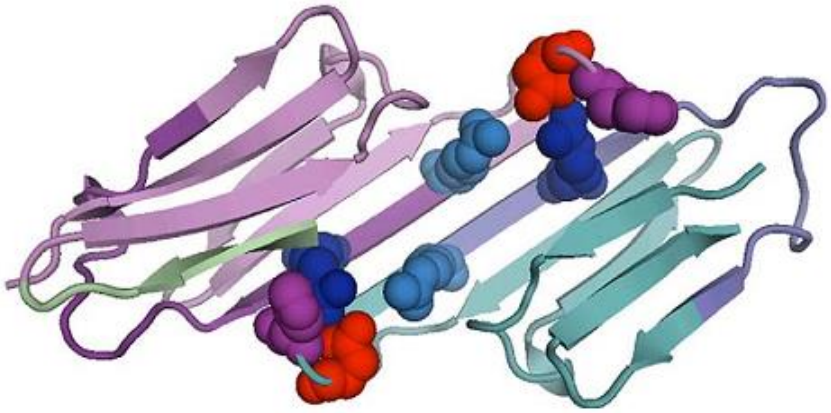


Figure 1.4: αB-crystallin dimer. (Red): D109; (Pink): H110; (Light blue): R116; (Blue): R120 (25).

The dimers form a spiral in the crystal lattice using both non-crystallographic and crystallographic axes, along with the additional β2 strand, result in the dimer being more compact and the groove being less accessible (Figure 1.4). Four residues situated at the dimer interface play important role in keeping this interface more compact D109, H110, R116 and R120 forms ion pair interaction with D109 and H110 whereas in the pH 9.0 R120 is close to partner R116 (25).

X-ray crystal structure of truncated αB-crystallin (sequence 68–162) reveals that the seventh β-strand extends away from the core domain and forms C-terminal extension. The C-terminal extension has two subunits called the hinge loop and C-terminal tail. The C-terminal tail has a highly conserved sequence that binds to beta sandwich (from top) of another molecule. The hinge loop is the

linker between the alpha crystallin core and C-terminal tail. The C-terminal extensions protrude from each end of the central dimer and these extended arms bind to adjacent dimers. The residues between 156–164 are palindomic (ERTIPITRE) and this appears to be the cause for polydispersity of the α-crystallin (27).

1.2.1.2. Chaperone-like function

The lens proteins undergo post-translational modifications with time, causing the proteins to unfold or denature during aging. These partially unfolded and denatured proteins are prone to form aggregates and this uncontrolled aggregation leads to light scattering particles that interfere with vision. In this condition α-crystallin plays a major role with its chaperone-like function by binding to partially opened and denatured β, γ-crystallins, thus attempting to prevent them from forming aggregates and light scattering particles (28, 29). The chaperone-like activity of α-crystallin has been extensively studied by Dr. Mohan Rao's group of Centre for Cellular and Molecular Biology (CCMB) and they show that chaperone-like activity of α-crystallin is temperature dependent. It does not prevent photo-aggregation at low temperatures. This protection starts around 30°C and steeply increases with temperature (30). With increasing temperature (above 30°C), a minor perturbation in its tertiary structure occurs that might lead to the exposure of its hydrophobic surfaces. The increasing temperature enhances or reorganizes the hydrophobic surfaces of α-crystallin and these structural alterations are important for its chaperone-like function and protective

ability (31) and it needs a molten-globule state or a partially unfolded stated of the target protein to bind and show its protective activity (32).

Swapping the N-terminal domain between human αA and αB-crystallins makes a more effective chaperone. In the case of N-terminal αB+ αA chimera this addition of αB N-terminal increases its accessible hydrophobic region, oligomer size and enhances chaperone-like activity, whereas N-terminal αA+αB chimera loses its protective abilities completely. This shows the importance of the N and C-terminal domains of αA and αB-crystallins in subunit oligomerization and chaperone-like activity (33).

1.2.2. βγ-Crystallins

The β-Crystallin family of proteins has seven members, four acidic (βA1, βA2, βA3 and βA4) and three basic (βB1, βB2 and βB3). These proteins account for 35% of the total crystallin content. In humans γ has three members, namely γC, γD and γS, contributing to 25% of the total crystallin content in the human lens (34). All βγ-crystallins, exist in the monomeric form, with molecular weights between 19-27 kDa. β-Crystallins exist as dimers, homo-oligomers and in many cases hetero-oligomers. They can consist of up to seven units (35-37). These two groups of proteins are essentially structural proteins and their main function in the lens is to provide transparency and necessary refractive index gradient so as to focus the light onto the retina (38).

1.2.2.1. Structure of βγ-crystallins

The amino acid sequence identity between β and γ-crystallins is around 30% and these polypeptides fold into two similar domains and each domain of the β and γ-crystallins contains two super-secondary structural arrangements, forming a β-sandwich of eight intercalated β-strands (39). Each motif consists of four sequential antiparallel β-strands, three from one motif and one from the other motif, and are arranged with -1, -1, +3 topology leaving C-terminal strand as an outer strand by generating $(3,1)_c$ type of structure (40) as shown in Figure 1.5 Left. This kind of β-sheet runs resembles the designs on Greek pottery and architecture and thus this arrangement in motif is called 'Greek key' (GK) motif (see Figure 1.5 Right).

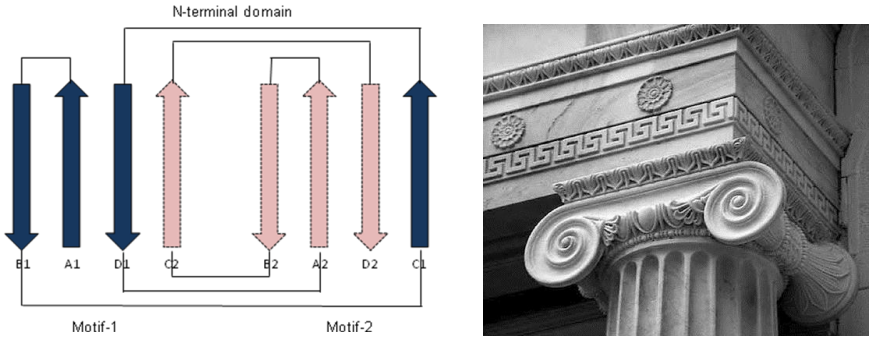


Figure 1.5 Left: Cartoon depicting the topology of the characteristic domain structure found in all βγ-crystallins showing how GK motifs, each comprising four β-strands (A, B, C and D) associate to form the overall chain fold. Figure 1.5 Right: An architecture showing the GK fold.

This kind of Greek key motif super-secondary folding, intra-molecular interactions between domains, inter-molecular interactions including 3D domain

swapping and local interactions involved in the close packing of co-operative units offers them stress resistance and structural stability (41).

Human γD-crystallin (HGDC) is a monomeric protein comprised of 173 amino acids, whose structure has been determined at 1.25 A° resolution, revealing two structurally homologous crystallin domains: N-terminal domain (residues 1-82) and C-terminal domain (residues 88-173). The two domains of HGDC are connected by a short stretch of six amino acids as the linker. These two domains consist of four Greek key motifs and contain 24 aromatic residues distributed throughout the protein: 14 tyr, 6 phe and 4 trp. 18 of these 24 residues have their aromatic partners within 5 A°.

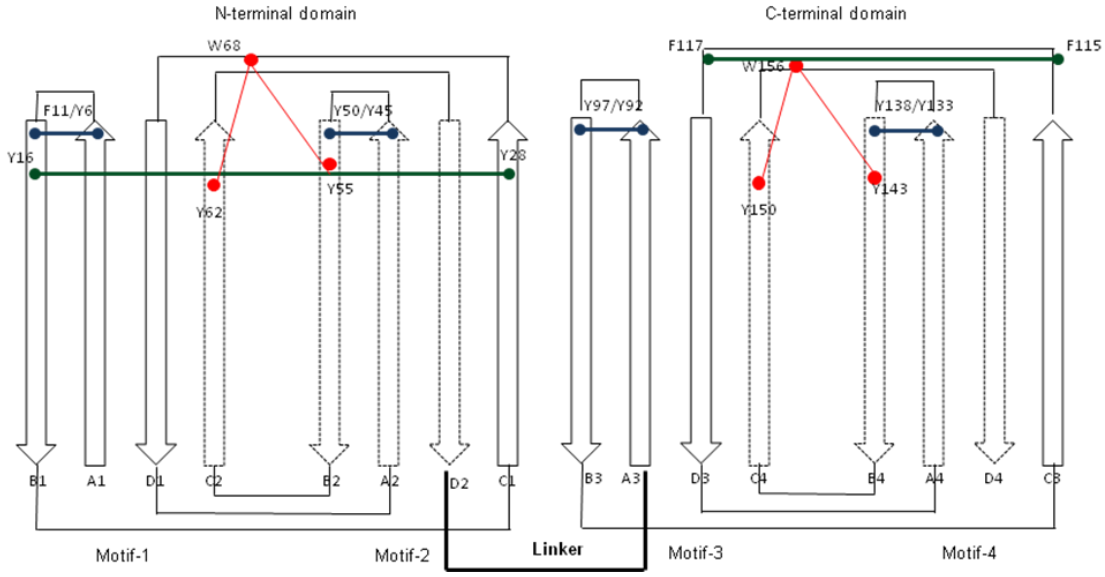


Figure 1.6: Topology diagram of HGDC and the locations of the aromatic residues. The six aromatic pairs and the two tyr-trp-tyr clusters are indicated by thick lines. Blue: four β-hairpin aromatic pairs (the ‘‘Greek key pairs’’); green: two additional β-hairpin aromatic pairs (the ‘‘non-Greek key pairs’’); arrows: β-strands.

These aromatic residues cluster into several structural elements. The tyrosine corner is one among them and is a classical feature of the Greek key motifs, bridging two β-strands by a hydrogen bond between the tyrosine hydroxyl group and a backbone carboxyl group or amino group of residue Y-3 (Δ3), Y-4 (Δ4) or Y-5 (Δ5). HGDC has two Y-4 (Δ4) type of homologous tyrosine corners, one in N-terminal domain around Y62 (Y55, Y62, W68) and another one in C-terminal domain around Y150 (Y143, Y150, W156) (Figure 1.6) (42).

Besides the tyrosine corners, six aromatic pairs, namely, four tyr/tyr, one tyr/phe and one phe/phe, are present in the β-hairpin sequences of the Greek key. Depending upon their location and close proximity these six pairs can be classified into “Greek key pairs” (Y6/F11, Y45/Y50, Y92/Y97 and Y133/138) and “non-Greek key pairs” (Y16/Y28 and F115/F117) (Figure 1.6). Substitution of these aromatic pairs with alanines revealed that Greek key pairs have larger contributions to the thermal stability than non-Greek key pairs, and among these the second Greek key pair (Y45/Y50) contribution is more important and may provide a nucleation site during the folding of the double Greek key crystallin domain (43). The tertiary domain fold is stabilized by two folded hairpin loops (formed from residues between A and B strands in each motif) that require glycine and serine residues at critical positions of the fold (44).

All the four trp residues are well buried at positions 42, 68, 130, 150 in HGDC (Figure 1.7). Of the four trp residues W68 and W156 show native state quenching while the others do not. W68 and W156 were surrounded by aromatic

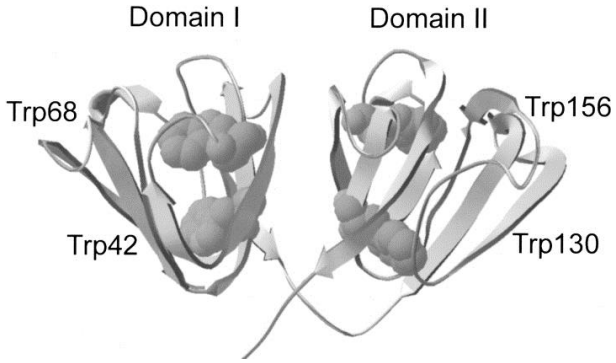


Figure 1.7: Wild type HGDC showing the location of the four native tryptophans at positions 42, 68, 130, and 156 (39).

amino acids Y55 and Y62 in the case of W68 and likewise Y143 and Y150 for W156. This aromatic amino acid shielding may be the source of the native state quenching of tryptophan fluorescence (45).

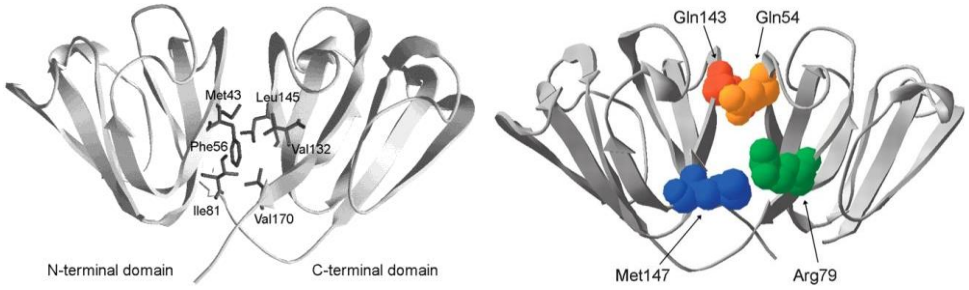


Figure 1.8: Wild type HGDC depicted in ribbon representation. Left: Amino acids contributing to the hydrophobic cluster of the domain interface (M43, F56, I81, V132, L145 and V170) are shown in wire frame (39). Right: Peripheral interface residues (Q54/Q143 and R79/M147) shown in space fill (39).

The two domains of HGDC are covalently linked by a six residue connecting peptide and interact non-covalently through interdomain amino acid chain contacts. In HGDC molecule the domain interface is composed of (a) a cluster of

six hydrophobic residues (Figure 1.8 left) and (b) two pairs of polar peripheral pairwise interactions surrounding the hydrophobic cluster (Figure 1.8 right). The hydrophobic cluster consists of M43, F56, and I81 from the N-terminal domain and V132, L145 and V170 from C-terminal domain. Peripheral pairwise interactions are between Q54/Q143 (top of the interface) and R79/M147 (nearer to the linker peptide) (39, 44). Substitutions of the N-terminal domain residues contributing to the interface destabilized the N-terminal. However substitutions of the C-terminal domain residues in the interface had little effect on the C-terminal but destabilized the N-terminal domain. This suggests that the C-terminal domain folds first and domain interface residues of the folded C-terminal domain probably nucleate the folding of the N-terminal domain (46) and the Q54/Q143 and R79/M147 interactions stabilize the native state by shielding the central hydrophobic cluster from solvent (47).

A non-native surface salt bridge between E135 and R142 located at the end of the AB folded hairpin turn appears to play a critical role in stabilizing the folding core. These highly conserved E135 and R142 of HGDC are crucial for stabilizing its hydrophobic domain interface; disruption of these charges on the HGDC surface might lead to unfolding and subsequent aggregation (48). The individual domains of HGDC exhibit differential stability with the C-terminal domain more stable than the N-terminal domain (49).

It is thanks to these extensive structural details of HGDC, made available through crystal and solution-state studies (39-49) which have helped us to

undertake the study of relating the structural disturbances arising due to mutations to the pathological phenotypes of congenital cataracts associated with mutations in γ-crystallins.

The major differences between β and γ-crystallin groups are as follows: (a) In β-crystallins the motifs are encoded by separate exons while in γ-cryatallin exons code whole domains (50), (b) β-crystallins have long N and C-terminal sequence extensions, and (c) a long linker (approximately 60 amino acid long).

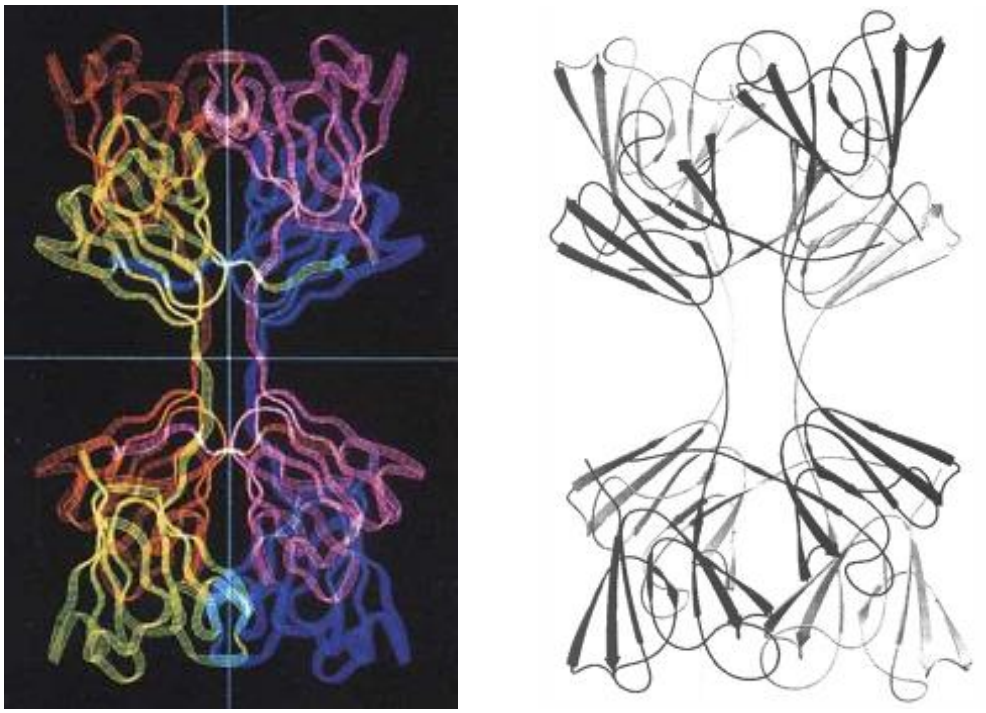


Figure1.9: Wild type βB2-tetramer showing the domain swapping of one domain of one subunit interacting with the other domain from a second subunit (51).

β-Crystallins have very similar domain structures to γ-crystallins. The X-ray structure of βB2-crystallin revealed that it has the same domain interface as the

γ-crystallins but they tend to form dimers and tetramers as the two domain assembly is changed into a four domain assembly by changing the conformation of the longer length linker and allowing domains to swap partners between two molecules as shown in Figure 1.9.

The N and C-terminal domains of a βB2 polypeptide are separated from each other with the extended linker peptide. The dimer is formed by the N-terminal domain of one subunit interacting with the C-terminal domain from a second subunit, while the C-terminal domain of the first subunit interacts with the N-terminal domain of the second subunit. The two sets of paired domains in the βB2 dimer have identical interfaces, and use the topologically equivalent residues to γ-crystallins. Domain swapping thus creates a four domain assembly, equivalent to two γ-crystallin molecules; likewise this dimer will form a tetramer (51).

Distribution of these three types of crystallins in the lens is asymmetric and biphasic (4). The central portion of the lens is rich in the β and γ-crystallins (See Figure 1.1) and the embryonic nuclear region is particularly rich in γ-crystallins (6). These proteins are present at high concentrations and are arranged compactly in a short range spatial order as in dense liquids or glasses to provide transparency and necessary refractive index (10). These two features are achieved by their being folded in a compact and highly stable manner using a series of super-secondary structural Greek key motifs.

1.2.3. Other proteins in the lens

1.2.3.1. Vimentin

Vimentin is a type III intermediate filament protein expressed in both the epithelial and fiber cells of the lens (52) and it interacts with the αB-crystallin and forms a structural framework (53).

1.2.3.2. Beaded filament structural protein 2

Beaded filament structural protein, also called as filensin, is a cytoskeletal intermediate filament protein and expressed specifically in the differentiated fiber cells of the lens (54). It has a nodular appearance and forms a meshwork underneath the plasma membrane of the lens fiber cells (55).

1.2.3.3. Aquaporin 0

The transparency of the eye lens depends upon its highly ordered tissue and architecture, the maintenance of this structure is dependent on the control of the hydration and volume of individual cells. Lens fiber cells deep in the center of lens cannot receive adequate nutrition by passive diffusion and are dependent on a transport system which can provide nutrients and remove the waste products. Each lens fiber cells is connected to its neighbour through a protein called major intrinsic polypeptide, also called as Aquaporin 0 (AQP 0), which transports water between cells and also regulates the volume. AQP 0 is the most abundant integral membrane protein in the differentiated lens fiber cell and functions as a

symmetrical water channel and provides lens transparency and homeostasis (56).

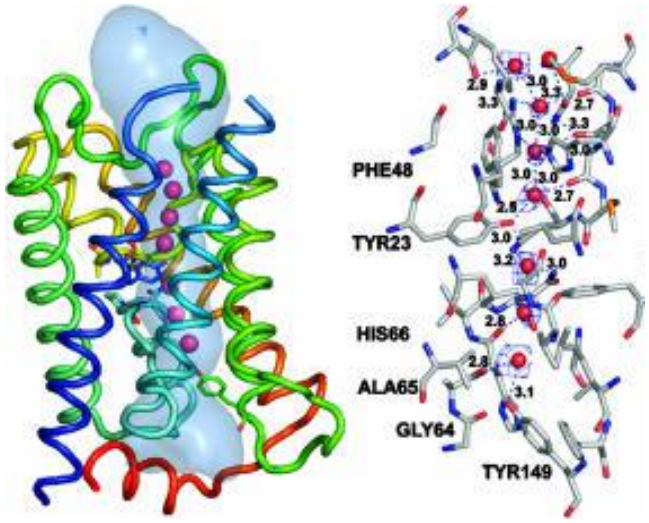


Figure 1.10: The crystal structure of AQP 0, highlighting some of the residues that generate the water channel (57).

The structure of bovine AQP 0 has been solved by X-ray crystallography and reveals an organization of the monomer molecule each of about 270 amino acid residues long forms into tetramers as shown in Figure 1.10. Each monomer is folded in such a fashion that a 28 A° long cylindrical channel ranging in diameter 1.99-2.5 A° is formed through which, water is transported. The structure is such that it also allows for cell-cell adhesion using AQP 0 (57).

Peter Agre was awarded the Nobel Prize in Chemistry for 2003 “for discoveries concerning channels in cell membranes”, with Roderick Mac Kinnon. This discovery opened the door to a whole series of biochemical, physiological and genetic studies of water channels in bacteria, plants and mammals.

1.2.3.4. Connexins

Connexins are a large family of transmembrane proteins. Three connexin proteins have been identified in the mammalian lens to date. Those are connexin 43 (Cx43), connexin 46 (Cx46) and connexin 50 (Cx50). Connexin 43 is present in the lens epithelium, connexin 46 is present in the fiber cells and connexin 50 is present in both epithelial and fiber cells.

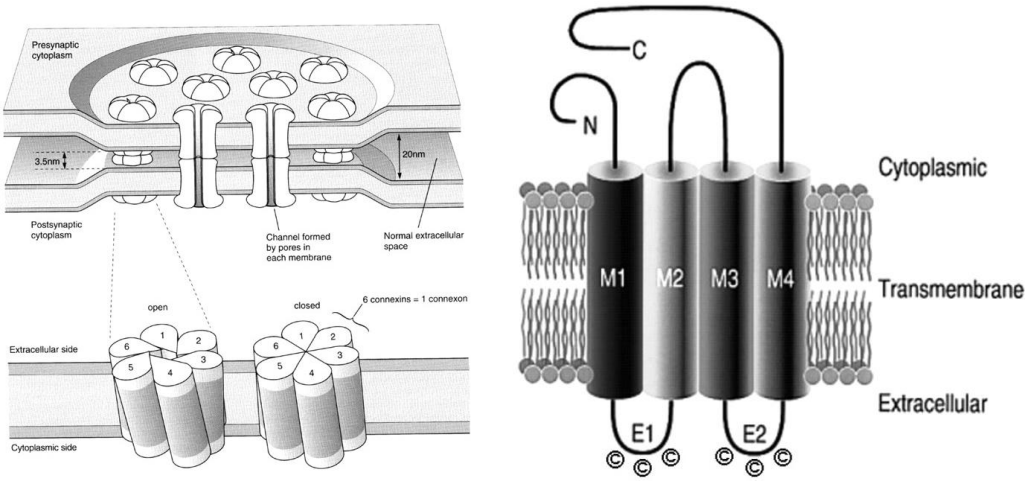


Figure 1.11: Right: Schematic diagram of the Cx50 polypeptide chain. M1-M4: transmembrane domains; E1 and E2: extracellular loop 1 and 2. The loops E1 and E2 each contain three cysteine residues. Left: Gap junction proteins organization in membranes (58).

Each connexin molecule is folded in a manner that it has four trans-membrane domains (M1- M4) and two extracellular domains (E1 and E2). While the transmembrane domains are largely hydrophobic in order to traverse the

non-polar membrane region, the E1 and E2 domains have two ‘loop’ motifs which hang out but are held together by disulfide bonds, as shown in Figure 1.11 Right panel. Connexins form multimeric complexes, actually hexamers, making what is referred to as a hemi-channel. One hemi-channel hexamer in one cell docks up with another hemi -channel in the neighbouring cell, to make a full channel, as shown in Figure 1.11 left panel. The connexins that make the channel act together by changing their conformation in synchrony in order to open and close the hemi channels. The channel closes when each subunit slides against one another and rotates in a screw-like manner, while opening is done by the reverse (58). Connexin proteins play a key role in connecting lens cells to the anterior surface of epithelial cells and maintaining inter-cellular transportation of small biomolecules, including ions, metabolites and messenger molecules (59).

1.2.3.5. Lens intrinsic membrane protein 2

Lens intrinsic membrane protein 2 (*LIM2*) is the second most abundant intrinsic membrane protein in lens fiber cells. The function of *LIM2* is not clearly understood. Its localization to the junctional region of the lens fiber cell membrane as well as throughout the fiber cell membrane suggests its role in lens junctional communication (60, 61)

1.3. Genetics of cataract

When mutations in crystallins or other lens proteins lead to protein aggregation they usually result in congenital cataract, while if they suffer

increased susceptibility to environmental insults, (such as light, oxidative damage, hyperglycaemic conditions) they might contribute to age related cataract. Mainly hereditary congenital cataracts tend to be inherited in a Mendelian fashion with high penetrance and all three classical types of Mendelian inheritance (autosomal dominant, autosomal recessive and X-linked) have been reported with congenital cataract. There are currently about 45 genetic loci and over 38 genes have been identified with isolated cataracts. Genes causing Mendelian cataracts are listed in Table 1.1 (62). Crystallin genes and connexin genes (*GJA3 and GJA8*) are among the most frequently mutated genes in hereditary cataracts.

1.3.1. Congenital cataracts with connexins

The genes coding for gap junction proteins *GJA3* (connexin 43) and *GJA8* (connexin 50) have been mapped to 13q11 and 1q21.1 respectively (63). At least 25 mutations in *GJ3* and 27 mutations in *GJA8* have been identified in families with congenital cataracts to date. Five mutations have been identified from our own institute from Dr. Chitra Kannabiran's group, and these include a homozygous missense mutation of c.649G>A (V196M) in *GJA8* in a family with autosomal recessive cataract, two heterozygous missense mutation, c.658C>T (P199S) in GJA8 and c.589C>T (P197S) in *GJA3* in two separate families with autosomal dominant cataract, a silent change c.84G>A (V28V) in *GJA8* with an autosomal dominant inheritance of cataract (64) and a single base insertion causing a frameshift at codon 203 c.670ins A (T203NfsX47) of connexin (65).

This insertion of A results in a frameshift at codon 203 with a stop codon after 46 amino acids of altered reading frame, producing a truncated protein consisting of 248 amino acid residues. Functional analysis of this frame shift mutant shows that it is localized to the ER but not in the plasma membrane and causes disintegration of the Endoplasmic reticulum Golgi intermediate compartment (ERGIC) and Golgi. Further analysis showed that the fourth transmembrane domain and a membrane proximal region (231–294 amino acids) of the cytoplasmic domain are needed for transport from the ER and localization to the plasma membrane (66).

1.3.2. Congenital cataracts with crystallins

Turning to the crystallins, mutations in 10 crystallin genes have been identified with isolated cataracts. Those are two α-crystallin genes (*CRYAA*, 21q; *CRYAB* 11q), two acidic β-crystallin genes (*CRYBA1*, 17q; *CRYBA4*, 22q), three basic β-crystallin genes (*CRYBB1*, 22q; *CRYBB2*, 22q; *CRYBB3*, 22q) and three γ-crystallin genes (*CRYGC*, 2q; *CRYGD*, 2q; *CRYGS*, 3q) (62).

The γ-crystallin coding genes are organized as a cluster of five genes (*CRYGA* → *CRYGE*) within an approximately 50 kb region and are arranged in a head to tail orientation. The sixth gene *CRYGF* is present close to this cluster and all these genes were found on chromosome 2. A typical *CRYG* gene is composed of three exons and two introns; the first exon contains only 9 bp followed by a short intron of about 100 bp. The second exon contains 243 bp and the third exon is composed of 273-276 bp and these two exons are separated by

a large intron of 1-2 kb. The non-translated 3'-end is short and is about 40bp (67). The *CRYGE* and *CRYGF* genes are not expressed in humans and therefore referred as pseudogenes (68). Cataract causing mutations in humans have so far been detected in *CRYGC*, *CRYGD* and *CRYGS* genes only. Mutations reported in α-crystallins, β-crystallins and γ-crystallins are reported in Table 1.2 (69-91), Table 1.3 (92-129) and Table 1.4 (130-156) respectively.

1.3.2.1. Mutations in α-crystallins

Of the 11 mutations in αA-crystallin (*CRYAA*) represented in Table 1.2A in 10 of them an arginine residue is involved and changes of a neutral or hydrophobic amino acid to or from arginine (R12C, R21L, R21Q, R21Q, R49C, R54C, R54P, G98R, R116C and R116H). This arginine involvement in these cataracts would tend to alter the molecular charge dispersion at the surface of the α-crystallin protein and show a deleterious effect on its chaperone-like activity.

The R116C mutation was the first human *CRYAA* mutation reported. A broad variety of experimental approaches have been undertaken to find out the functional consequences of this mutation. The mutant protein shows reduced chaperone-like activity (157), interacts with βB2 and γC-crystallins (158) as well as with actin (159). The R49C mutation induces abnormal localization of αA-crystallin to the nucleus (73). In the case of G98R mutation, the resulting protein forms larger oligomers that are prone to precipitate over time and diminished chaperone-like activity (160). An autosomal recessive cataract has been

associated with a chain termination mutation near the beginning of the protein (W9X), converting the tryptophan codon at position 9 into a termination codon. In heterozygous condition this early chain termination mutation would be expected to not to affect the protein synthesized from the normal allele, suggesting that half the normal level of α-crystalllin can provide sufficient chaperone-like activity and structural packing to establish and maintain lens transparency (69).

Half of the mutations in *CRYAB* gene (V152fs, Q151X, G154S, R157H) are associated with myopothies and two mutations reported so far are in regulatory elements (249C>G, 650C>G and 652A>G (promoter region); and 2398T>G (intronic region)), associated with multiple sclerosis and others (R11H, P20S, R56W, D109H, R120G, D140N, R149fs and A171T) are associated with cataracts.

Table 1.2B lists 14 mutations reported in αB-crystallin (*CRYAB*) associated with congenital cataract. Here again 4 of them involve replacement of arginine. The first human mutation reported in αB-crystallin (R120G) was associated with desmin-related myopathy as well as to cataract (83). This is a missense mutation, which reduces the chaperone-like activity dramatically, by affecting the tertiary structural packing (157) and causing aggregation and precipitation of the mutant protein under stress (161). This again (just as the above arginine related mutants in αA-crystallin) suggests that conserved arginines of α-crystallin chains are essential for the protein's structural integrity which in turn affects its protein's function. The P20S mutation in *CRYAB* can

induce crystallin subunit exchange, markedly decreased chaperone-like activity, abnormal nuclear localization, and increased apoptosis compared to the wild type gene (162). The D140N mutation alters the tertiary and quaternary structures and increases protein aggregation and susceptibility to thermal denaturation. (84)

1.3.2.2. Mutations in β-crystallins

These include missense mutations, insertions changing the reading frame and causing expression of aberrant peptides with premature termination and splice mutations as shown in Table 1.3. *CRYBA1* codes for two β-crystallins, βA1- and βA3-crystallin, which differ by the length of their N-terminal extensions (163). Three mutations have been observed so far in βA3/βA1. Of these two of them affect the same splice site of intron 3: in the first case, it is a G→A transition and in the second case a G→C transversion. The third mutation causes G91 deletion. These three mutant proteins lead to zonular, pulverulent and lamellar cataracts, respectively (117-127).

1.3.2.3. Mutations in γ-crystallins

Mutations in γ-crystallins tend to produce nuclear or peripheral cataracts consistent with their degree of damage to their structure. The P24T mutation in *CRYGD* induces markedly lower solubility due to hydrophobic patches on its surface ((164, 165).Two mutations in *CRYGD*, R36S and R58H, have been shown not to alter the protein fold but to lower the solubility and enhance the

crystal nucleation rate of these mutants. This causes them to precipitate out of solution through hydrophobic interactions in the R36S mutant, and by forming crystals in the lens in the case of the R58H mutation ((166, 39). In another mutant of *CRYGD*, R14C, this protein too maintains its tertiary structure but is susceptible to thiol-mediated aggregation (167, 168).

The T5P mutation in *CRYGC* affects the protein conformation and decreases its solubility and stability; the mutation also alters the interaction between γC-crytallin and other crystallins and increases aggregation (169, 170). The R168W mutation in *CRYGC* induces changes in tertiary structure of the molecule and enhances self-aggregation to form light scattering particles (171), and in the mutant W156X deletion of 18 residues of *CRYGC* causes self-aggregation since the side chains of several hydrophobic residues in the sequence are exposed to the solvent; a similar kind of aggregation happens in the truncated mutant W157X of *CRYGD* (172).

1.4. Scope of the study

The role of the Greek key motif, based on the protein's unique topology has helped us understand the structural perspective as described in detail above in section 1.2.2, but how this Greek key-derived dense packing of the β and γ-crystallins in the eye lens translates into transparency is an issue of functional interest. It is with this in mind that we have chosen congenital cataracts associated with mutations in β and γ-crystallins to study the functional role of the Greek key. The main reason for selecting congenital cataract mutations to study

this issue is twofold. One is that congenital cataracts are essentially due to genetic mutations in origin and two is the fact that the 52 mutations that have been reported so far in human β and γ-crystallins are associated with an interesting phenotypic dichotomy (Table 1.3 and 1.4). About half of them generate nuclear cataract, blocking the central visual axis (Figure 1.12.A) and hence causing complications which can disrupt the development of the visual pathways leading to conditions such as amblyopia (visual stimulation either fails to transmit or is poorly transmitted through the optic nerve to the brain), nystagmus (involuntary movement of the eye often called as dancing eye) and possibly alter association deficits in the growing infant, if not treated at the earliest (173-177). The rest are associated with peripheral cataracts which do not demand early action, since they do not block the visual axis (Figure 1.12B).

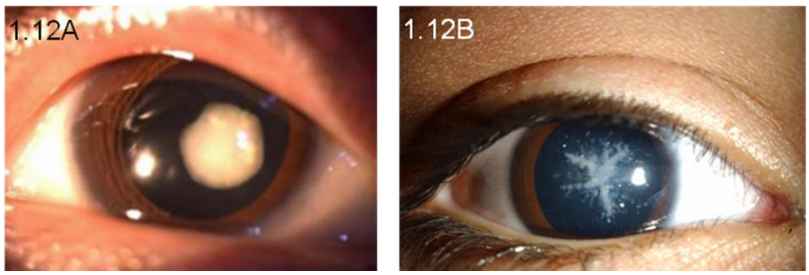


Figure 1.12A: Slit lamp photograph of the eye of a nuclear cataract patient (153). Figure 1.12B: Coralliform cataract is similarly shown (Coralliform: round or elongated processes radiating out of the center of the lens) (138).

In an effort to understand the molecular structural basis behind this phenotypic dichotomy (nuclear cataract-peripheral cataract), we have chosen to study representative mutations in HGDC and Human γS-crystallin (HGSC). The reasons for selecting these two are: (a) they are monomers; the crystal and

solution structures are well known for HGDC (39) and the crystal structure of the C-terminal domain of HGSC (178) has been studied in detail and the solution structure of the full length murine γS-crystallin has been resolved by NMR (179), (b) structural analysis of several mutants (R14C, P24T, R36S, R58H, G61C, E107A and W157X of HGDC and G18V of HGSC are already available to extend our analysis and (c) these cataracts caused by the above mutations are Mendelian monogenic disorders.

It thus becomes possible to go beyond reporting the mutations in βγ-crystallins associated with congenital cataract and to study the functional genetics or attempt a protein structural rationale. We have chosen three mutants, two from HGDC (P24T and R77S) and one from HGSC (D26G) which are associated with peripheral cataracts, and three mutants (A36P, R140X and G165fs) of HGDC and the mutant V42M of HGSC associated with nuclear cataract. In addition, we have prepared two full length chain mutants of HGDC, not reported naturally, namely, Y134A, which is a mutation in the fourth Greek key motif, but which still keeps all the four motifs intact, and the double mutant L45PL54P which disrupts the second Greek key motif, of HGDC. We have then compared their properties with the other mutants and their respective wild types (WT) by cloning, expressing and isolating the proteins, purifying them and studying their structural details in solution (P24T, R77S, A36P, R140X, G165fs L45PL54P and Y134A of HGDC and D26G and V42M of HGSC). To this end, we have compared their solubility, their secondary and tertiary structural features using circular dichroism (CD) and fluorescence spectroscopy, their structural

stability towards chemical denaturation (using the commonly used chemical denaturing agent guanidine hydrochloride, (GuHCl)) and thermal denaturation, their tendency to form light scattering aggregate particles *in vitro* using extrinsic spectral probes, and *in cellulo* by transfecting their cDNAs into a lens cell line, as well as by molecular modeling *in silico* by standard methods.

Table 1.1: Genes causing non-syndromic Mendelian cataract

Chromosome	Locus/gene	Cataract type
1p36	CCV	AD
1p36.1	115665/EPHA2	AD
1p32	176946/FOXE3	AR and AD
1q21.1	601094/GJA8	AD
1q25-q31	600897	AD
2p12	115800/CCNR	AD
2q33-q35	607304/CRYGC	AD
2q33-q35	123680/CRYGD	AD
3p21.3-p22.3	123690/FYCO1(CATC2)	AR
3q22.3-q25.2	607182/ADCCC	AD
3q21-q22	BFSP2	AD
3q25	603212/CRYGS	AD
6p24	123730/GCNT2	AR
7q34	600429/AGK	AR
8q13.3	610345/EYA1	AD
9q13-q22	601653/CAAR	AR
9q22.33	605749/TDRD7	AR
10p13	611258/VIM	AD
10q23.31	193060/SLC16A12	AD
10q25	605749/PITX3	AD
11p13	602669/PAX6	AD
11q22.1- q23.2	607108/CRYAB	AD
12q13	123590/AQPO	AD
12q24	154050/CCA5	AD
13q11-q12	614422/GJA3	AD
14q22-q23	121015/CTPP5	AD
14q24.3	610634/VSX2	AR
15q21-q22	142993/CCSSO	AD
16p13.2	605728/TMEM114	AD
16q22.12	611579/HSF4	AD
16q22-q23	602438/MAF	AR and AD
17p13-p12	177075/CTAA2	AD
17q11.2-q12	601202/CRYBA3/A1	AD
17q24	123610/GALK1	AR
17q24	604313/CCA1	AD
19q13	609376/CATCN1	AR
19q13.4	115660/FTL	AD
19q13.4	134790/LIM2	AR
20p11.23	154045/BFSP1	AR
20p12-q12	603307/CHMP4B	AD
21q22.3	605387/CRYAA	AD
22q11.2	123580/CRYBB1 600929/CRYBB2	AD AD

	123620/CRYBB3 123630/CRYBA4	AR and AD AD
Xp22	123631/CXN3X	XL
Xq24	600929/X-linked	XL

AD: autosomal dominant; AR: autosomal recessive. (Taken from [62]).

Table 1.2: Mutations in human α-crystallins associated with congenital cataracts

Table 1.2A: Mutations in human αA-crystallins associated with congenital cataracts

S.No	Mutation	Type of cataract seen	Reference
1	W9X	Congenital	(69).
2	R12C	Nuclear	(70).
3	R21L	Nuclear	(71).
4	R21Q	Nuclear	(72).
5	R21W	Nuclear	(70).
6	R49C	Nuclear	(73).
7	R54C	Punctate lenticular	(74).
8	R54P	Y shape suture	(75).
9	G98R	Presenile	(76).
10	R116C	Nuclear	(77).
11	R116H	Nuclear	(70, 78).

Table 1.2B: Mutations in human αB-crystallins associated with congenital cataracts

S.No	Mutation	Type of cataract seen	Reference
1	R11H	Nuclear	(79).
2	P20S	Posterior polar	(80).
3	R56W	Congenital	(81).
4	D109H	Posterior polar	(82).
5	R120G	Myopathy	(83).
6	D140N	Lamellar	(84).
7	450 del A;R149fs	Posterior polar	(85).
8	464 del CT; V152fs	Myopathy	(86).
9	Q151X	Myopathy	(86).
10	G154S	Myopathy	(87).
11	R157H	Myopathy	(88).
12	A171T	Lamellar	(99).
13	249C>G, 650C>G and 652A>G	Multiple sclerosis	(90).
14	249C>G, 650C>G and 652A>G (promoter region) and 2398T>G (intronic region)	Multiple sclerosis	(91).

Table 1.3: Mutations in human β-crystallins associated with congenital cataracts

Table 1.3A: Mutations in human βB1-crystallins associated with congenital cataracts

S.No	Mutation	Phenotype	Comments
1	MIK	Nuclear and Pulverulant	Abrogates initiate codon, even if translation occurs, it loses GK-1 and part of GK-2 (92).
2	N58Tfs106	Nuclear	Elongates the C-terminus, disrupting the β-crystallin interactions and truncation (93).
3	S129R	Nuclear and microcornea	Destabilizes the βB1/βA3-crystallin heteromer but not the βB1 homomer and also sensitizes βB1 to proteolysis (94, 95).
4	G220X	Pulverulant (affecting fetal nucleus)	Truncation of 33% of the GK-4, and the entire C-terminus, losing solubility (96).
5	Q 223X	Nuclear	Truncation, losing part of GK-4 motif (97).
6	S228P	Nuclear	Loss of the conserved S228, which, along with a distal glycine, enables the super secondary fold by packing the hairpin over the β-sheet in the GK motif. Replacing it by the structure breaker P disturbs the supersecondary fold (98).
7	R233H	Nuclear	PolyPhen analysis shows damage to the protein structure, and loss of solubility due to loss of the conserved cationic arginine (99).
8	X253R	Nuclear and microcornea	Elongates the C-terminal extension and disrupts β-crystallin interactions (100).

Table 1.3B: Mutations in human βB2-crystallin associated with congenital cataracts

S.No.	Mutation	Phenotype
1	A2V	Posterior subcapsular (101).
2	I21N	Dense white opacities throughout the embryonic and fetal nuclei (99).
3	S31W	Coronary (102).
4	cDNA-G54A splice site variation	Zonular (103).
5	W59C	Central opacity (103).
6	D128V	Nuclear with by ring-shaped opacity (104).
7	V146M	Nuclear (99).
8	W151C	Nuclear (105).
9	Q155X	Variable genotype, but predominantly Nuclear, Coppock-like, Pulverulant opacity in embryonic nucleus (106-113).
10	Y159X	Not reported (114).
11	V187M	Nuclear (115).

Table 1.3C: Mutations in human βB3-crystallin associated with congenital cataracts

S.No	Mutation	Phenotype	Comments
1	R75H	Microcornea cataract	Mutation affects the GK-2 motif, and destroys a highly conserved amino acid R75 (114).
2	G165R	Nuclear cataract with cortical riders	Mutation Destabilizes GK-4 Motif and increases electrostatic charge surface (116).

Table 1.3D: Mutations in human βA3/A1-crystallin associated with congenital cataracts

S.No	Mutation	Phenotype	Comments
1	Splice site mutation in first base in intron 3: CRYBA1-IVS3+1 G>A	Nuclear, Pulverulant, Zonular, Y-Sutural opacity.	Exon skipping abolishes GK-3 (117-121).
2	Splice site mutation in first base in intron 3: CryBA1-IVS3+1 G>C	Pulverulant opacities and Sutural opacities in embryonic nucleus.	(122).
3	3 bp deletion in exon 4 leading to deletion of glycine 91; ∆91	Nuclear.	Defective folding of two β- strands making the GK motif (123-127).

Table 1.3E: Mutations in human βA4-crystallin associated with congenital cataracts

S.No	Mutation	Phenotype	Comments
1	F94S	Lamellar	(128).
2	L69P	Lamellar and microcornea,	(128).
3	G64W	Nuclear and microcornea	G64 is a highly conserved residue across species (129).

Table 1.4: Mutations reported in human γD, γC and γS-crystallins associated with congenital cataracts.

Table 1.4A: Mutations reported in human γD-crystallin

S.No	Mutation	Type of cataract seen	Comments
1	R14C	Nuclear and perinuclear	Only surface change; extensive disulfide Inter-molecular bridges (130, 131).
2	R14S	Coralliform	Slight change in hydrophobic content, opening up a phosphorylation site (no conformational work) (132).
3	P24T	Coralliform, Cerulean and Fasciculiform	Extensive conformational analysis done by many shows no change in 2°and3° structure, minor surface changes, drop in solubility (133-139).
4	A36P	Nuclear	Our study shows GK compactness distorted; higher hydrophobic exposure upon mutation; and solubility drops (140).
5	R37P	Nuclear	No data yet (141).
6	R37S	Birefringent crystals	Crystal structure reveals no changes (142).
7	W43R	Nuclear	Minor 3° structural changes; pI must be different, solubility decreased (143).
8	M44V	Blue dot opacity	No structural data yet (103).
9	Y56X	Nuclear	Large scale truncation of chain; three GK motifs lost (144).
10	R58H	Aculeiform	H bonding lost. But no other change. pI altered (145).
11	G61C	Coralliform	Stability is compromised (146, 147).
12	R77S	Polar coronary	No change seen in structure upon mutation. Solubility high (105).
13	E107A	Nuclear	Solubility high. pI change is seen to lead to hetero aggregation with α-crystallin (148).
14	Y134X	Not reported	No data, but GK-4 gone (70).
15	R140X	Nuclear	Our data shows GK-4 lost; solubility lost; surface exposure extensive (89).
16	W157X	Nuclear	GK-4 gone; solubility lost; surface exposure extensive (117).
17	G165fs	Nuclear	Our data shows high hydrophobic exposure; GK-4 distorted and solubility lost (149).

Table 1.4B: Mutations reported in human γC-crystallin

S.No	Mutation	Type of cataract seen	Comments
1	T5P	Coppock-like	Less solubility and lower thermal stability (145).
2	G41fs	Zonular Pulverulant	Less solubility (150).
3	C109X	Nuclear	Truncation and loss of GK (151).
4	S119S	Nuclear	(144).
5	W157X	Nuclear	Truncation mutant; data shows solubility loss, exposure of residues to surface (152).
6	R168W	Nuclear	pI change, as in E107A? (153).
7	R48H	Nuclear	Similar to the γD-mutant (154).

Table 1.4C: Mutations reported in human γS-crystallin

S.No	Mutation	Type of cataract seen	Comments
1	G18V	Progressive cortical	Extensive data shows no change in 2° or 3°, but compaction of GK-1 likely to alter (155).
2	D26G	Coppock	Our data shows very little change in 2° or 3°, solubility high (140).
3	S39C	Microcornea-cataract	No data (89).
4	V42M	Autosomal dominant cataract	Our data shows compact packing of GK is lost (156).

Materials & Methods

2.1. Materials

2.1.1. Vectors used for cloning and expression

Three plasmid vectors pBSSK+ (Stratagene), pET-21a (+) (Novagen) and pCDNA3.1 (+) (Invitrogen) were used in this study. pBSSK+ was used for cloning and maintenance of various cDNAs; pET-21a (+) for over-expression of recombinant proteins in *E.coli* and pCDNA3.1 (+) for checking aggregation properties in mammalian cells.

2.1.1.1. pBSSK+vector: This consists of a pUC origin with the *lac* promoter, β-galactosidase α-fragment and ampicillin resistance gene as selection marker plus two blunt end cut enzymes *EcoRV* and *SmaI* which are useful in incorporating the blunt end product into the vector. Further details are shown in Figures 2.1 and 2.2.

2.1.1.2. pET-21a (+) vector: Of pBR322 origin, it has T7 promoter, *lacI* coding sequence, N-terminal T7 tag coding sequence, C-terminal His tag coding sequence and ampicillin resistance gene as selection marker. Its expression can be regulated by the chemical isopropyl-1-thio-D-galactopyranoside (IPTG) using which high levels of recombinant protein expression can be achieved in *E.coli*. Further details are shown in Figure 2.3 and 2.4.

2.1.1.3. pCDNA3.1 (+) vector: It has pUC origin, human CMV and SV40 promoters, C-terminal His tag coding sequence, ampicillin and neomycin resistance genes as selection markers and BGH polyadenylation sequence. The

pUC origin gives high copy number growth and replication; human CMV promoter permits efficient high-level expression of recombinant protein; ampicillin resistance gene for selection of vectors in *E.coli*; neomycin resistance gene for selection of stable transfectants in mammalian cells and BGH polyadenylation sequence for efficient transcription termination and polyadenylation of mRNA. Further details are shown in Figure 2.5 and 2.6.

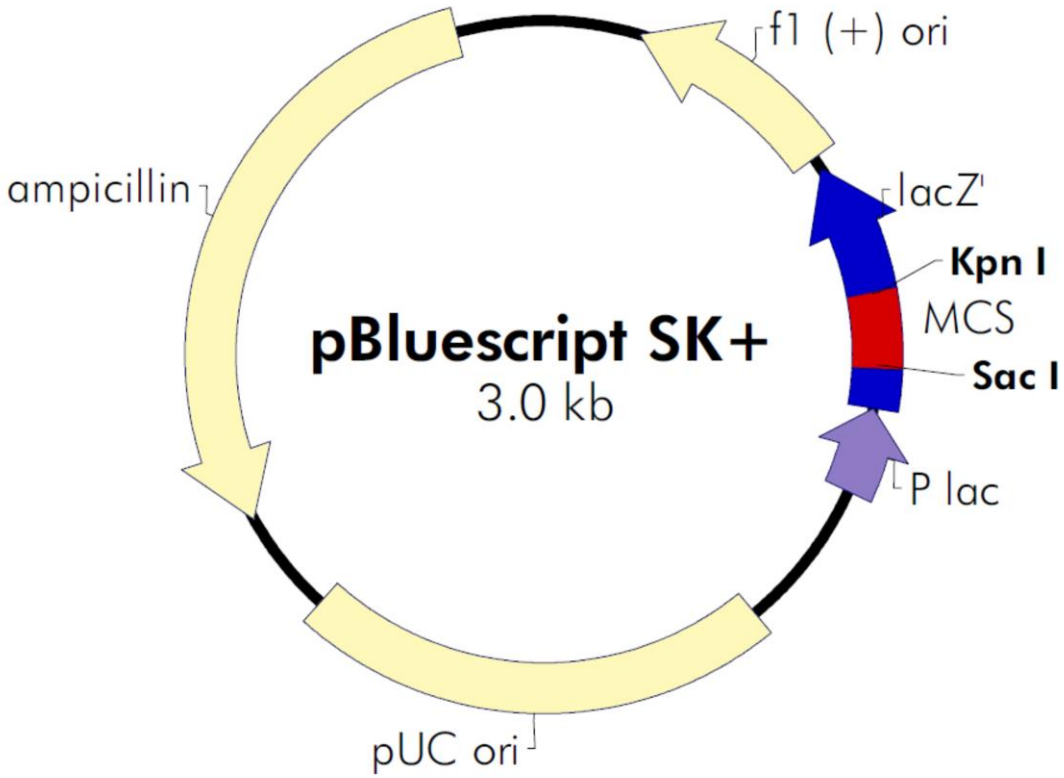


Figure 2.1: Vector map of pBSSK+ showing pUC origin, multiple cloning site region, *lacZ* gene and ampicillin resistance gene marker.

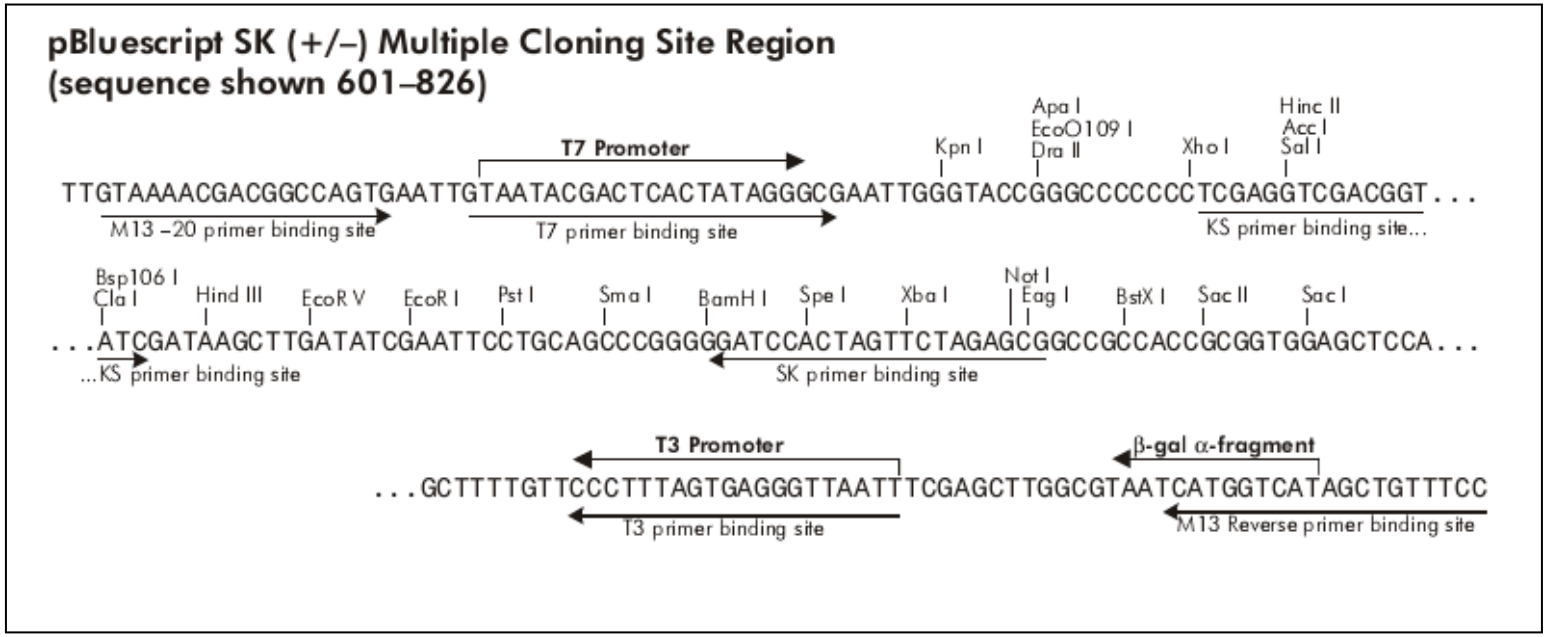


Figure 2.2: Multiple cloning site region of pBSSK+ vector along with T7 promoter and β-galactosidase α-fragment. Human γD-crystallin (HGDC) and human γS-crystallin (HGSC) cDNAs were cloned into this vector at *SmaI* site.

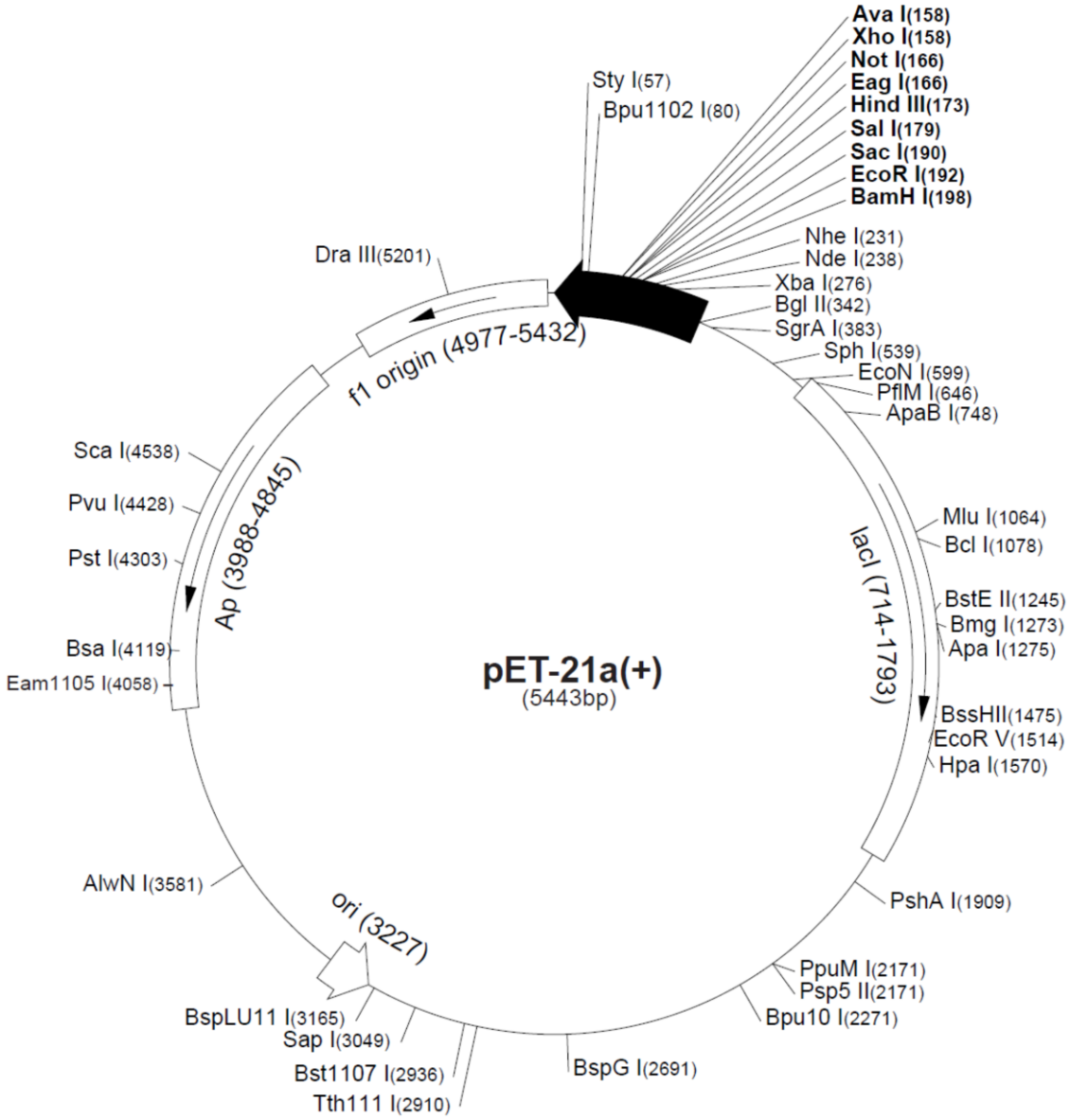


Figure 2.3: Vector map of pET-21a (+) showing pBR322 origin, *lacI* gene, multiple cloning site and ampicillin resistance gene.

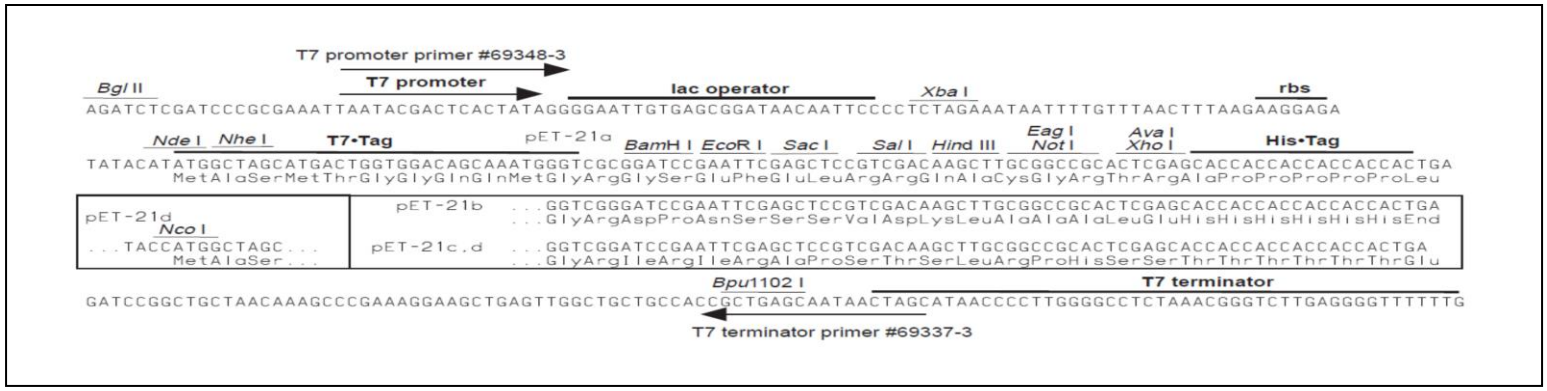


Figure 2.4: Multiple cloning site region of pET-21a (+) vector along with T7 promoter, *lac* operator, ribosomal binding site, T7 tag and His tag. HGDC and HGSC cDNAs were cloned into this vector at *NdeI* and *HindIII* sites.

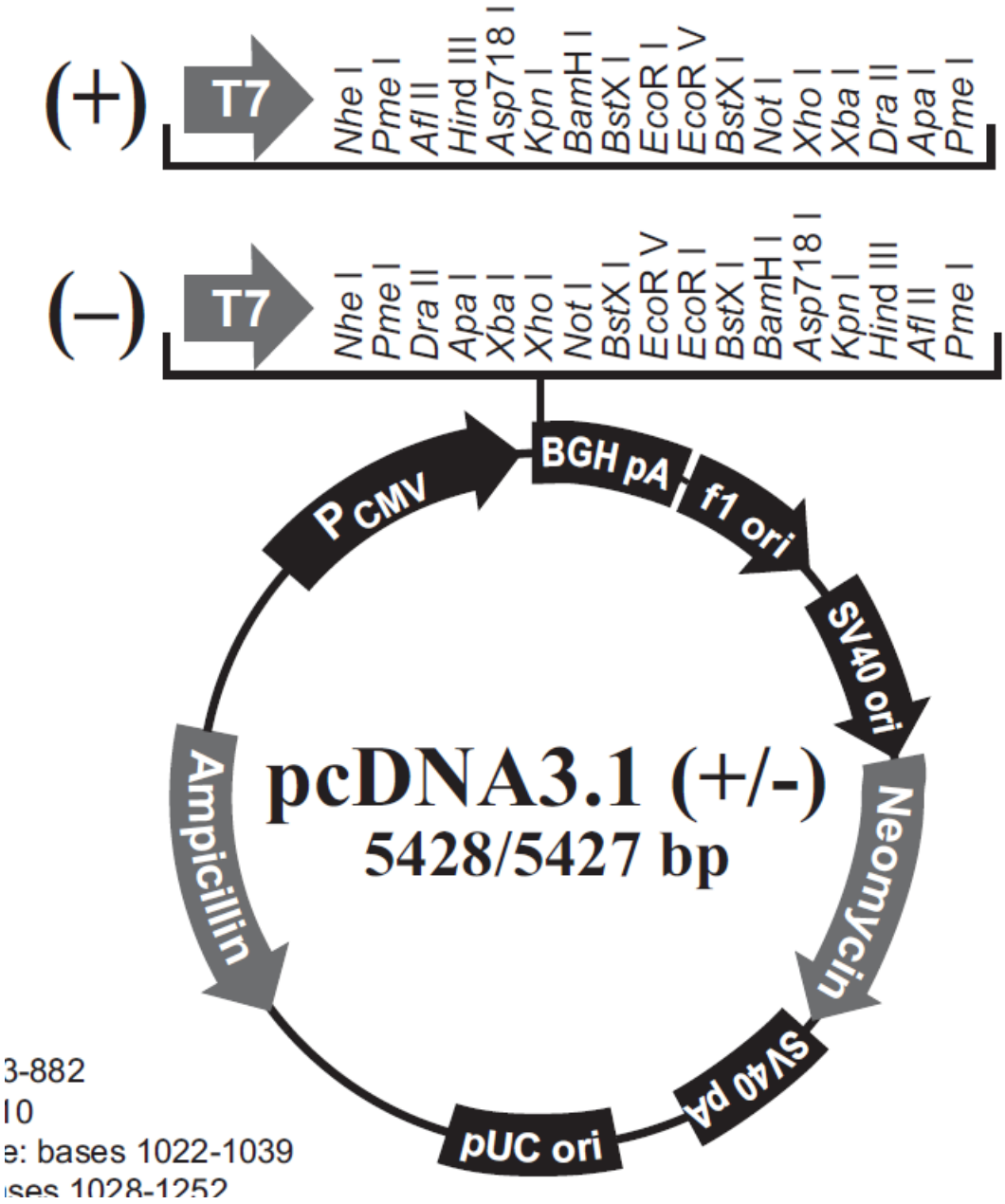


Figure 2.5: Vector map of pCDNA3.1 (+) showing pUC origin, CMV promoter, multiple cloning site, BGH polyadenylation site, ampicillin and neomycin resistance genes.

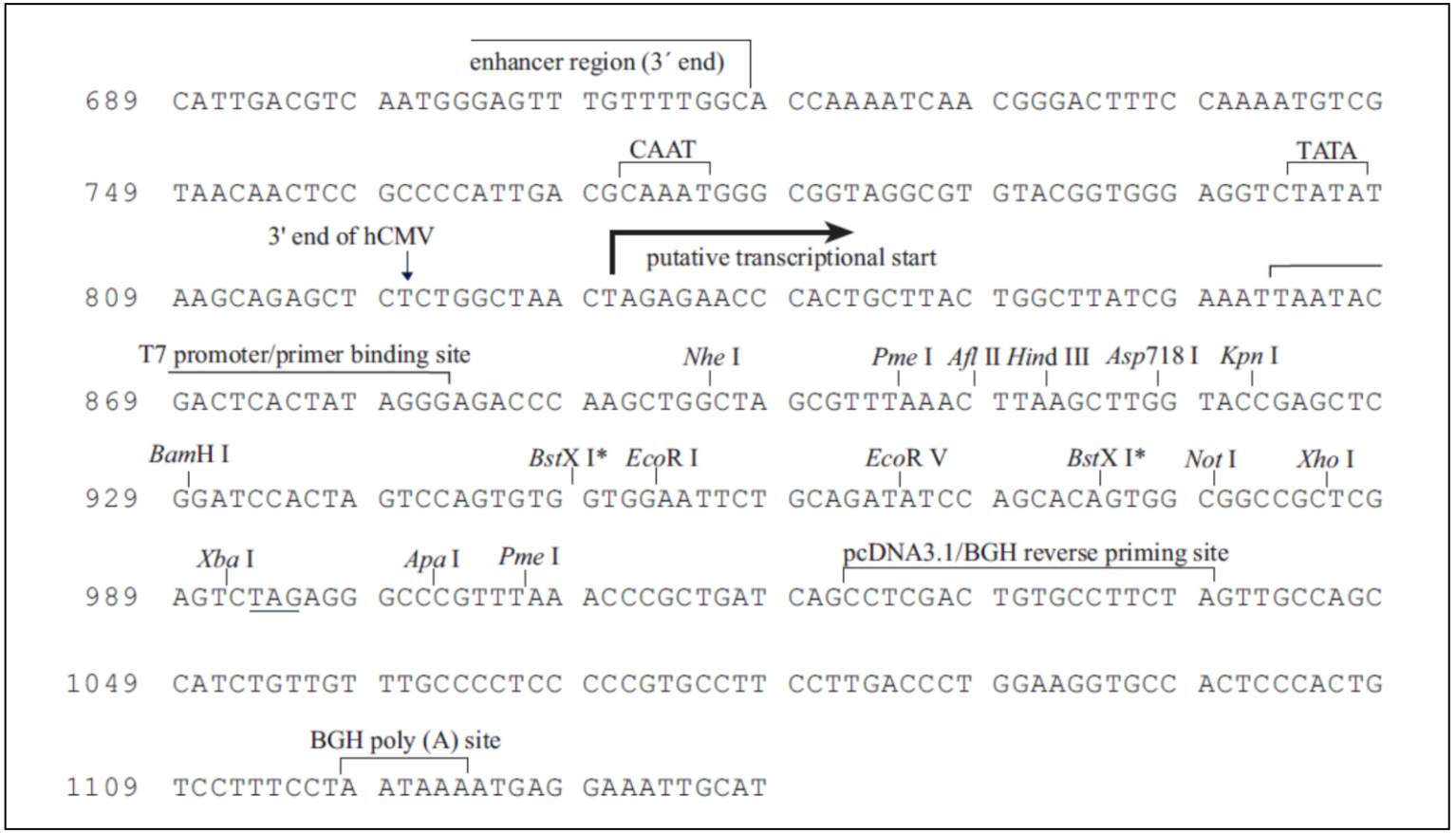


Figure 2.6: Multiple cloning site region of pCDNA3.1 (+) vector along with TATA box of CMV promoter and BGH polyadenylation site. HGDC and HGSC cDNAs were cloned into the *EcoRI* and *XhoI* sites of this vector.

2.1.2. Bacterial strains used for cloning and expression

DH5-α strain of *E.coli* (Novagen) was used for screening and maintenance of the recombinant plasmids. BL21 (DE3) pLys (S) strain of *E.coli* (Novagen) was used for the over-expression of recombinant protein.

2.1.3. Cell lines

The Human Lens Epithelial Cell line (HLE-3B) was used for checking the aggregation properties.

2.2. Methods

2.2.1. Cloning of wild type and mutant constructs

Human cadaveric eye lens was collected from the Ramayamma International Eye Bank, L.V. Prasad Eye Institute, after due ethical and scientific approval from the Institutional review board. Total RNA was isolated from the lens using Trizol reagent. The first strand was synthesized by Reverse Transcription Polymerase Chain Reaction (RT-PCR) using oligo-dT primer and Superscript III reverse transcriptase (Invitrogen). Human γD-crystallin (HGDC) and Human γS-crystallin (HGSC) cDNAs were amplified from the first strand using forward primer with *Nde1* restriction site and reverse primer with *HindIII* restriction site. The amplified WT HGDC and HGSC cDNAs were cloned into a *SmaI* digested pBSSK+ vector using T4-DNA ligase. The recombinant clones were confirmed by Polymerase Chain Reaction (PCR) and restriction digestion.

Wild type cDNAs were released from the pBSSK+ vector by restriction digestion with *Nde1* and *HindIII* restriction enzymes. The released cDNAs were ligated into the *Nde1* and *HindIII* sites of the pET-21a (+) vector using T4-DNA ligase. Mutant clones pET-21a γD P24T, pET-21a γD R77S, pET-21a γD A36P, pET-21a γD R140X, pET-21a γD G165fs, pET-21a γD L45PL54P, pET-21a γD Y134A were generated from the pET-21a γD WT template and pET-21a γS D26G and pET-21a γS V42M were generated from the pET-21a γS WT template by PCR-based site directed mutagenesis using Phusion DNA polymerase. The amplification conditions were as follows: an initial denaturation step at 98°C for 30 s, followed by 16 cycles of denaturation, annealing and extension at 98°C (10 s), 55°C (30 s), 72°C (3 min), respectively, with a final extension step at 72°C for 10 min. The PCR product was digested with *Dpn1* for 1 h and transformed into DH5-α and the plasmids were isolated.

His-tagged WT HGDC, R140X and G165fs mutants were amplified from the previously cloned pET-21a γD WT template using Phusion DNA polymerase using forward primer having *EcoR1* restriction site, and reverse primer having codons for His tag and *XhoI* restriction site and cloned into the *EcoR1* and *XhoI* sites of the pCDNA3.1 (+) vector by the above-mentioned method. His-tagged mutant clones, pCDNA3.1 (+) γD P24T, pCDNA3.1 (+) γD R77S, pCDNA3.1 (+) γD A36P, pCDNA3.1 (+) γD L45PL54P and pCDNA3.1 (+) γD Y134A were generated from the pCDNA3.1 (+) γD WT template by site-directed mutagenesis using Phusion DNA polymerase by the above mentioned method.

Table 2.1: List of primers used for cloning and sequencing of human HGDC, HGSC and their various mutants.

Clone	Primer sequence
pET-21a γD	F-5'TCC*CATATG*GGGAAGATCACCCTCTACGAG3'
	R-5'GC*AAGCTT*TCAGGAGAAATCTATGACTCT3'
pET-21a γD P24T& pCDNA3.1 (+) γD P24T	F-5'CAGCAGCGACCACACCAACCTGCAGCCCTACTTGAG3'
	R-5'GGCTGCAGGTTGGTGTGGTCGCTGCTGCATTCATAG3'
pET-21a γD R77S & pCDNA3.1 (+) γD R77S	F-5'GACTCGGTCAGCTCCTGCCGCCTCATCC3'
	R-5'GCGGCAGGAGCTGACCGAGTCGCTGAG3'
pET-21a γD A36P & pCDNA3.1 (+) γD A36P	F-5'CTGCAACTCGCCGCGCGTGGACAGCGGC 3'
	R-5'GTCCACGCGCGGCGAGTTGCAGCGGCTC 3'
pET-21a γD G165fs	F-5'CGAATGCCAGAGTGGCTCTCTGAGGAGAGTCATAGAT TTCTCC3'
	R-5'CTCTCCTCAGAGAGCCACTCTGGCATTCGTGGC3'
pET-21a γD L45PL54P & pCDNA3.1 (+) γD L45PL54P	F-5'GCGGCTGCTGGATGCCCTATGAGCAGCCCAA CTACTCGGGCCCCCGTACTTCCTGCGCCGCGG3'
	R-5'CGCAGGAAGTACTGGGGGCCCGAGTAGTTGGGCTGCT CATAGGGC ATCCAGCAGCCGCTGTCCAC3'
pET-21a γD Y134A & pCDNA3.1 (+) γD Y134A	F-5'CTGGGTCCTCGCCGAGCTGTCCAACTACCGAGGACG3'
	R-5'GTAGTTGGACAGCTCGGCGAGGACCCAGGAGCCCTC3'
pET-21a γS	F-5'GGGAGTTCCATATGTCTAAAACTGGAACC3'
	R-5'CCGGAATTCTTACTCCACAATGCG3'
pET-21a γS D26G	F-5'CTATGACTGTGATTGCGGCTGTGCAGATTTCCACACAT AC 3'
	R-5'GAAATCTGCACAGCCGCAATCACAGTCATAGCGAC3'
pET-21a γS V42M	F-5'GCTGCAACTCCATTAAAATGGAAGGAGGCACCTGGGCTG3'
	R-5'CCAGGTGCCTCCTTCCATTTTAATGGAGTTGCAGCGA CTTAGG3'
T7	F-5'TAATACGACTCACTATAGG3'
T7	R-5'TATGCTAGTTATTGCTCAG3'
BGH	R-5'TAGAAGGCACAGTCGAGG3'
pCDNA3.1 (+) γD	F-5'CG*GAATTC* ATGGGGAAGATCACCCTC3'
	R-5'CCG*CTCGAG*TTAATGATGATGATGATGATGGGAGAAA TCTATGACTCTCCTCAG3'
pCDNA3.1 (+) γD R140X	R-5'CCG*CTCGAG*TTAATGATGATGATGATGATGGTAGTTG GACAGCTC GTA GAGG3'
pCDNA3.1 (+) γD G165fs	R-5'CCG*CTCGAG*TTAATGATGATGATGATGATGGAGAGCC ACTCTGGCA TTC-3'

The sequences of the WT and mutant clones were verified by sequencing using an ABI 3130 Genetic Analyzer, with T7 forward,T7 reverse and BGH reverse primers. The primers used for cloning and sequencing are listed in Table 2.1.

2.2.2. Over-expression of recombinant proteins

The recombinant constructs pET-21a γD WT, pET-21a γS WT pET-21a γD P24T, pET-21a γD R77S, pET-21a γD A36P, pET-21a γD R140X, pET-21a γD G165fs, pET-21a γD L45PL54P, pET-21a γD Y134A, pET-21a γS D26G and pET-21a γS V42M, were transformed into *E.coli* BL21 (DE3) pLys (S) cells. A single colony containing the recombinant construct was picked, inoculated into 15 ml of Luria-Bertoni (LB) medium containing 50 µg/ml ampicillin and 34 µg/ml chloramphenicol and grown for 8 h by shaking at 225 rpm, 37°C. After 8 h, 10 ml of the culture was transferred into 1L of LB medium containing 50 µg/ml ampicillin and 34 µg/ml chloramphenicol. The cultures were grown at 37°C to an absorbance value of 0.6 at 600 nm. Protein synthesis was induced by the addition of IPTG to a final concentration of 1 mM and the cultures were grown for an additional 3.5 h. Cells were pelleted down from the 1L culture by centrifugation at 6000*g* for 10 min at 4°C.

The cell pellets were suspended in 40 ml of lysis buffer containing 50 mM Tris hydroxymethyl aminomethane chloride (Tris-Cl) (pH 7.3), 100 mM KCl, 1 mM Ethylene diamine tetra acetic acid (EDTA), 1 mM dithiothreitol (DTT), 1 mM phenyl methyl sulfonyl fluoride (PMSF) and 20 µg/ml aprotinin. In the case of γS

WT, D26G and V42M mutant pellets were suspended in 40 ml of 50 mM sodium acetate buffer (pH 4.75) containing 1 mM DTT and 1 mM PMSF. The cell suspension was extensively sonicated for 40 cycles (30s bursts of sonications, followed by 30s rest cycles) at 35% amplitude at 4°C using a high intensity Ultrasonic processor (Sonics Vibra Cell; Sonics & Materials Inc, Newton, MA).The cell lysate was centrifuged at 30,000*g* for 20 min at 4°C. The resulting supernatant and the pellet were checked for the presence of the recombinant protein on 14% Sodium dedecyl sulphate poly acrylamide gel electrophoresis (SDS-PAGE). Wild type γD, WT γS, P24T, R77S, D26G and V42M were predominantly found in the soluble fraction, whereas A36P, R140X, G165fs, L45PL54P and Y134A were found in the inclusion bodies. We over-expressed all these proteins at different concentrations of the inducer molecule IPTG (0.25, 0.5 and 1.0 mM), at various temperatures (18°C and 37°C) and various time points (2.5 h, 3.0 h and 3.5 h). Mutants A36P and Y134A of γD were largely found as inclusion bodies in all these conditions, though some amount of protein was found in the soluble fraction when they were over-expressed at 18°C, but the mutants G165fs, R140X and L45PL54P were found in the insoluble fraction in all these conditions, hence we had to purify the proteins from inclusion bodies. The expression profiles of R140X and Y134A at different conditions are shown in Figures 2.7 and 2.8.

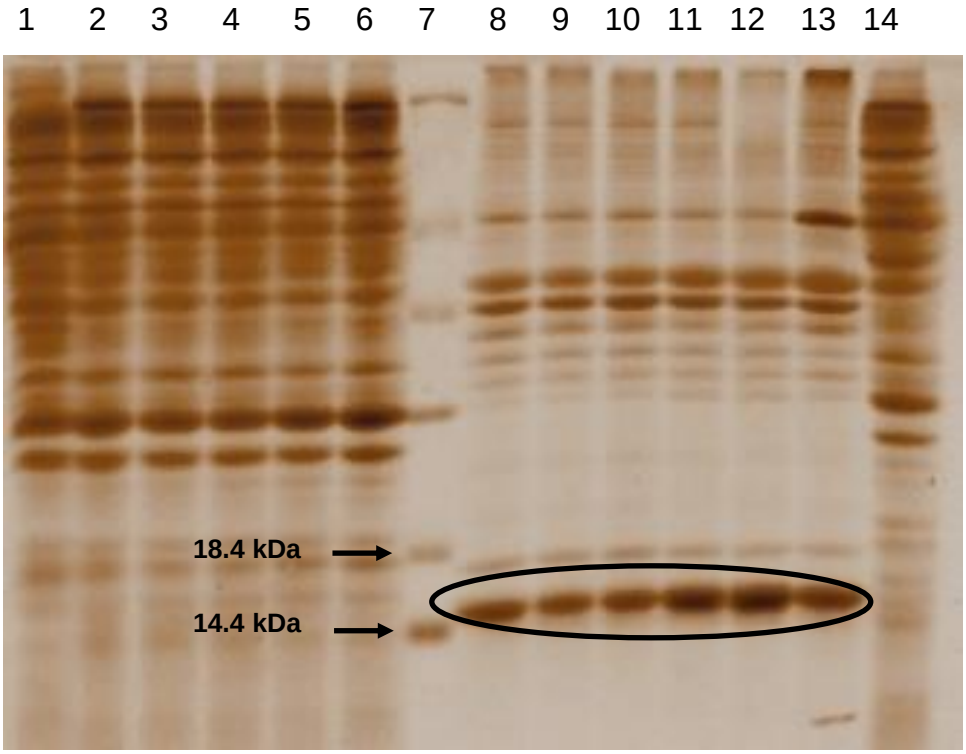


Figure 2.7: Expression profile of R140X γD at different IPTG concentrations (0.25, 0.5 and 1.0 mM), at various temperatures (18°C and 37°C) and various time points (2.5 h, 3.0 h and 3.5 h). Lane 1: uninduced; Lane 2-6 and 14 are soluble fractions; Lane 2: 0.25 mM IPTG, 37°C, 3.5 h; Lane 3: 0.5 mM IPTG, 37°C, 3.5 h; Lane 4: 1.0 mM IPTG, 37°C, 3.5 h; Lane 5: 1.0 mM IPTG, 37°C, 2.5 h; Lane 6: 1.0 mM IPTG, 37°C, 3.0 h; Lane 14: 1.0 mM IPTG, 18°C, 10.0 h; Lane 7: marker; Lane 8-13 insoluble fractions; Lane 8: 0.25 mM IPTG, 37°C, 3.5 h; Lane 9: 0.5 mM IPTG, 37°C, 3.5 h; Lane 10: 1.0 mM IPTG, 37°C, 3.5 h; Lane 11: 1.0 mM IPTG, 37°C, 2.5 h; Lane 12: 1.0 mM IPTG, 37°C, 3.0 h; Lane 13: 1.0 mM IPTG, 18°C, 10.0 h. Highlighted circle shows the over-expressed R140X protein found in the insoluble fraction.

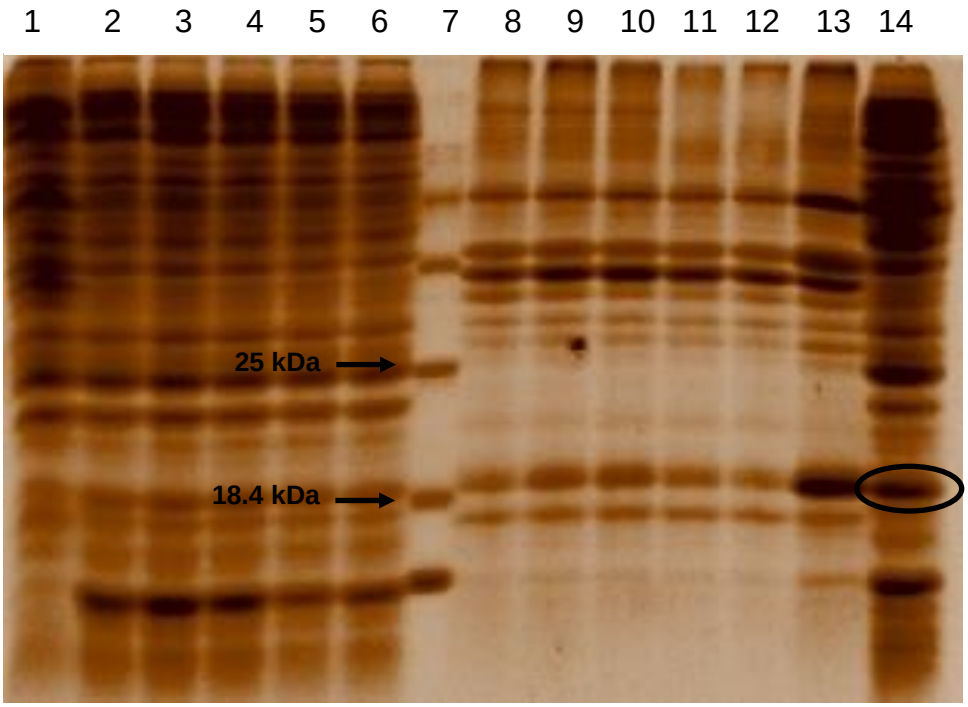


Figure 2.8: Expression profile of Y134A γD at different IPTG concentrations (0.25, 0.5 and 1.0 mM), at various temperatures (18°C and 37°C) and various time points (2.5 h, 3.0 h and 3.5 h). Lane 1: uninduced; Lane 2-6 and 14 are soluble fractions; Lane 2: 0.25 mM IPTG, 37°C, 3.5 h; Lane 3: 0.5 mM IPTG, 37°C, 3.5 h; Lane 4: 1.0 mM IPTG, 37°C, 3.5 h; Lane 5: 1.0 mM IPTG, 37°C, 2.5 h; Lane 6: 1.0 mM IPTG, 37°C, 3.0 h; Lane 14: 1.0 mM IPTG, 18°C, 10.0 h; Lane 7: marker; Lane 8-13 insoluble fractions; Lane 8: 0.25 mM IPTG, 37°C, 3.5 h; Lane 9: 0.5 mM IPTG, 37°C, 3.5 h; Lane 10: 1.0 mM IPTG, 37°C, 3.5 h; Lane 11: 1.0 mM IPTG, 37°C, 2.5 h; Lane 12: 1.0 mM IPTG, 37°C, 3.0 h; Lane 13: 1.0 mM IPTG, 18°C, 10.0 h. Highlighted circle shows the recombinant Y134A protein over-expressed at 18°C, found in the soluble fraction.

2.2.3. Purification of recombinant proteins

2.2.3.1. Purification of γD-WT, P24T, R77S, Y134A and A36P mutant proteins

The supernatant was subjected to ammonium sulfate fractionation at 30% concentration. Ammonium sulfate was added pinch by pinch, the solution incubated on ice for 2 h and centrifuged at 30,000*g* for 20 min at 4°C. The pellet and supernatant were checked for the presence of the recombinant protein using 14% SDS-PAGE. The proteins γD-WT, P24T and R77S were predominantly found in the supernatant. The supernatant was loaded onto a Phenyl-Sepharose (P-Sepharose) column equilibrated with 50 mM Tris-Cl buffer (pH 7.3) containing 30% of ammonium sulfate and eluted using a gradient of 0-30% of ammonium sulfate in 50 mM Tris-Cl buffer (pH 7.3). Individual fractions were checked on 14% SDS-PAGE. Fractions containing the required protein were pooled, concentrated using an Amicon stirred ultrafiltration cell with 3 kDa cut-off membrane and the concentrated protein was further purified to homogeneity by gel filtration chromatography using a Sephadex G-75 column.

In the case of A36P and Y134A, the supernatant obtained after lysis was directly loaded onto a G-75 column and fractions containing protein of interest were further purified using a Sulfo Propyl-Sepharose (SP-Sepharose) column.

2.2.3.2. Purification of γS-WT, D26G and V42M proteins

The supernatant was chromatographed using a SP-Sepharose ion-exchange column. The column was equilibrated with 50 mM sodium acetate buffer (pH 4.75) and the supernatant was loaded onto the column and eluted using a salt gradient of 0-1 M KCl. The fractions containing the required protein were pooled, concentrated using an Amicon stirred ultrafiltration cell with a 3 kDa cut-off membrane and further purified to homogeneity on a Sephadex G-75 column.

2.2.3.3. Purification and refolding of R140X, G165fs and L45PL54P mutant proteins

The insoluble pellet in each case was resuspended in 15 ml of washing buffer (25 mM Tris-Cl (pH 8.0), 100 mM KCl, 0.1% Triton X100, 1.5 M urea and 2 mM DTT), homogenized and washed 5-6 times by centrifugation at 30,000*g* for 20 min at 4°C until a clear supernatant was obtained. The pellet was then washed with 15 ml of wash buffer lacking urea and Triton X100 and dissolved in 20 ml of buffer containing 50 mM Tris-Cl (pH 7.3) and 7.5 M urea and stirred continuously to solubilize the protein and incubated at 4°C for 12 h.

We were not successful in refolding the denatured protein by On-column refolding using Quaternary ammonium Sepharose (Q-Sepharose) or SP-Sepharose matrix and rapid dilution methods; hence we adopted the stepwise dialysis approach. The protein mixture was dialyzed initially against 50 mM Tris-

Cl buffer (pH 7.3) and 7.5 M urea under continuous stirring at 4°C. 50 mM Tris-Cl buffer (pH 7.3) was added drop by drop till the protein completely loses urea.

The purity and specificity of the proteins was assessed by SDS-PAGE, western blotting and also by Matrix Assisted Laser Desorption Ionization (MALDI) mass spectral analysis at the Centre for Cellular and Molecular Biology (CCMB), Hyderabad, thanks to the help and advice of Dr. Suman Thakur there. The concentration of each protein was measured by A_{280} in 4.5 M GuHCl, using its molar extinction coefficient (εM) calculated from Expasy http://web.expasy.org/cgi-bin/protparam/protparam). The estimated mass numbers of each protein obtained using MALDI mass spectral analysis were: WT: observed 20597 (expected 20610), P24T: 20615 (expected 20614), R77S: 20532 (expected 20540), A36P: 20580 (expected 20636), R140X: 16448 (expected 16453), G165fs: 19659 (expected 19662), Y134A: 20488 (expected 20518) and L45PL54P: 20550 (expected 20578).

We thank Mr. Srinivasu K for helping us in cloning, expressing and purifying D26G protein.

2.2.4. Spectroscopic analysis

Circular dichroism (CD) spectra were recorded using a dichroigraph instrument (J-810; Jasco, Easton, MD) at room temperature (27°C). Far-UV CD spectra were recorded in the region between 250–193 nm with 2 mm path length quartz cells and the near-UV CD spectra in 320–250 nm region were recorded

with 1 cm path length quartz cells, CD was measured for every 0.5 nm between 193-320 nm with 2 s response time at 100 nm/s speed. At least three scans of each spectrum were averaged, and baselines of the buffer alone were subtracted. The protein concentration used for determining the far-UV spectra was 6 μM (0.125 mg/ml), 10 μM (0.2 mg/ml) and 12μM (0.250 mg/ml) in 50 mM Tris-Cl and 10 mM sodium phosphate buffer (pH 7.3) and for the near-UV spectra, it was 24 μM (0.500 mg/ml) and 40 μM (0.8 mg/ml) in 50 mM Tris-Cl and 100 mM sodium phosphate buffer (pH 7.3). We thank CCMB for allowing us the use of the CD spectrometer.

Intrinsic fluorescence spectra were recorded at room temperature (27°C) using a fluorescence spectrophotometer (F-2500; Hitachi, Yokohama, Japan) and the spectra were recorded in the range 300 to 400 nm using an excitation wavelength of 295 nm, with 2.5 nm excitation and emission slits. The protein concentrations used were 6 μM (0.125 mg/ml), 10 μM (0.2 mg/ml) in 50 mM Tris-Cl and 100 mM sodium phosphate buffer (pH 7.3). At least three scans of each spectrum were averaged and baselines of the buffer alone were subtracted. Fluorescence quenching experiments were done using KI and acrylamide as quenchers, and the results were analyzed using the Stern-Volmer approach, following Augusteyn et al (180).

Extrinsic fluorescence spectra of proteins were recorded at room temperature (27°C) using two surface hydrophobicity probes, namely 4,4'-dianilino-1,1'-binaphthyl-5,5'-disulfonate (bis-ANS) (181) and 9-diethylamino-5H-

benzo[alpha]phenoxazin-5-one (Nile Red) (182). With bis-ANS, spectra were recorded in the range of 400 to 600 nm, using an excitation wavelength of 390 nm with 2.5 nm excitation and emission slits. With Nile Red, the excitation was at 540 nm and the emission recorded between 570 and 700 nm, using 10 nm slits. The protein concentrations used in each case were 6 μM (0.125 mg/ml) and 5 μM (0.1 mg/ml) in 50 mM Tris-Cl and 100 mM sodium phosphate buffer (pH 7.3). Stock solutions of bis-ANS and Nile Red were prepared in methanol and the final alcohol concentration was maintained below 7% v/v when the reagents were mixed with the proteins. Concentrations of bis-ANS and Nile Red were measured using extinction coefficients of 16.8 mM^{-1} cm^{-1} at 385 nm and 45 mM^{-1} cm^{-1} at 552 nm respectively.

Formation of amyloid-type fibrils was monitored at room temperature (27°C) using the probe Thioflavin-T (183). Spectra were recorded in the range of 470 to 570 nm, using an excitation wavelength of 444 nm with 10 nm excitation and emission slits. Protein concentrations used in each case were 6 μM (0.125 mg/ml) and 5 μM (0.1 mg/ml) in 50 mM Tris-Cl and 100 mM sodium phosphate buffer (pH 7.3). Baselines of the buffer alone were subtracted.

Equilibrium unfolding and refolding experiments were performed at ambient temperature (27°C) by diluting the purified proteins to 0.2 mg/ml in a series of 46 different concentration solutions in the range of 0 to 4.5 M guanidine hydrochloride (GuHCl) with increasing intervals of 0.05 or 0.1M GuHCl in a buffer containing 100 mM sodium phosphate, 1 mM EDTA and 5 mM DTT. Samples

were incubated at 37°C for 16 h. The procedure used here was the same as that of Mills et al. (49). Fluorescence emission spectra were recorded for each unfolding sample using a spectrofluorimeter as described above. Data were analyzed by plotting the concentration of GuHCl for each sample versus the ratio of fluorescence intensities at 360 and 320 nm. The ratio of fluorescence intensities at these wavelengths was chosen for the analysis in order to simultaneously monitor changes in native and unfolded maxima. Equilibrium unfolding data was analyzed to determine transition midpoints and $\Delta G°$ by fitting to the two-state model of Greene and Pace (184), or the three-state model of Clark et al. (185), using the Graphpad prism software. The model that best fit the data was selected based on a random distribution of residuals. Transition midpoints and ΔG° were calculated for all transitions from these fits. In all fluorescence experiments, the response time used was 0.08 s, scan speed 60 nm/s and the Photomultiplier tube (PMT) voltage was below 400 V.

Differential scanning calorimetry (DSC) was done at CCMB, Hyderabad using a VP-DSC Micro Calorimeter instrument (VP-DSC; Piscataway, NJ, USA), with the protein concentration of 0.5 mg/ml in 50 mM Tris-Cl buffer (pH 7.3), in the range 20- 95°C, at a speed of 1°C per minute. The samples and references were degassed immediately before use. C_p (kcal/mole/°C) and C_p (kcal/°C) was plotted against temperature. Micro Cal LLC DSC software was used for data acquisition and analysis and data were corrected for buffer baseline prior to concentration normalization.

Time-dependent light scattering measurements of both WT and V42M were done by monitoring the protein solution turbidity at 600 nm, using the spectrofluorimeter mentioned above, with 2.5 nm or 5 nm excitation and emission slits, at the constant temperature of 61°C with the protein concentration of 0.1 mg/ml in 50 mM Tris-Cl buffer (pH 7.3).

2.2.5. Cell culture, transfections and Immunoflouroscence

The human lens epithelial cell line HLE-3B was cultured in Complete Medium consisting of Dulbecco's Modified Eagle Medium (DMEM) (Sigma) supplemented with 20% fetal bovine serum (FBS) (HyClone) containing antibiotics in a 5% humidified CO_2 incubator at 37°C. At 12 h prior to transfection, 1,50,000 cells were seeded onto a 18 mm coverslip in a six-well culture plate and incubated in a 5% humidified CO_2 incubator at 37°C. The Complete Medium was removed after 12 h and replaced with 1 ml of antibiotic-free medium and the cells were transfected with the recombinant constructs using Fugene HD (Promega) at 1:6 ratio (1 µg vector/6 µl fugene). After incubation for 4 h, the antibiotic-free medium was replaced with 2 ml of Complete Medium and incubation was continued up to 24 h for imaging. For cell culture, Immunofluorescence and Imaging we followed the same procedure as described in our earlier paper (172).

After incubation, the transfected cells were washed with phosphate buffered saline (PBS) and fixed with absolute ice-cold methanol for 3 min. The cells were then washed thrice with PBS, blocked with 2.5% Bovine serum

albumin (BSA) for one hour and incubated with anti-His antibody (1:200 dilution) raised in mouse for two hours, washed three times with PBS, and incubated with Fluorescein-isothiocyanate (FITC)-conjugated anti-mouse antibody (1:500 dilution). Excess unbound antibody was removed by washing the coverslips thrice with PBS, and they were stained with propidium iodide (PI) for 2 min. Excess stain was removed by subsequent PBS washes. The coverslips with the cells were then mounted using 50% glycerol in PBS and images collected using a laser scanning confocal microscope (LSM510; Carl Zeiss, Jena, Germany). The excitation laser used was 483 nm. The emission of green fluorescence was collected using the 505 to 530 nm band-pass filter and that of red fluorescence was collected using the 585 to 615nm band-pass filter.

2.2.6. Molecular modeling and dynamics simulations

This work was done by Dr. Sushil Chandani of Novarus Discoveries Pvt Ltd, Hyderabad, with whom we have collaborated in this part of the study. While the detailed structure in the crystallin and solution states is already available for HGDC, similar information for HGSC is also available hence we have used the solved structure of the murine protein (PDB ID 2A5M), which shares 89.9% identity with the human molecule, was used as a template for molecular modeling techniques. The Swiss-Model workspace was used for this (186). Structures obtained were soaked in ~1200 molecules of water, and the charge of the ensemble adjusted to 0 using sodium ions. One cycle of minimization (200 iterations, steepest descent) was performed with the protein restrained, followed

by another cycle (500 iterations, steepest descent) without any restraint. The structure was heated to 300 K over 40 ps (time step 1 fs), and finally minimized for 200 iterations using the adopted basis Newton-Ralphson method. These protocols were implemented in the CHARMM simulation environment, version c33b2 (187). Structures were visualized using an Accelrys Discovery Studio system.

The modeling and dynamics of HGSC and its mutants D26G and V42M were done by the following approaches (188, 189). The molecular dynamics simulations were performed using Gromacs 4.5.5, developed by the Berendsen group (190), and run on Apple Computer's operating system OS 10.7. The selection of suitable rotamers of D26, N54 and R84 amino acids were performed in the Discovery Suite (Accelrys Inc., San Diego, CA.), as was the generation of the mutant G26. The Charmm force field (187) terms were used throughout. Visualization of structures was done with Insight II and Discovery (Accelrys Inc. San Diego, CA.). Initial structures were prepared for simulations by soaking in ~21000 water molecules in a dodecahedral box, the net charge of the system being set to zero by the addition of counter ions. After a 2000-iteration minimization, of the ensemble, the water molecules were equilibrated for 200 ps with harmonic constraints on the protein being gradually released. Following this molecular dynamics simulations were continued for 2 ns at 300 K and with periodic boundary conditions.

Results

3.1-3.6. Solubility and conformational features in solution

Substantial differences were seen in the solubility of the various mutants. WT Human γD-crystallin (HGDC) is highly soluble in water (450 mg/ml) and its mutants display lower solubility; P24T and R77S are quite soluble up to 30 mg/ml (164) and 320 mg/ml (191) respectively. A36P was found to be soluble only up to 5 mg/ml and precipitates beyond that. Several others such as R140X, G165fs, L45PL54P and Y134A are far less soluble, and with solubility less than 200 µg/ml. In fact, these sparingly soluble molecules came out as inclusion body when they were expressed. In the case of γS-crystallins, it was possible to recover the human γS-crystallin (HGSC) WT, mutants D26G and V42M from the soluble fractions during the purification and isolation steps. However, upon standing at 4°C, mutant V42M starts precipitating, unlike the D26G mutant and WT. This tendency to precipitate upon standing at 4°C was also seen in the HGDC mutants A36P, R140X, G165fs, L45PL54P and Y134A.

Solubility is seen to be decreased at least 100 times in A36P, R140X, G165fs, L45PL54P and Y134A, but not so badly affected in the case of P24T, R77S and D26G mutants.

We also studied the mutant D26G in detail (published) and this part of the work was done along with our colleague Mr. Srinivasu K of our lab and some of those results (Figure 3.4) we report here for better comparison.

To determine how the mutation affects the overall conformation of the molecule, we examined all the mutants using Circular Dichroism (CD) and fluorescence spectroscopy. CD is a well-established form of spectroscopy that allows us to investigate the manner in which a protein chain is folded in its backbone conformation, i.e., the secondary structural features of the chain.

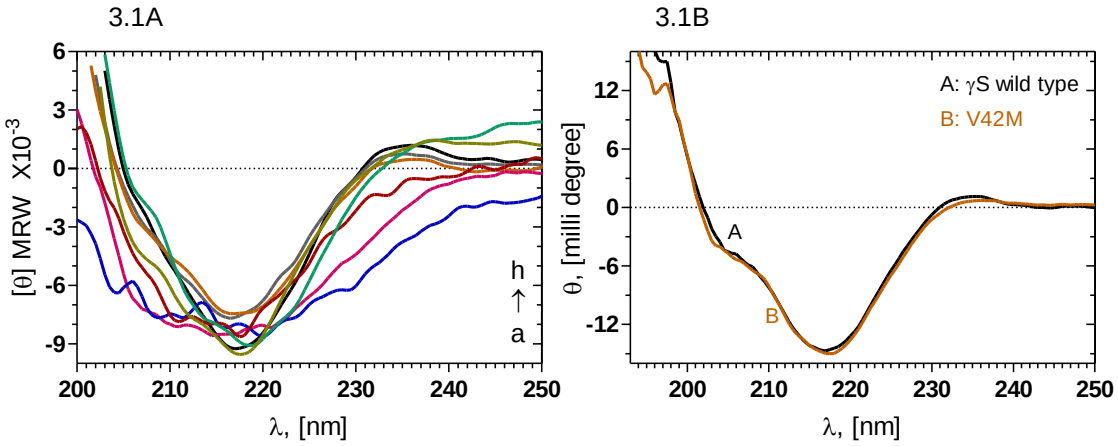


Figure 3.1: Secondary structural features of HGDC, HGSC and its mutants. **Figure 3.1A:** θ: Mean residue molar ellipticities, in milli degrees, of A (Blue): G165fs; B (Violet): L45PL54P; C (Maroon): R140X; D (Brown): R77S; E (Grey): P24T; F (Black): γD-WT; G (Olive): A36P; and H (Green): Y134A. Protein concentration in the case of γD-WT, P24T, R77S, and A36P was 12 μM (0.250 mg/ml) and in the case of R140X, G165fs, L45PL54P and Y134A it was 6 μM (0.125 mg/ml) in 50 mM Tris-Cl buffer (pH 7.3). MRW = mean residue molecular weight, taken as 110 Da. **Figure3.1B:** θ: Elipticity in milli degrees, of A (Black): γS-WT; B (Brown):V42M. Protein concentration in each of the case was 10 μM (0.2 mg/ml) in 10 mM sodium phosphate buffer (pH 7.3). The cell path length was 2 mm and all spectra were recorded at room temperature (27°C), corrected for background buffer signal and each reported spectrum is an average of 3 independent runs.

The backbone conformations of γ-crystallins are rich in β-pleated sheets, with each monomer composed of 16 β-pleated sheet segments arranged in antiparallel fashion and these β-pleated sheets are expected to show a

prominent characteristic minimum at 218 nm band in their far-UV CD spectra (192). Figure 3.1A compares the far-UV CD spectra of the WT and the various mutants of HGDC. A strong 218 negative peak is seen in all cases, with minor changes between the spectra of the mutants in comparison to that of the WT, suggesting that the overall basic chain conformation is not altered in any significant manner upon mutation. Despite the difference in intensity, the characteristic minimum of all proteins was the same and lying around 218 nm. The discrepancy in intensity may reflect disturbances of the domain interface or slight differences in solution or other structural rearrangements but the overall β-sheet chain conformation or secondary structure is not disturbed in any great manner.

Turning to HGSC, Figures 3.1B&3.4A compares the far-UV CD spectra of WT HGSC, V42M and D26G; all the spectra are superimposable. The negative band at 218 nm, the shoulder at around 206 nm and the positive peak at 195 nm are very similar in these cases showing backbone conformation is not disturbed in any manner upon mutation.

The sidechain packing or the tertiary structural feature was monitored using near-UV CD and fluorescence emission spectroscopy. Near-UV CD spectra (Figure 3.2A, 3.2B and 3.4B) reveal minor changes in the microenvironment around the aromatic residues of the mutants in comparison to the WT. However, changes in this region of the CD spectrum are not as informative; indeed it is the preferred practice to study the intrinsic fluorescence

spectral features of the aromatic residues (trp, tyr, phe) rather than the near-UV CD.

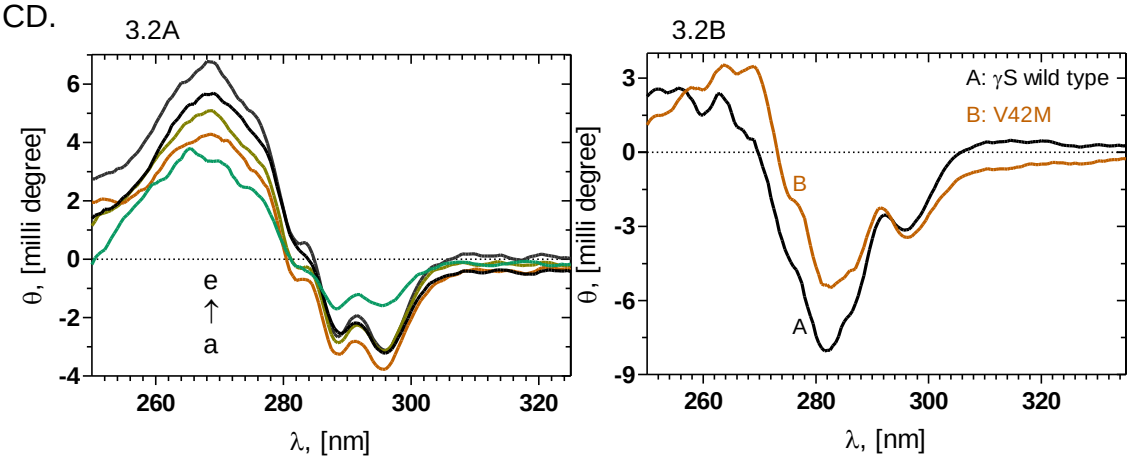


Figure 3.2: Tertiary structural features of HGDC, HGSC and its mutants. **Figure 3.2A:** θ: Ellipticities, in milli degrees, of A (Green): Y134A; B (Brown): R77S; C (Olive): A36P; D (Black): γD-WT; and E (Grey): P24T. Protein concentration in each case was 24 μM (0.5 mg/ml) in 50 mM Tris-Cl buffer (pH 7.3). **Figure 3.2B:** A (Black): γS-WT; B (Brown): V42M. Protein concentrations was 40 μM (0.8 mg/ml) in 100 mM sodium phosphate buffer, (pH 7.3), cell path length 10 mm and the other conditions of measurement were the same as above.

HGDC has 4 trp, 6 phe and 12 tyr residues and HGSC has 4 trp, 9 phe and 14 tyr residues in their sequence. Of these, the trp residues exhibit the most notable fluorescence quantum yields, and hence most often studied by exciting at 295 nm and monitoring the emission in the 320-350 nm region (193). All the four trp residues are well buried at positions 43, 69, 131, 157 in HGDC (194, 195) and 47, 73,137 and 163 in HGSC (196) in the hydrophobic core of the native protein and thus they emit in the 320-327 nm region. When structural perturbation occurs due to mutation, making one or more tryptophan residues accessible to the surface, the emission wavelength maximum is expected to be red-shifted and the intensity increased.

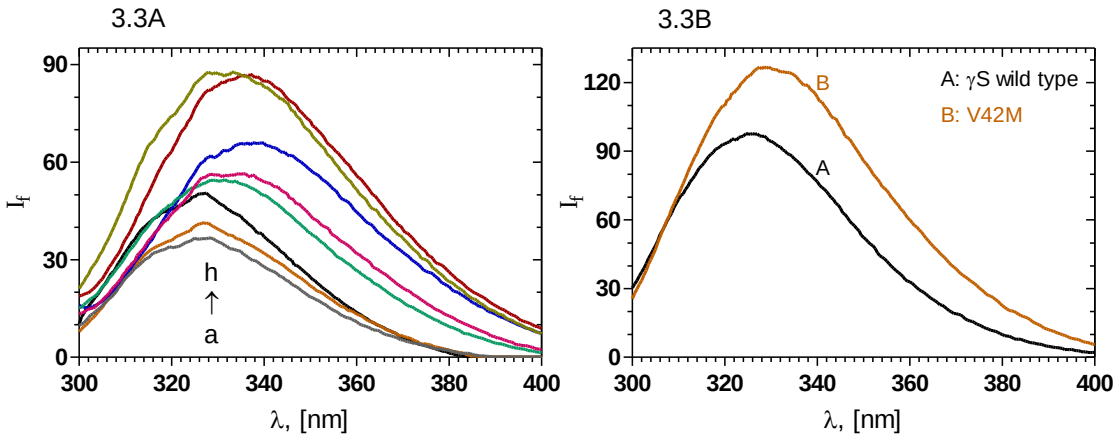


Figure 3.3: Intrinsic fluorescence of HGDC, HGSC and its mutants. **Figure 3.3A:** I_f: Emission intensity in arbitrary units. A (Grey): P24T; B (Brown): R77S; C (Black): γD-WT; D (Green): Y134A; E (Violet): L45PL54P; F (Blue): G165fs; G (Maroon): R140X; and H (Olive): A36P. Protein concentration in each case was 6 µM (0.125 mg/ml) in 50 mM Tris-Cl (pH 7.3). **Figure 3.3B:** A (Black): γS-WT; B (Brown): V42M. The protein concentrations used was 10 µM (0.2 mg/ml) in 100 mM sodium phosphate buffer, (pH 7.3). λ_{exc}: 295 nm, cell path length 3 mm, excitation and emission slits 2.5 nm and recorded at room temperature (27°C). Each reported spectrum is an average of 3 independent runs.

Figure 3.3A monitors the tertiary structural features of the proteins around their aromatic side chains through the intrinsic fluorescence spectra in the 300-400 nm region. The emission maximum of WT HGDC molecule is seen to occur at 327.5 nm with a relative intensity of emission (I_f) of 50.5 arbitrary units. P24T and R77S are seen to emit at 328.0 (I_f 37) and 327.0 nm (I_f 41) respectively. In comparison, the mutant A36P displays a slightly red-shifted intrinsic emission at 333 nm (I_f 87), while Y134A emits at 333 nm (I_f 55), and R140X, G165fs and the N-terminal domain double mutant L45PL54P emit around 337 nm with intensities of 85, 65 and 55 respectively. This suggests that the aromatic side chains in these mutants are somewhat more exposed to the solvent than in the WT, or in

the mutants P24T and R77S. This interpretation is supported by experiments using the ionic quencher KI which quenches the emission of only those fluorophores residues which are surface-accessible and not those buried in the apolar interior (180). The Stern-Volmer quenching constants of the KI quenching are seen to vary in the order A36P > Y134A >> R77S > P24T > WT (HGDC). We were not able to do the KI quenching for the mutants R140X, G165fs and L45PL54P due to precipitation (salting out) of these mutant proteins upon KI addition.

When we turn to γS-crystallin, we note that the WT HGSC emits at 326 nm with an intensity of 98 arbitrary units (Figure 3.3B), while the mutant V42M emits at 327.5 nm with an enhanced intensity of 127 units. This 25% enhancement in the emission intensity in V42M suggests a somewhat higher exposure of the aromatic residue to the solvent. Figure 3.4C shows the intrinsic fluorescence of the WT HGSC and mutant D26G. WT HGSC and mutant D26G (5 µM protein was used) both emit at 326 nm with an intensity of 44 units suggesting aromatic residues are buried and not exposed to the solvent.

In addition, we found that the ionic quencher KI had little effect on the WT HGSC molecule with a Stern-Volmer value of K_{SV} = 0.12, while that for D26G was 0.21 but had a noticeable effect on the mutant V42M with a Stern-Volmer value of K_{SV} = 0.88 (180). The neutral quencher acrylamide too showed some slight difference between the WT HGSC (K_{SV} = 1.12), D26G (K_{SV} = 1.20) and V42M (K_{SV} = 1.93). These results suggest that tryptophan and its

microenvironment is significantly changed in A36P, R140X, G165fs, L45P54P, Y134A and V42M but not in P24T, R77S, D26G in comparison to the wild types.

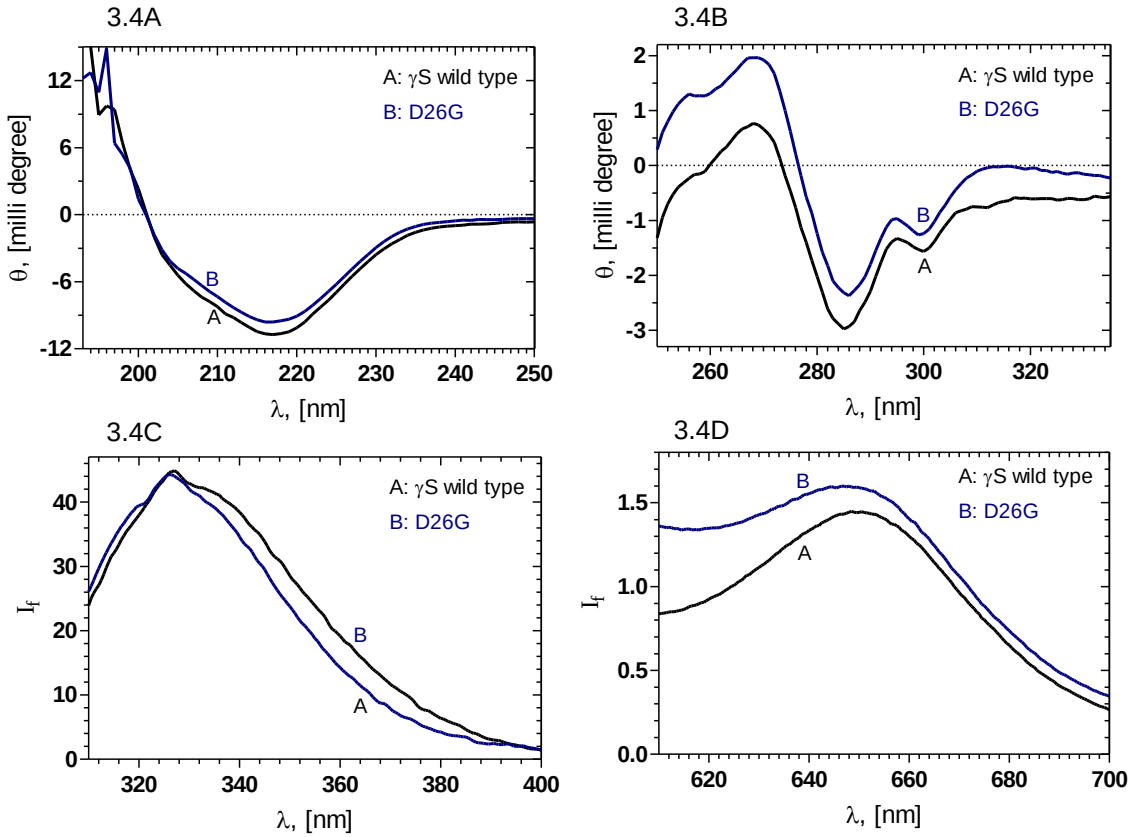


Figure 3.4A&B: Secondary and tertiary structural features of the WT HGSC and its mutant D26G. (Black): γS-WT; (Blue): D26G. **Figure 3.4C: Intrinsic fluorescence of the WT HGSC and D26G**. **Figure 3.4D: Extrinsic fluorescence of the WT HGSC and D26G with surface probe Nile Red.** The protein concentrations used for D and E were 5 µM (0.1 mg/ml) in 50 mM Tris-Cl buffer (pH 7.3), cell path length 3 mm, excitation and emission slits are 2.5 nm for 3.4C, 10 nm for 3.4D and spectra were recorded at room temperature (27°C).

The extent of surface exposure is better monitored using the extrinsic reporters bis-ANS (181) and Nile Red (182), which display enhanced emission intensity upon binding to hydrophobic surfaces.

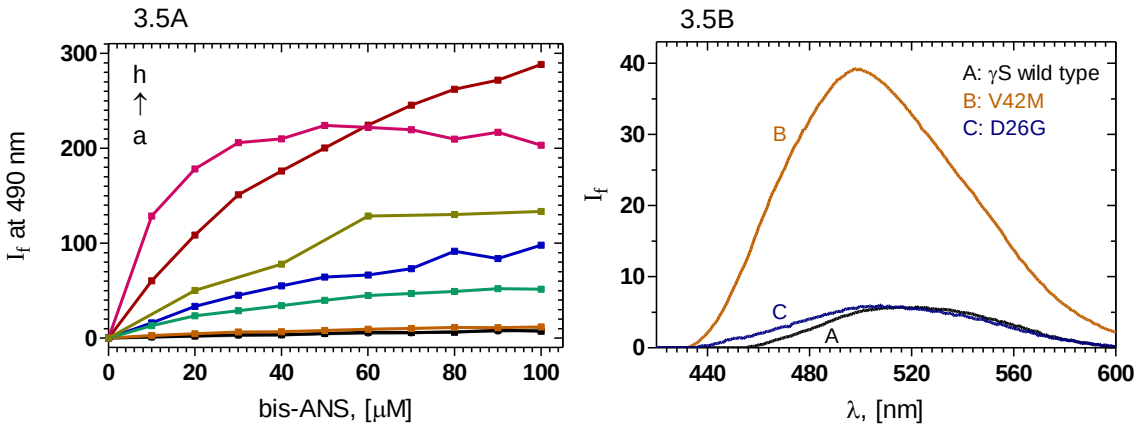


Figure 3.5: Surface exposure of non polar residues in the proteins, monitored using bis-ANS as the extrinsic probe. **Figure 3.5A:** A (Black): γD-WT; B (Grey): P24T; C (Brown): R77S; D (Green): Y134A; E (Blue): G165fs; F (Olive): A36P; G (Violet): L45PL54P; and H (Maroon): R140X. I_f at 490 nm of the probe was measured as a function of its increasing concentration. Protein concentration in each case was 6 μM (0.125 mg/ml) in 50 mM Tris-Cl buffer (pH 7.3). **Figure 3.5B:** A (Black): γS-WT; B (Brown): V42M; C (Blue): D26G; Protein concentration in each case was 5 μM (0.1 mg/ml) in 100 mM sodium phosphate buffer (pH 7.3). λ_{exc}: 390 nm, cell path length 3 mm, excitation and emission slits 2.5 nm and spectra were recorded at room temperature (27°C).

Figure 3.5A shows that WT HGDC and the mutants P24T and R77S generate little or no extrinsic fluorescence with the increasing concentration of bis-ANS but the other mutants display higher fluorescence in the order: R140X (I_f 262.2) > L45PL54P (I_f 209.6) > A36P (I_f 130.2) > G165fs (I_f 91.4) > >Y134A (I_f 49.1) > R77S (I_f 11.2) > P24T (I_f 6.8) > WT (HGDC) (I_f 6.2) at 80 μM of dye.

In the case of the γS-crystallin, the mutant V42M exhibits a 24 nm blue-shift (Figure 3.5B) in its emission maximum, and an eight-fold increase (I_f 39.3) in its emission intensity compared to the WT HGSC (I_f 5.7), mutant D26G shows a 8 nm blue-shift with same intensity (I_f 6.0) compared to WT HGSC when it bound to 60 μM bis-ANS, indicating that V42M has a higher degree of surface hydrophobicity. These results suggests that in mutants A36P, R140X, G165fs, L45PL54P, Y134A and V42M display a greater amount of surface exposure of otherwise buried nonpolar residues, compared to the mutants P24T, R77S, D26G and the respective wild types.

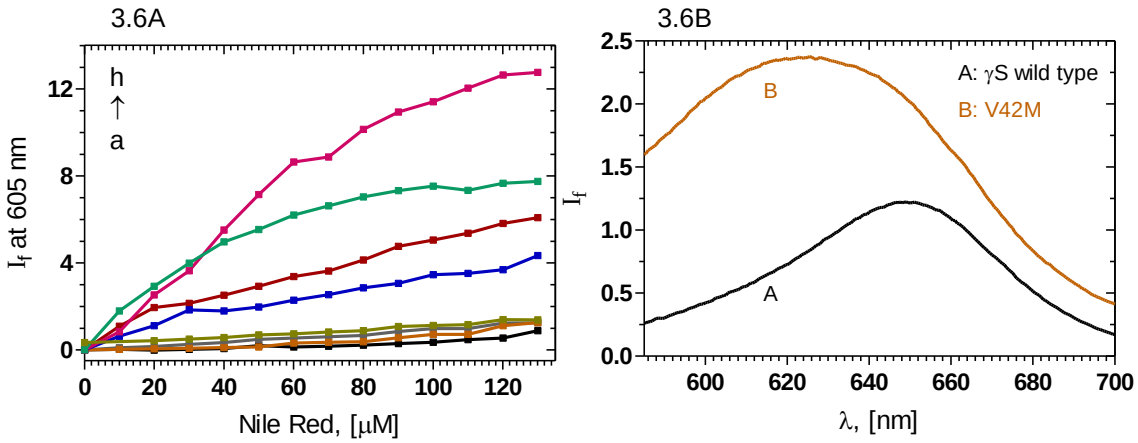


Figure 3.6: Aggregation tendencies of the proteins, estimated using Nile Red as the extrinsic probe. Figure 3.6A: A (Black): γD-WT; B (Brown): R77S; C (Grey): P24T; D (Olive): A36P; E (Blue): G165fs; F (Maroon): R140X; G (Green): Y134A; and H (Violet): L45PL54P. I_f at 605 nm of the probe was measured as a function of its increasing concentration. Protein concentration in each case was 6 μM (0.125 mg/ml) in 50 mM Tris-Cl buffer (pH 7.3). **Figure 3.6B:** A (Black): γS-WT; B (Brown): V42M; λ_{exc}: 540 nm, cell path length 3 mm, excitation emission slits 10 nm, and spectra were recorded at room temperature (27°C).

We have used another probe, the neutral dye Nile Red, which is also known to be a sensitive detector of protein self-aggregation. Figure 3.6A shows

that in the case of HGDC and its mutants, the extrinsic fluorescence intensity of the added surface probe Nile Red at 605 nm varies in the order L45PL54P (I_f 10.1) > Y134A (I_f 7.0) > R140X (I_f 4.1) > G165fs (I_f 2.8) > A36P (I_f 0.8) > P24T (I_f 0.6) > R77S (I_f 0.3) > WT (HGDC) (I_f 0.2) at 80 µM of dye.

Figure 3.6B shows a blue-shift in the emission band maximum of Nile Red when bound to the γS mutant V42M (625.5 nm cf. 651 nm with WT) and over twofold increase (I_f 2.3) in its emission intensity than D26G (I_f 1.5) (Figure 3.4D) and WT HGSC (I_f 1.2); this result too suggests a greater exposure of non polar residues to the solvent in these (R140X, G165fs, L45PL54P, Y134A of HGDC and V42M of HGSC) mutants.

3.7. Stability of the mutants towards chemical denaturation

The fluorescence spectrum of denatured HGDC is higher in intensity and red shifted compared to that of the native state. These spectral differences have offered themselves as sensitive reporters of the transition between the native and unfolded states of the molecule (193). In order to assess the structural stability of the WT and mutant proteins, unfolding experiments using the well known chemical denaturing agent guanidinium hydrochloride (GuHCl) were performed.

Figures 3.7A-C show the GuHCl induced denaturation profiles of WT HGDC, A36P and R140X. WT HGDC denatures in single step and shows the classical behaviour of two state transition (Figure 3.7A) with only the folded and unfolded species, while A36P denatures in two steps, or a three state transition

(Figure 3.7B). The WT HGDC and A36P denature at transition mid points 2.81 M and 2.73 M GuHCl concentration respectively. However A36P displays an earlier transition, with a midpoint around 0.45 M, indicating the population of a partially unfolded intermediate. However, its second transition occurs at the same 2.8 M GuHCl as the WT indicating their C-terminal domain is perhaps not affected upon mutation. The free energy $\Delta G°$ value of the WT HGDC was estimated using the two-state denaturation model (184) to be 9.4 kcal mol^{-1}. The free energy values of A36P were $\Delta G°_1$ value of 2.61 kcal mol^{-1} for the first transition and $\Delta G°_2$ value of 7.61 kcal mol^{-1} for the second transition.

The truncation mutant R140X appears to undergo a simple two-state unfolding (Figure 3.7C), with a transition midpoint at 1.83 M GuHCl with a $\Delta G°$ value of 2.65 kcal mol^{-1}. This is reminiscent of what Mills IA *et al*., (49) found for the N-terminal domain alone of HGDC ($\Delta G°$ 3.7 kcal mol^{-1} and equilibrium unfolding transition midpoints in units of M GuHCl [C_m] 1.21 M GuHCl). The loss of 4^{th} Greek key appears to make R140X behave somewhat similar to the N-terminal domain of HGDC, possibly because it might not display a strong inter-domain contact due to chain truncation.

The denaturation profiles of R77S (191) and P24T using GuHCl have been reported by others (164), and they all follow the simple two-state model of unfolding, though each of them denatures at slightly lower GuHCl concentrations than the WT HGDC. We were not able to study the denaturation of L45PL54P and G165fs due to their solubility and precipitation problems. This denaturation

experiments show that A36P, V42M and R140X are less stable and the former two are capable of forming intermediates at less concentration of GuHCl with lower ΔG° values.

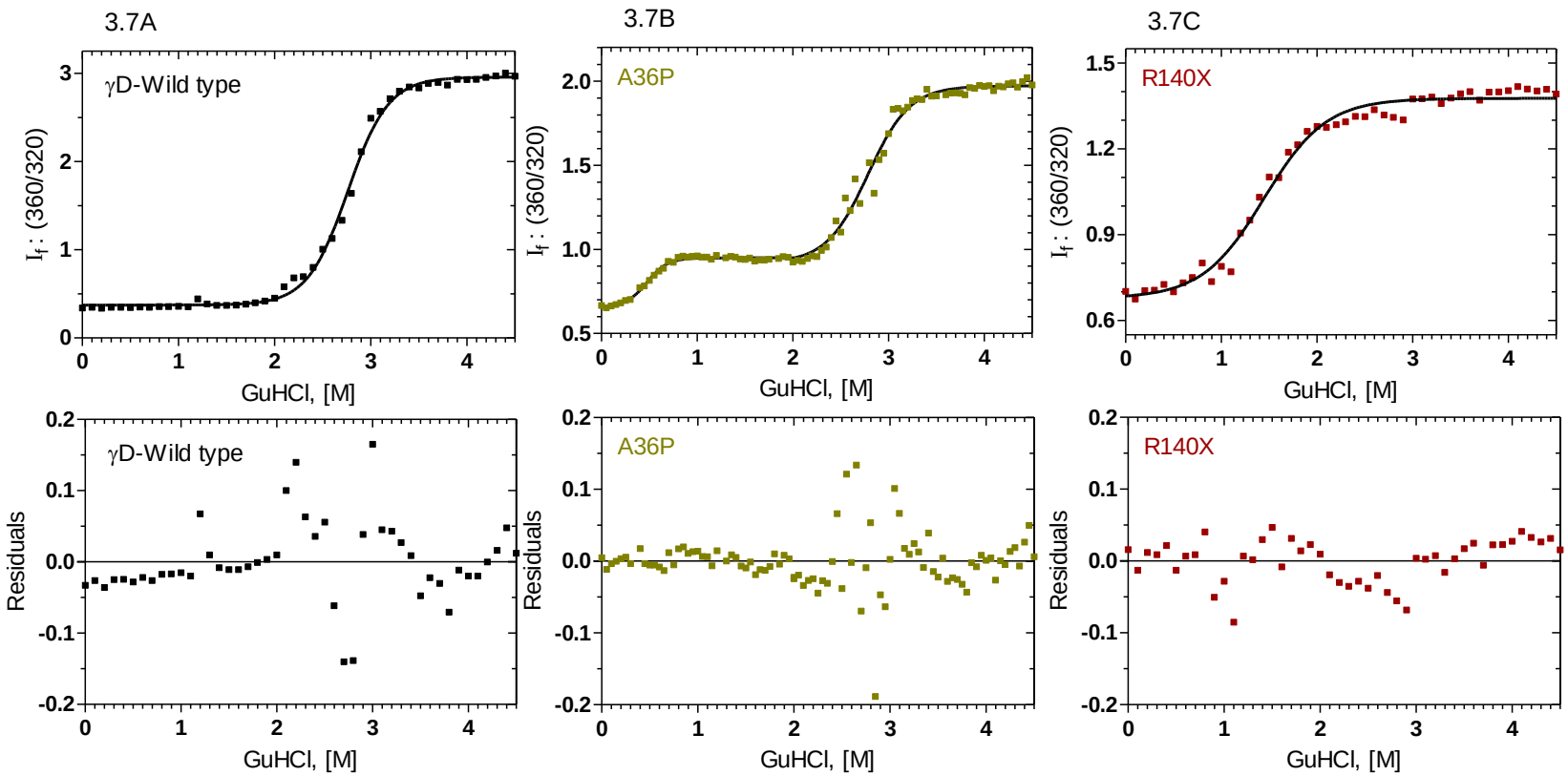


Figure 3.7 A-C: GuHCl induced denaturation of WT, A36P and R140X mutant γD-crystallins. (Black): γD-WT; (Olive): A36P; (Red): R140X; Relative emission intensity of the 360 nm band (of the denatured form) was compared to that of the 320 nm band (of the native protein) and monitored as a function of denaturant concentration. Solid line indicates the fitted data and solid blocks stand for raw data. Protein concentration in each sample was 0.2 mg/ml in 50 mM Tris-Cl buffer (pH7.3), 1 mM EDTA and 5 mM DTT.

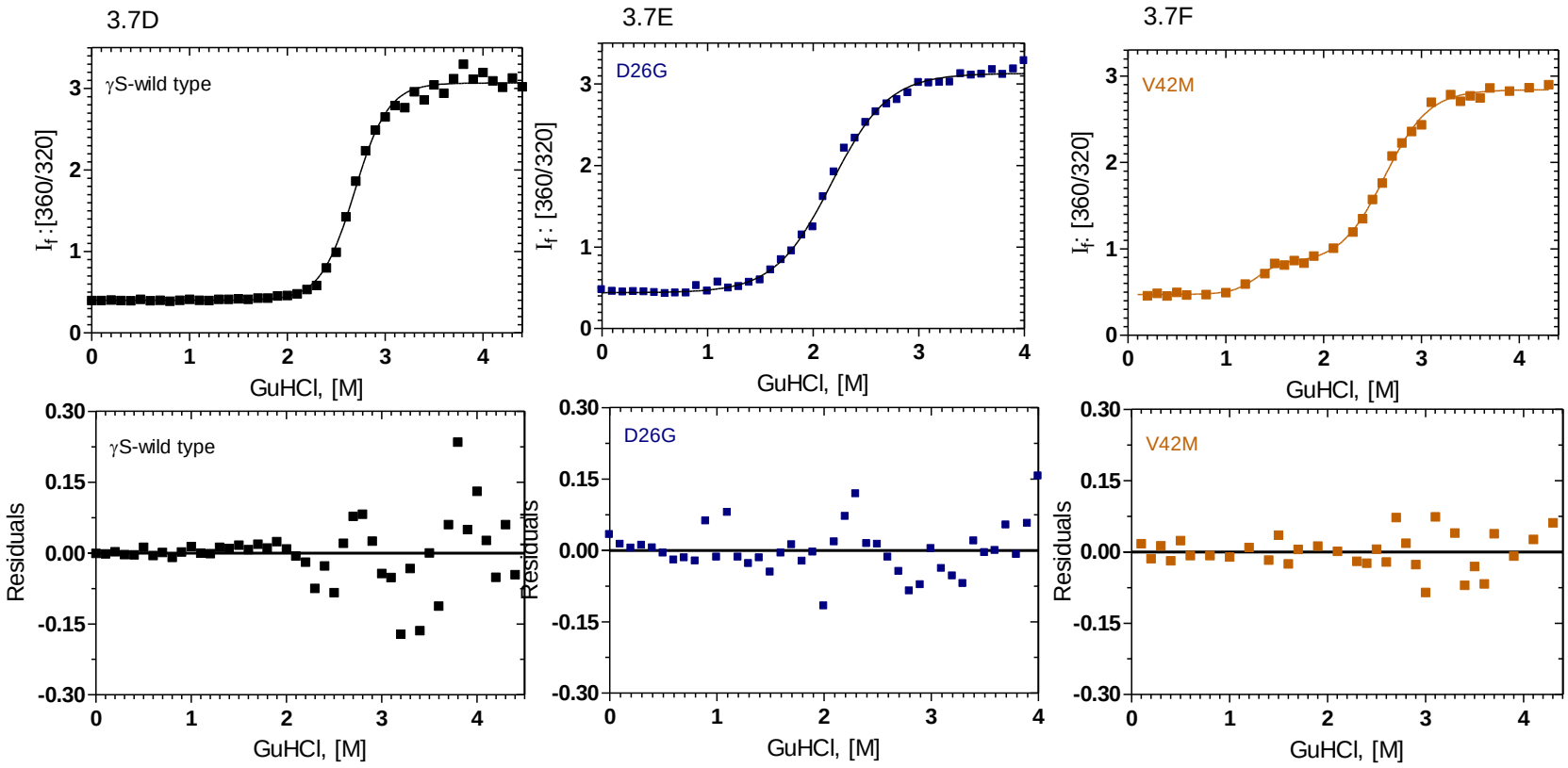


Figure 3.7 D-F: GuHCl induced denaturation of WT, D26G and V42M mutant γS-crystallins. (Black): γS-WT; (Blue): D26G; (Brown): V42M; Solid line indicates the fitted data and solid blocks stand for raw data Residuals of WT and mutants are also shown below the graphs λ_{exc}: 295 nm, cell path length 3 mm, excitation emission slits 5 nm. Other conditions were same as above.

Figures 3.7D-F show the GuHCl induced denaturation profiles of WT HGSC, D26G and V42M. Wild type HGSC and D26G (Figures 3.7D and 3.7E) denature in a single step and show two state transition, while V42M denatures (Figure 3.7F) in two steps and shows three state transition with an intermediate. A major unfolding transition occurs between 2 to 3.5 M GuHCl concentration, with the midpoint transition around 2.8 M GuHCl. The WT HGSC, and its mutants D26G and V42M denature at transition midpoints 2.8 M, 2.05 M and 2.8 M GuHCl concentration respectively. V42M displays an earlier transition, with a midpoint around 1.2 M GuHCl, indicating the population of a partially unfolded intermediate. However, its second transition occurs at the same 2.8 M GuHCl as the WT HGSC indicating C-terminal domain is perhaps not affected upon mutation. The free energy $\Delta G°$ values of WT HGSC and D26G were estimated using the two-state denaturation model (184) and they are of 7.76 kcal mol^{-1} and 4.80 kcal mol^{-1}. The $\Delta G°$ values for the first transition in V42M was estimated to be about 4.2 kcal mol^{-1}, while that for the second transition was 6.8 kcal mol^{-1}.

3.8-3.9. Thermal stability of mutants

We used differential scanning calorimetry (DSC) in order to study the thermal denaturation profiles of the WT and mutant molecules. Figure 3.8A shows a thermal melting temperature (T_m value) of 81.5°C for the WT HGDC, 78.9°C for P24T and 82.5°C for R77S. In contrast, the mutant A36P displayed a T_m value of 48.5°C (but the protein started precipitating after 55°C and hence we had to stop collecting data after 60°C) and since the mutants L45PL54P, R140X

and G165fs started precipitating upon raising the temperature beyond 40°C, we could not estimate their T_m values. Thermal denaturation curves obtained by DSC show the mutant A36P to have a lower thermal stability compared to WT, P24T and R77S mutants.

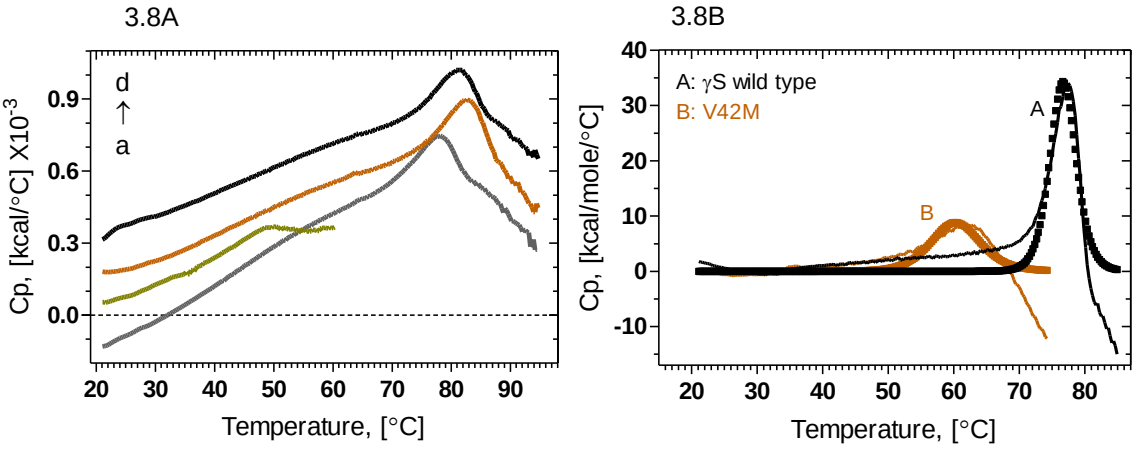


Figure 3.8A: DSC profiles of HGDC and its mutants. A (Grey): P24T; B (Olive): A36P; C (Brown): R77S; and D (Black): γD-WT. **Figure 3.8B: DSC profiles of HGSC and its mutant V42M**. A (Black): γS-WT and B (Brown): V42M. Fitted curves were obtained from cursor initiative fitting procedures by normalized data. Solid line indicates the raw data and blocks stand for fitted data. Protein concentration of 0.5 mg/ml was used and heated in the range of 20-95ºC with a rate of 1ºC / min.

Figure 3.8B shows the thermal unfolding profiles of the WT HGSC and V42M, obtained using differential scanning calorimetry. WT HGSC shows the thermal transition at 76.7ºC, while the mutant displays the transition far earlier at 60.3ºC, showing thereby that the mutant is weaker in stability. However, estimation of the enthalpies of unfolding was given up, since the protein solutions began to turn cloudy and starting to precipitate just before, during and after the

transition. This led to a poor goodness of fit (unacceptable χ^2 values), thus allowing us only to report the T_m values.

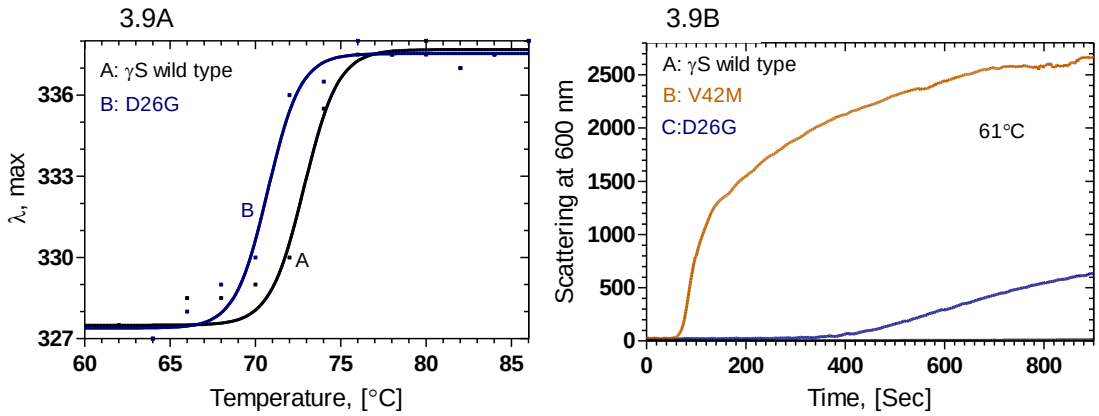


Figure 3.9A: Thermal unfolding of WT HGSC and mutant D26G. γS-WT; (Blue): D26G; monitored by following the change in the wavelength of emission (from 327 nm for the native form to 338 nm for the denatured form), with temperature, using the spectrofluorimeter. Protein concentrations were 5 µM (100 µg/ml) in 50 mM Tris-Cl buffer (pH 7.3), with 1 mM EDTA and 5 mM DTT. **Figure 3.9B: Mutant V42M starts scattering light while heated at 61°C while the WT HGSC and D26G do not.** (Black): γS-WT; (Blue): D26G; and (Brown): V42M. Time course of light scattering by the proteins at 600 nm light measured as a function of time, using 2.5 nm excitation and emission slits. A protein concentration of 5 µM (0.1 mg/ml) in 50 mM Tris-Cl buffer (pH 7.3) was used.

Figure 3.9A shows the thermal unfolding curves of WT HGSC and its mutant D26G. Fluorescence maximum value was noted when the proteins were denatured upon heating and plotted as transition curves. Thermal transition curves for both proteins were quite similar, and essentially two-state in shape. The WT showed the midpoint of transition at 72.0ºC, agreeing with Brubaker et al. (197), while D26G denatured at a lower temperature, with the midpoint at 70.8ºC, suggesting that the mutation lowers the stability of the molecule slightly.

In the calorimetry experiment, we found the mutant V42M to progressively start forming a cloudy suspension with increasing temperature, even before its denaturation temperature of 60.3ºC, while the WT HGSC was clear until it reached its denaturation temperature of 76.7ºC. This led us to monitor the time dependence of light scattering, i.e., turbidity at 600 nm, at the constant temperature of 61°C of both WT HGSC and mutants D26G and V42M. This is shown in Figure 3.9B where we note that the WT does not scatter light even after 800 s but the mutant V42M does so quite rapidly by 400 s of incubation. Notice too that the mutant D26G behaves somewhat similar to the WT HGSC than with V42M.

3.10. Nature of the protein aggregates

The mouse γB-crystallin mutant γB^{nop} (at physiological conditions) and WT human γD and γC (at lower pH) are all known to form amyloid type filaments when they are partially unfolded and expose their β-pleated sheets *in vitro*, (198-200). Keeping this in mind, we monitored this tendency with the aggregates formed by all the mutants. A ready method to check this is the use of the dye Thioflavin-T which, upon binding to an amyloid-forming protein, displays a significant red shift and enhanced intensity in its emission band in the 470- 570 nm region when excited at 444 nm (201).

HGDC, P24T, R77S, D26G and HGSC show very little Thioflavin-T fluorescence (Figure 3.10A and Figure 3.10B) at 490 nm but G165fs, L45PL54P, R140X and V42M show more fluorescence when bound to Thioflavin-T and the

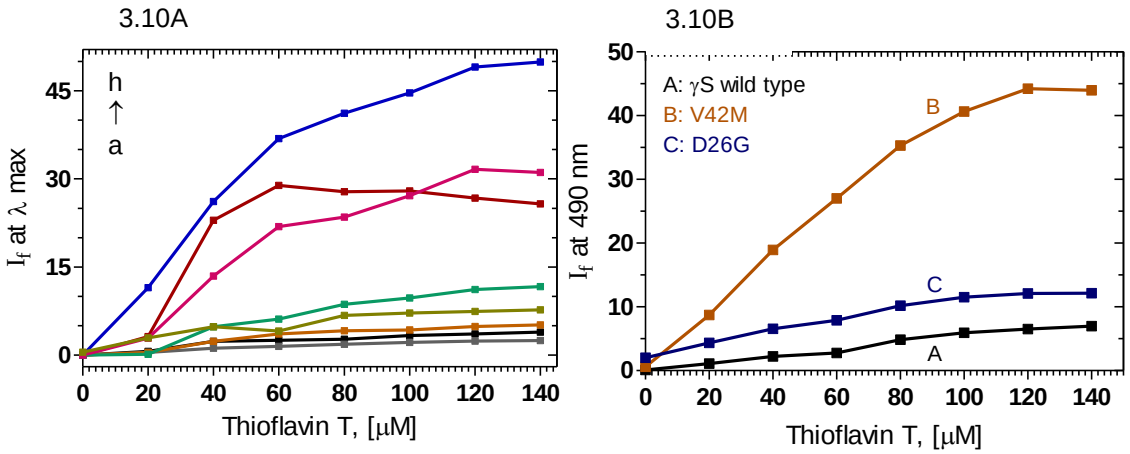


Figure 3.10: Aggregation in mutants appears amyloidogenic. Figure 3.10A: A (Grey): P24T; B (Black): γD-WT; C (Brown): R77S; D: (Olive): A36P; E (Green): Y134A; F (Maroon): R140X; G (Violet): L45PL54P; and H (Blue): G165fs. I_f of the probe at λ_{max} was measured as a function of increasing concentration. Protein concentration in each case was fixed at 6 μM (0.125 mg/ml) in 50 mM Tris-Cl buffer, cell path length 3 mm, excitation and emission slits 5 nm. **Figure 3.10B:** A (Black): γS-WT; B (Blue): D26G; C (Brown): V42M; I_f of the probe at 490 was measured as a function of increasing concentration. Protein concentration in each case was 5 μM (0.1 mg/ml) in 100 mM sodium phosphate buffer. Excitation and emission slits 10

the effect is in the following order: G165fs (I_f 41.1) > L45PL54P (I_f 23.5) > R140X (I_f 27.7) > >>Y134A (I_f 8.6) > A36P (I_f 6.7) > R77S (I_f 4.1) > P24T (I_f 1.8) > WT (HGDC) (I_f 2.6) at 80 μM of dye, V42M mutant of HGSC triggers almost eight-fold higher intensity (I_f 35.3) of Thioflavin-T emission (monitored at 490 nm) than when the dye is bound to the D26G (I_f 10.1) and WT HGSC protein (I_f 4.8)

suggesting that the aggregates formed from G165fs, L45PL54P, R140X, Y134A of γD and V42M of γS seems to have some amyloid-type character; but this needs to be confirmed by transmission electron microscopy (TEM).

3.11. *In situ* studies: does the mutant protein aggregate in cells?

The *in situ* aggregation property of WT HGDC and its truncated mutant W157X had already been studied in the HLE-3B cell line (172). This encouraged us to check the aggregation properties of the other mutants in HLE-3B cells. Figure 3.11 shows the confocal microscopic images of human lens epithelial cells HLE-3B transfected with WT HGDC and the various mutant cDNAs, tagged with 6X His tag and probed with anti-His antibody (mouse), and FITC-conjugated anti-mouse secondary antibody. The nuclei of the cells were counterstained with propidium iodide (PI) and visualized in red color. The figure shows that the WT HGDC, P24T and R77S distributed uniformly across the cell whereas A36P, R140X, G165fs, L45PL54P, and Y134A form punctate particles, suggesting *in situ* aggregation. Mutant Y134A too shows some punctate particles, perhaps due to the report that the sequence V126-Y134 in the molecule has an intrinsic propensity for aggregation (202). It is, however, not clear whether the aggregates seen in the cells are self-aggregates or hetero aggregates but we did not pursue this matter further.

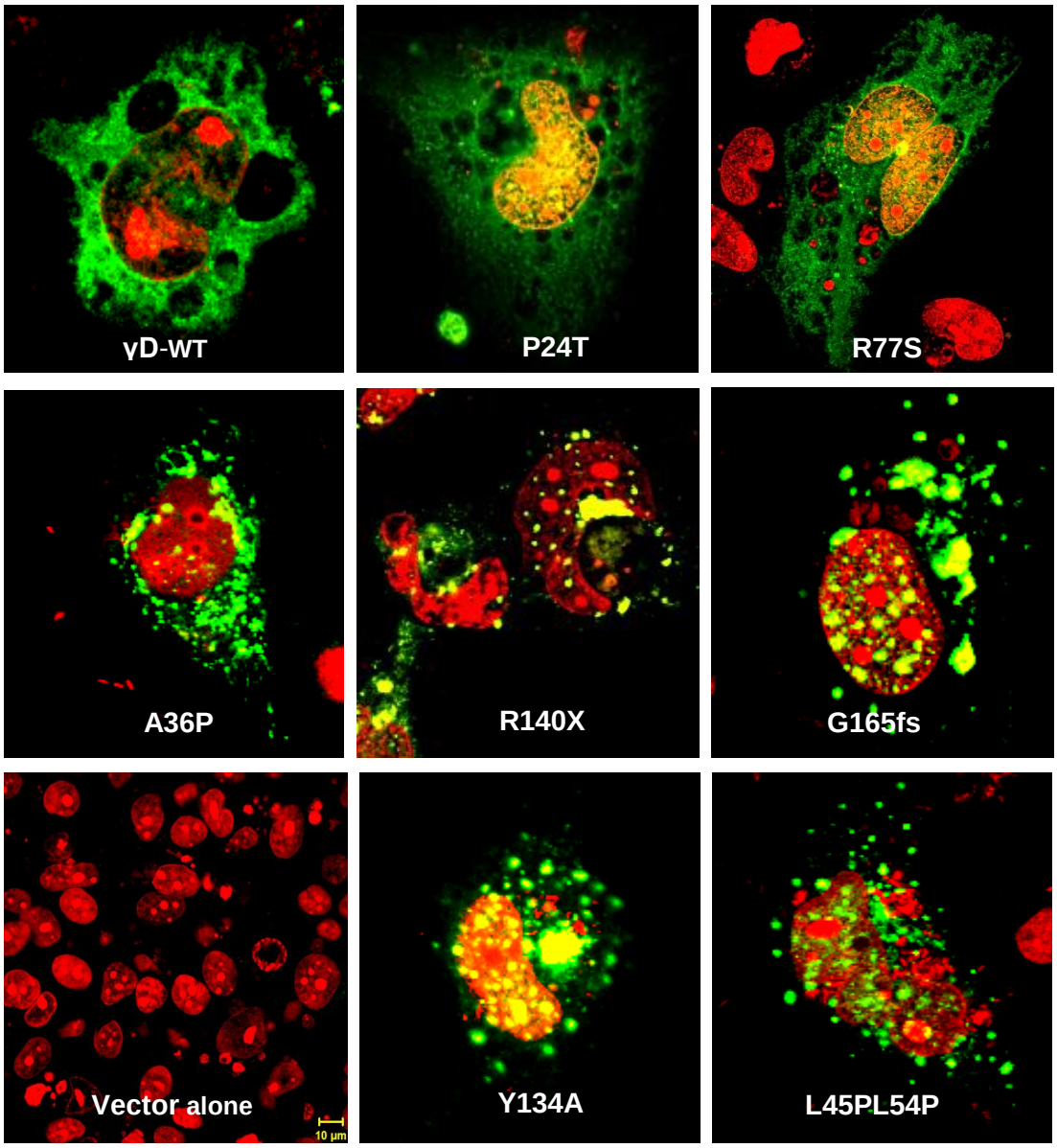


Figure 3.11: Visualizing the aggregation of the protein *in situ* in the human lens epithelial cell line HLE-3B using confocal microscopy. In each case, 6- His-tagged cDNA of the protein was transfected and visualized using anti-His antibody and FITC-conjugated secondary antibody. The nuclei of the cells were visualized using PI; Magnification: 200X for vector alone and 630X for others.

3.12. Molecular modeling and dynamics analysis of mutants

Note: The modeling work was done by our collaborators Prof. N. Srinivasan and Dr. Garima Agarwal of Molecular Biophysics Unit (MBU), Indian Institute of Sciences (IISc), Bangalore. We report the results here.

The high resolution crystal structure of WT HGDC (39) makes it possible to model the WT, as well as its mutants. We have used the interactive graphics software SETOR (203) to understand the consequences of each mutation on the structure of the molecule, including the accessible surface areas of exposed residues (204). The figures have been rendered using PyMOL (205), and Table 3.1 gives a summary of the structural analysis of the WT and various mutants of HGDC, including the solvent exposure of relevant residues, possible inter-domain interactions, and other features.

Figure 3.12A shows the modelled structures of the various mutants, where residues contributing to the increased nonpolar surface in the mutant were shown in magenta colour, and those contributing to increased polar surface in green.

Looking first at the mutant P24T, the threonine residue at the site of mutation has been labelled. Exposure of numerous nonpolar residues might reduce the solubility of the molecule. Increased exposure of polar residues also from C111 (labelled in the figure), would predispose the protein to aggregation.

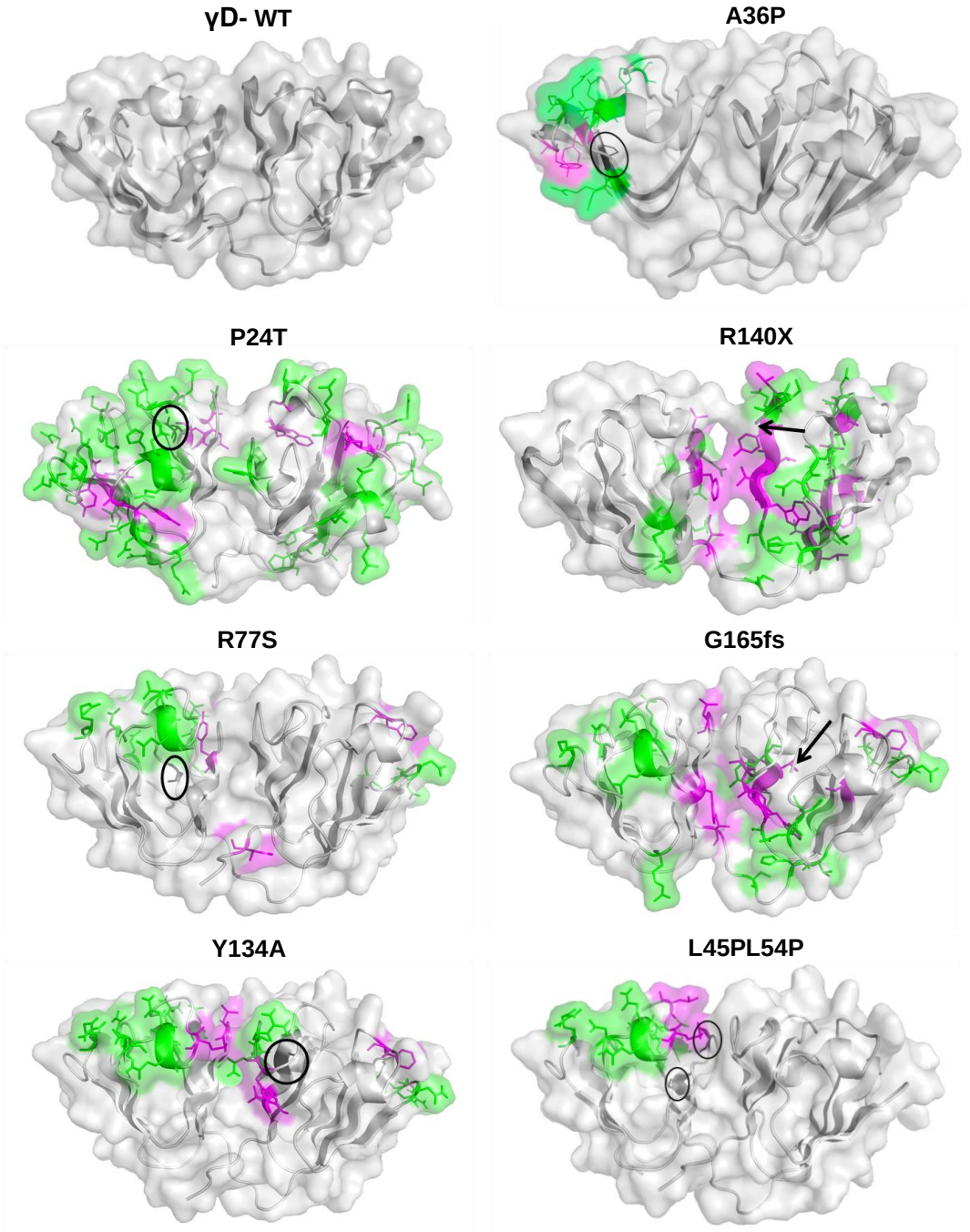


Figure 3.12A: Modeled structures of HGDC and its mutants. The residues with increased nonpolar surface exposed to the solvent are indicated in magenta, and polar residues indicated in green; the site of mutation is highlighted in black.

In the mutant A36P, the backbone φ angle at the site of mutation in the γD- WT is -156.4°. As the residue proline constrains φ to -60°, drastic alteration in the backbone conformation is expected, thus disrupting the tertiary structure as well. Additionally, the coulombic interaction that A36 has with surface D residues of the γD-WT molecule is lost upon replacing A36 by P. The exposure of nonpolar residues P43, L53, F118 and F173 (see Table 3.1) would also encourage aggregation.

On the other hand, in the double mutant L45PL54P, the Greek key motif 2 is disrupted, affecting the conformation and inter-domain interactions. Interestingly, more polar residues than nonpolar ones are exposed (see Table 3.1).

When we next look at R77S, we find that mutation of arginine 77 to serine reduces the solvent exposure as well as the positive charge at the site. The altered surface potential as well as the exposure of nonpolar groups might affect its association with other protein molecules and reduce its solubility.

The mutation Y134A is located in the middle strand of the 4th Greek key motif, and while it does not disrupt the topology, it does expose as many as 5 apolar and 13 polar side chains, which would lead to decreased solubility and possible aggregation.

Table 3.1: Summary of the structural analysis of various mutants of HGDC.

Mutant	Secondary structure and solvent accessibility of the residue	Nonpolar residues with increased solvent exposure in the mutant	Polar residues with increased solvent exposure in the mutant	Effect on Inter-domain interactions	Intermolecular interactions at the site of mutation
P24T	Edge strand in the GK-2 motif in the N-terminal domain is solvent-exposed	Y6,Y16,Y45,Y50 ,Y62,A63,Y98,F 118,W157,A159	T4,E7,R9,Q12,R14,H15,E17,D21 ,P23,S30,N33,R36,D38,Y45,Q47 ,S51,R59,D61,H65Q66,Q67,S72, D73,R76,S87,H88,R89,R95,R99, Q101,C111,Q113,R115,N125,E1 28,R140,T160,R163	Indirectly	Not directly affected
A36P	GK-1motif distorted due to Proline	P43,L53,F118,F 173	R9,N33, C41,Q47, H65,Q66, D 97,R99,C111,	GK-1 not involved in interdomain interactions. However,the conformation change and destabilization of GK-1 might indirectly affect.	Altered stability and conformation of the GK-1 might affect Intermolecular interactions
L45PL54P	Middle and edge strands,respectively , in GK-2 are solvent exposed	L53, M69, L71,	R9, E46, D64, Q66, Q67, S74, R76, R99	GK-2 involved in domain-domain interactions. The conformation and stability of this β-sheet is affected due to the presence of	Altered stability and conformation of the GK-2 might affect Intermolecular

	L45 is buried, L54 is solvent-exposed			prolines	interactions
R77S	One of the middle strands in GK-2 in N-terminal domain is solvent exposed	Y55, F118, I171	R9,F24,Q26,N33,E46,Q47,N49, Q66,Q67,S72,S74,S77,D97,R99, Q101	No	Occurs close to a positively charged patch in symmetry-related molecule
Y134A	Located at the middle strand in GK-4; buried	L53,M69,F118, V132,Y139,	R9,N33, E46, D64, N66, N67, S74,R76, D97, R99, Q101,R140, Q143	Possibly affected as the site occurs in GK-4 involved in inter-domain interactions	Does not seem to be affected
R140X		L53,F56,I81,I90, L92,Y98,L112,I121,L124,V126,L127,W131,V132, L133,Y134,L136	Y45,Q54,R59,R79,S84,H88,E96, D97,Y98,T106,D108,C109,Q113, N119,E120,H122,N125,S130,E135,S137,N138 *#Y144,L145,L146,W157,A159,A162,V164,L167,V170*	Affected	Could be affected due to unfolding of the molecule
G165fs		L53,F56,I81,L92,F118,W131,V132,L133,Y139,	R9,N33,R59,Q66,Q67,R76,S84, H88,D97,R99,Q101,T106,D108, C109,S130,Y134,*#V170*	Affected	Possibly affected due to unfolding

The residues indicated in italics are the residues buried in the WT but absent in the respective mutants.

In the mutant R140X, a significant fraction of the region contributing to the C-terminal domain is lost upon truncation. The deleted region includes three β-strands which form integral part of the jellyroll fold of the C-terminal domain. Loss of three β-strands and five buried apolar residues in the deleted region is expected to result in the near complete loss of structure of the C-terminal domain. Significant loss of tertiary structure, greater exposure of nonpolar residues and of C109 would be expected to lead to aggregation of the molecule. This would also be the case with the other mutant Y134X in comparison to R140X, this mutant has lost the hexa-peptide sequence ELSNYR, comprising almost all surface-seeking residues. The situation with Y56X, which has lost even greater part of the chain, is expected to be far more severe.

The frameshift mutation in G165fs results in premature termination, resulting in the loss of a C-terminal β-strand. The exposure of C109 and of multiple nonpolar residues along with the loss of tertiary contacts due to the missing β-strand would affect the stability of the molecule.

We next did molecular modeling and molecular dynamics work on HGSC and its mutants. This part of the work was done by our collaborator Dr. Sushil Chandani of Novarus Discoveries Pvt Ltd, Hyderabad. Turning to V42M molecule, bioinformatic analysis, using PolyPhen (206), predicted a high probability of distortion of the structure. To examine the structural effects of V42M substitution, homology models were built using as template murine γS-crystallin, with which the human protein shares an 89.8% identity. Analysis of the

WT HGSC showed the residue V42 to be in close proximity with I8, F10, W47, and L62, making a compact nonpolar core in the N-terminal domain. It is worth noting that in addition to V42, the residues I8, F10, W47 and L62 are also conserved not only among the γ-crystallins of humans, but also in dogs, rabbits, guinea pigs and mice. (In HGDC, F10 is replaced by L10, an equally hydrophobic residue of comparable van der Waals radius and nonpolar surface area). It would thus appear that these residues are vital for the compact, nonpolar core structure of this domain of these molecules. An analysis of the mutant V42M HGSC showed that fitting the larger methionine residue into the confines of this core causes a strain on the integrity of the structure.

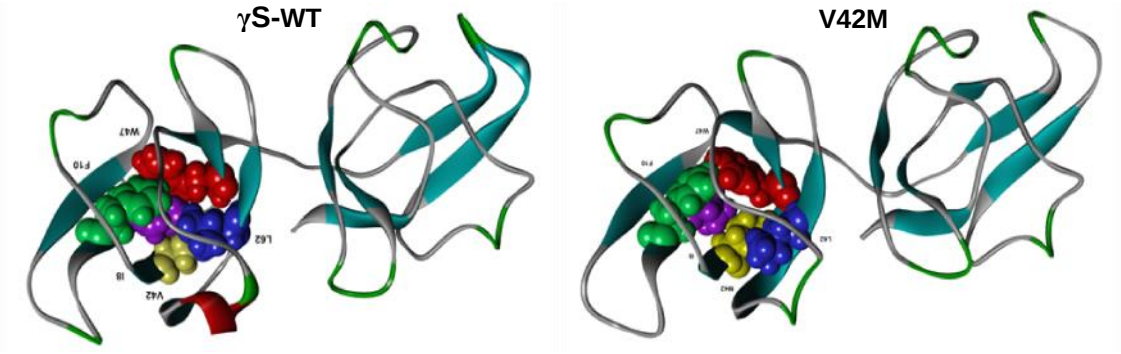


Figure 3.12B: Modeled structures of HGSC and its mutant V42M. The residues in the core of the N-terminal domain of the protein and in the vicinity of V42 and M42 as space-filling CPK models (Violet: I8; Green: F10; Red: W47; Blue: L62 and Yellow: V/M42).The rest of the protein is represented as a ribbon. The tighter packing of these residues is evident in the WT (HGSC) molecule, which is loosened in the V42M mutant. Note that the C-terminal domain is essentially unaffected by the mutation.

Figures 3.12B reveal this strain and 'opening up' of the compact core. Preliminary molecular dynamics analysis has suggested this strain is not relieved

in the course of annealing the molecule from low temperature to ambient. This *in silico* analysis suggests that this strain and consequent opening up causes a distortion of the compact β-sheet compact core topology of the N-terminal domain of the molecule, though the canonical Greek key fold itself is strained but not lost. Note from a comparison of the figures, too, that the C-terminal domain does not appear to have been affected by the mutation.

In the model of the WT HGSC molecule, the anionic side chain of D26 is seen to interact with the cationic side chain of residue R84 (Figure 3.12C, Panel A). In the selected rotameric conformation of R84, N54 is also found to lie within interacting distance of R84. In structures obtained after 2 ns of molecular dynamics too, the WT structure exhibited a strong D26-R84 interaction. In the mutant (Figure 3.12C, Panel B), the R84 side chain moved closer to N54, with the absence of R-D interaction leading to a widening of the distance between the adjacent anti parallel β-strands that bear these two amino acids. However, the Greek key folds are not lost or distorted in any major manner (see the overlays of WT HGSC and D26G in (Figure 3.12C, Panel C). It should be noted that the initial selection of the Arginine rotamer has a strong bearing on the observed interactions, as it is difficult to sample the conformational space of an amino acid with a long side-chain. Our examination of possible R84 conformations show that in the G26 mutant, aside from contact with N54 it might enter into a weaker, long-distance contact (~4 A$^{\circ}$) with the anionic side chain of D153 in the C-terminal domain (Figure 3.12C, Panel D).

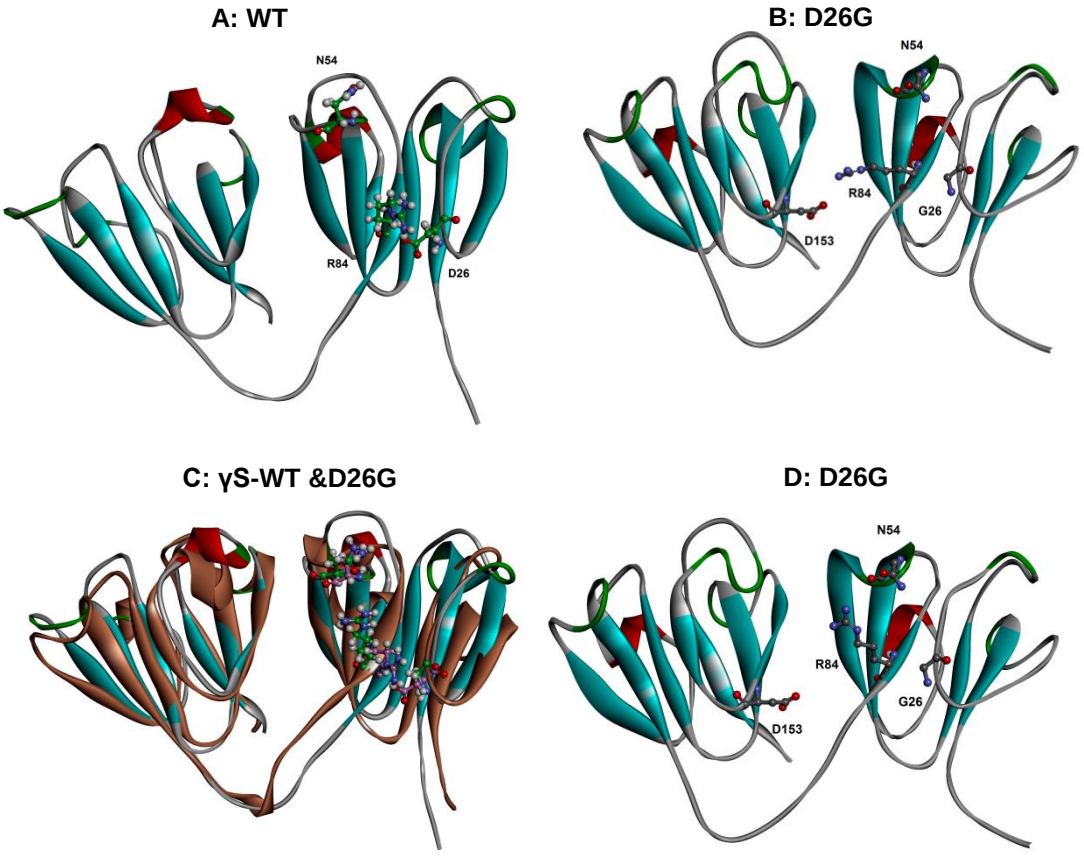


Figure 3.12C: Details of local packing changes caused by replacing aspartate by glycine in position 26 of the γS-Crystallin molecule. Panel A shows the interaction between D26 and R84 in the WT HGSC. Panel B shows the effect of replacing D26 by G. Panel C overlays the two structures (mutant D26G in brown and γS-WT in green). Panel D shows how R84 in the mutant, having lost contact with D26, might enter into a long distance contact with D153 in the C-terminal domain.

Discussion

4.1. Mutations in human γ-crystallins

As many as 29 mutations have been reported so far in human γ-crystallins,17 mutations (R14C, R14S, P24T, A36P, R37S, R37P, W43R, M44V, Y56X, R58H, G61C, R77S, E107A, Y134X, Y140X, Y156X and G165fs) in human γD-crystallin (HGDC), 8 mutations in human γC-crystallin (HGCC) (T5P, R48H, G38fs, C109X, S119S, G129C, W157X and R168W) and 4 mutations in human γS-crystallin (HGSC) (G18V, D26G, S39C and V42M). Table 1.4 in the introduction section lists these mutations reported in human HGDC, HGCC and HGSC and the associated cataract types.

Though our study has focused on some representative mutants of HGDC and HGSC, this interesting phenotypic dichotomy of nuclear cataract vs peripheral cataract is also seen in HGCC and its structurally similar family member, β-crystallin mutants as well (Tables 1.4B and 1.3).

We start our discussion with peripheral cataract associated mutant proteins P24T, R37S, R58H, G61C, R77S, G18V and R168W of HGDC. The mutant P24T is a very well studied molecule and is associated with different types of cataracts such as cerulean, lamellar, coralliform, fasiculiform and unclassified silica-like nuclear cataract. proline to threonine change at position 24 affects its solubility. Microconcentrator and lyophilization methods have shown its solubility to be 10-30 mg/ml compared to 450 mg/ml of the WT (164). NMR study

has revealed that the mutant has developed three surface hydrophobic patches, one with D21, H22, Y50 and R79, the second patch with N49, Y50 and V75 and the third one involving T4, L5, L6, H15, E17 and C18. These three surface hydrophobic patches may play a role in reducing its solubility (165). However, this change does not affect its stability significantly.

R36S is associated with crystal cataract. A cataractous lens from a patient with R36S mutation was studied by X-ray diffraction method. It was found that this crystal packing (2.25 A°) is only possible with serine and not with the bulky arginine due to its steric hindrances imposed at this 36 position (none of its rotamer fits at this position for the formation of crystals). In the mutant, serine side chain interacts with N24 position of the other molecule, while the other molecule's serine side chain interacts with N24 of the first molecule and losing the charge by replacing arginine with serine at this position also helps the molecule in forming crystals (142).

R58H is also associated with crystal cataracts and this too behaves similar to R36S. Basak's crystallography studies show that at position 58 replacement of arginine with histidine makes the mutant lose the direct ion-pair intermolecular interactions. In the WT, two carbon atoms of arginine form a strong ion-pair interaction with two nitrogen atoms of D156 of a second molecule. In the mutant, these intermolecular interactions are altered. The replacement of arginine with histidine kills the bond between D156 and allows it to interact with D97 with a

H_2O molecule interaction. This causes loss of the strong protein-protein, inter molecular ion-pair interactions at the mutation site (39).

G61C is associated with coralliform cataract. Here, the cysteine is not exposed to the surface and the mutant's thermal stability is close to that of the WT but its stability towards the chemical denaturant GuHCl is compromised. The mutant undergoes a two step unfolding, with a partially unfolded intermediate around 2.0 M of GuHCl (147).

R77S is associated with polar coronary cataract. NMR spectra show that the arginine to serine change at position 77 affects the local environment around L25 and Q47 and fails to form any interaction with E46 causing hydrophobic patches on the surface. However, the stability of this mutant is quite similar to that of the WT (191).

Turning to human HGSC the mutant G18V is less stable than WT, having a T_m of 65°C cf. 75°C for WT and its stability towards the chemical denaturant GuHCl is also altered, showing an intermediate around 0.8M GuHCl (207). This partial opening of mutant at low concentration of GuHCl is probably due to the distortion caused by valine in place of glycine at position 18. This position 18 is located in a solvent exposed loop of the protein with the torsion angles that are only allowable for glycine. Replacement of this glycine by the bulkier valine causes distortion (197).

Turning to HGCC, five mutants causing severe nuclear cataract phenotype and the other two peripheral cataracts. The mutants C109X and W157X can be explained very easily by their truncated forms in the similar way how truncated mutants in HGDC (Y56X, Y134X, Y140X and Y156X) which causes nuclear cataract.

The R168W mutant is reported to be associated with lamellar cataract. Tertiary fold of this mutant is slightly altered, though stability (monitored by thermal and chemical denaturation) is not altered significantly (171).

The T to P change at position 5 in the mutant T5P of HGCC is associated with Coppock cataract. This change restricts the solubility to 0.4 mg/ml. Secondary structure is slightly altered, β-pleated conformation is decreased significantly as expected because of newly introduced proline into the first β-pleated sheet. Results quantified by PROSEC for α, β-sheet, β-turn and random coil in the mutant are 8, 72, 2, 19% and for the WT they are 8, 51, 15 and 27% respectively. Thermal stability is also compromised (169). This change also forms aggregates in cells affects the protein-protein interaction with α-crystallin (170). There is no significant data available with extrinsic flourophores like bis-ANS, Nile Red and Thioflavin-T which actually reveals the surface hydrophobicity and aggregating properties of the molecule. It's lower solubility, loss of stability to heat and decreased protein-protein interactions with other crystallins may lead to the cataract formation.

G41fs, a frame shift mutant of HGCC, heterozygous in condition (5-bp repeat has been found in only one allele of each affected family member) is associated with zonular pulverulant cataract. The 5-bp repeat disrupts the reading frame of the coding sequence and results in translation of truncated protein consisting of 41 amino acid residues in first Greek key motif followed by 68 random amino acids. The first 41 amino acid are identical to γC-crystallin but, rest of the protein is completely different. The two motifs forming N-terminal domain are inter-dependent of hydrogen bonding (44) so this truncated polypeptide is predicted not to have the ability to fold properly to form the first Greek key motif. It has random coil in its secondary structure, no ordered tertiary structure and solubility is also severely affected (150).

S119S is associated with nuclear cataract. This polymorphism was observed in heterozygous state in the affected members. So far more than twenty genes have been reported to be associated with cataract. Here the authors had screened only for three genes i.e., *CRYAA, CRYGD* and *CRYGC*. Screening the other candidate genes may give the clue for its severe phenotype. No functional study is available to show whether this silent mutation affects its expression levels thus leading to cataract (144).

In short, there appears to be a set of common features of mutant human γ-crystallins associated with peripheral cataracts. These are: The Greek key

motif is not distorted in any major manner and the stability and solubility of the protein is marginally affected.

We now turn our discussion to nuclear cataract-causing mutant proteins W43R of HGDC and W157X of HGCC whose conformational studies are already available. W43R is associated with nuclear cataract. Its solubility is affected and restricted to 20 mg/ml. Thermal and stability towards a denaturant GuHCl is compromised. The mutant undergoes two step unfolding with a highly populated partially unfolded intermediate around 0.19M GuHCl with a ΔG° value of 2.4 kcal mol^{-1}. Unfolding experiments with 3.6M urea also shows mutants opening up indicating its stability to be compromised. Its NMR spectra are similar to those of another mutant, V75D, is known to undergo a two step denaturation with partially unfolded intermediate, indicating a prominent intermediate in its unfolding curve. The mutant is susceptible to protease cleaving, with the cleavage happening between R79 and L80 producing a 11 kDa band. The same cleaved band is also found in the WT HGDC upon UV irradiation damage that leads to formation of a partially unfolded species that can be easily cleaved by proteases (208). Its stability towards thermal stress is also compromised, with the mutant showing a T_m value of 77.8°C compared to a T_m value of 84.3°C for the WT (143).

In W157X, the last Greek key is broken. Its solubility is reduced to micrograms, tertiary fold altered, nonpolar side chains are exposed to the solvent

to a great extent and aggregation was observed in HLE-3B and COS-1 cells in comparison to the WT protein (172).

We now notice a common set of properties in mutants associated with nuclear cataracts. Their solubility and structural stabilities are greatly altered. They aggregate readily, forming light scattering particles *in vitro* and *in cellulo*, and displaying amyloidtype aggregates. All these properties are in sharp contrast to those associated with peripheral cataract.

4.2. Mutations in human β-crystallins

We now look at the structurally homologous human β-crystallins, where too the reported mutations (listed in Table 1.3) show a phenotypic dichotomy. Interestingly, here too we notice the putative connection between the structural integrity of the Greek key motif and the phenotypic dichotomy. Note that all the eight mutations reported in βB1-crystallin associated with nuclear cataracts, six of these (S129R, G220X, Q223X, S228P, R233H and X253R) are C-terminal domain mutants which disturb the Greek key motifs 3 or 4. Many of them (G220X, Q223X, S228P and X253R) were predicted to result in an abnormally elongated or truncated C-terminus and the production of a mutant protein.

The serine to arginine mutation at 129 position revealed that the mutation slightly altered the structures of both βB1-crystallin homomer and βB1/βA3-crystallin heteromer and importantly, the mutation significantly decreased the

thermal stability of βB1/βA3-crystallin but not βB1-crystallin affecting the protein-protein interactions among β-crystallins which are necessary in maintaining lens transparency (94).

Sequence-structure alignment of all the Greek key motifs shows that two quite distal residues, a glycine and a serine, are the most conserved and are involved in stabilizing the supersecondary fold by packing this β-hairpin over the β-sheet, replacing the serine with structure breaker proline at 228 disturbs Greek key motif's super secondary fold (98).

In R233H, the strongly positively charged guanidino group of arginine is replaced by the positively charged histidine, which may result in instability and poor solubility of the molecule (99).

The other two mutants M1K and N58Tfs both in the N-terminal domain might be predicted to lead to an absence of a functional protein product either by abrogation of translation or by nonsense mediated mRNA decay. M1K abrogates the initiation codon. The codon AUG is optimum for translational initiation. In M1K the mutation alters the first codon AUG as AAG codon. In this condition translation might start at a downstream 113 codon AUG, producing the translated protein lacking the first 112 amino acids of the WT protein including the Greek key I and part of the Greek key 2 motif (92). In N58Tfs, deletion of G takes place at 168 position of exon 2 and generates a 106 amino acid long truncated

molecule. There is a strong possibility that this frame shift and premature truncation may lead to nonsense-mediated decay and thus no protein product (93).

Turning to human βB2-crystallin, mutations A2V, I21N and S31W do not affect any of the Greek key motifs, while the others would be predicted to do and lead to nuclear cataracts. A2V change retards the tetramerization of βB2-crystallin at high protein concentrations (209). The 31st residue serine substituted by tryptophan (S31W) enhances hydrophobicity and creates a lipocalin signature motif (lipocalins share characteristic conserved sequence motifs and bind to small hydrophobic molecules, such as steroids, bilins, retinoids, and lipids and are capable to bind cell-surface receptors and form complexes with soluble macromolecules). This is suspected to alter the local binding ability which would affect its intermolecular interactions (102).

D128V and V187M are associated with nuclear cataract phenotype. D to V replacement at 128 position makes the mutant lose the negative electrostatic potential surrounding the D and is replaced by a remarkably enlarged positive potential in this region, which may affect protein-protein interactions (104). The V187M substitution is thought to affect the fourth Greek key motif, potentially disturbing the integrity of the subunit interaction domains which is important in sustaining their stability and dimerization (115).

In βB3, R75H affects the second motif and destroys a highly conserved R75 and may affect the solubility (114), while in G165R mutation, which is associated with nuclear cataract. Glycine at 165 position is located at the tip of the loop connecting β-strands 1 and 2 of the fourth Greek key motif. Maintenance of the hairpin fold requires amino acid residues with small (alanine) or no (glycine) side chains, which can give the acute torsion angles needed to form the Greek key motif. The G165R change would be expected to open up the hairpin fold and destabilize the fourth Greek key motif. The residue G165 resides 8 A° and 11 A° from residues R167 and R195 respectively, on substitution with R165 results electrostatic repulsion and helps in destabilizing the Greek key motif integrity (116).

Turning to human βA3/A1-crystallin, the first 2 exons encode the amino terminal arm, and exons 3-6 encode Greek keys 1-4. Three mutations have been reported- two splice site mutations and one 3 bp deletion which thus loses sequences containing one or more Greek key motifs, and are invariably associated with nuclear cataract. With βA4-crystallin, 3 mutations are reported and these impair the β-strands forming the Greek key, destabilizing the protein by as much as 5 kcal mol^{-1}. All these lead to nuclear cataract accompanied by microcornea/microphthalmia (128). G64W substitution takes place at a corner of the backbone structure. Glycine is often found in β-sheet secondary structures and is the amino acid appearing most frequently at position i+2 of β-turn, glycine

at codon 64 probably forms similar secondary structures as β-sheet or β-turn. The substitution may result in damage of secondary structure during the CRYBA4 protein folding process so that the structure of the protein has reduced stability (129). L69P change which is predicted to disrupt the β-sheet structure in CRYBA4, protein folding would be impaired, most probably leading to a structure with reduced stability in the mutant (128). F94S substitution would significantly reduce the intrinsic stability of the crystallin monomer (128).

4.3. Caveats

There are other possible ways leading to nuclear cataract phenotype. One is that protein aggregation can also arise due to intermolecular disulfide cross-linking, as reflected in mutant R14C. Here cysteine at 14 is surface exposed and encourages dimer and oligomer formation. A freshly prepared R14C of HGDC crystallin showed 4% dimers and prolonged incubation at 37°C for 24 h its dimer percentage increased to 40%, 10% trimers and 5% higher oligomers were also seen. High critical temperature is an indication of oligomerization. The mutant showed a critical temperature of 20°C while the WT has showed 4°C (167). Raman spectroscopy study also shows about half of the cysteines in the mutant are involved in oxidation (168). Mutants W59C and W151C of βB2 might also do likewise.

The other is the formation of hetero-molecular aggregates through Coulombic interactions; one example is E107A of HGDC, associated with nuclear cataract. This glutamate to alanine change at 107 position alters attractive interactions between α-crystallin due to change in surface electrostatic potential in the mutant protein (pI change from 7.2 to 8.2). This pI change doesn't affect the γ-γ interactions but affects the γ-α interactions. Experiments with liquid-liquid phase separation in solution of pure γ crystallins as well as E107A+α and WT+α mixtures clearly shows γ-γ homologous interactions were not disturbed but shows strong interaction towards α-crystallin. This could be well explained by mutant's positive charge and α- crystallins negative charge at physiological pH (210). Likewise, R48H of human HGCC, and mutant I21N in βB2, associated with nuclear cataracts, might display a similar charge-driven intermolecular aggregate, through coulombic forces.

4.4. Congenital cataracts in Mice

While our discussion so far has been on human cataracts, many similar congenital cataracts have been reported in mice (211). Though there are differences in the length and actual sequences between mouse and human crystallins, there is considerable sequence homology between them. In addition even here, mutations affecting the Greek key folding or altering the pI of the mutant significantly (e.g., W43R γA, V76D γD, D77G γA, W168R βA1) seem to be associated with nuclear cataracts.

Summary of our results and analysis of all mutations reported so far in βγ-crystallins suggests that when the mutation distorts even one of the Greek key folds in the βγ-crystallin family, the resultant protein (a) loses its structural compactness and stability, (b) unfolds, exposing several buried residues to the surface causing self-aggregates, (c) becomes sparingly soluble in water, (d) tends to generate amyloidogenic aggregates *in vitro*, and (e) displays light scattering particles when transfected in cell lines, all of which are consistent with the phenotype of nuclear cataract. These mutations lead to what has been termed as 'protein disorder disease'. In contrast, mutations that do not disturb the Greek key motifs maintain the overall chain folding and only produce local disturbances around the mutation site leading to peripheral cataracts due to reduction in solubility. This situation has recently been described as one involving well-folded proteins but with aberrant homologous protein interactions (212) also termed 'native state aggregation' (198) or 'protein condensation disease' (213, 214). The structural integrity of the Greek key motif in βγ-crystallins thus appears to be an essential element in packing the crystallins in a compact, close-packed manner. Such packing appears to offer long term stability and stress resistance to the βγ-crystallins (40) and the short-range order they exhibit in the lens is thought to be responsible for the transparency of the lens (10).Our results suggest that distortion of even one of these motifs in the chain leads to nuclear cataract and central vision loss.

Summary

Cataract or eye lens opacification is the leading cause of blindness worldwide. It occurs either because of mutations in the genes coding for the lens proteins (particularly in cases of congenital cataract) or due to environmental and metabolic changes that occur to the lens constituents that accumulate over a period of time, compromising transparency (age-related cataract). Congenital cataract, afflicting over 1.4 million children in the world with a prevalence of 1 to 15/10,000 children (with birth prevalence of bilateral cataract being 1 to 3/10,000 births), is essentially a genetic disorder and is an important cause of childhood blindness affecting about 20,000-40,000 newborns worldwide yearly (2).

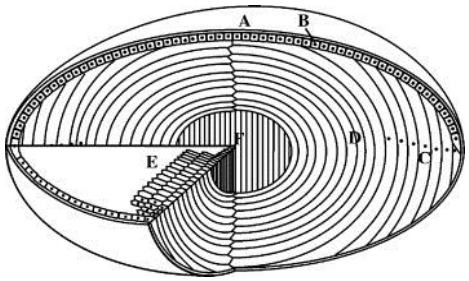


Figure 5.1: The human eye lens. A: capsule, B: epithelial cells, C: equatorial region, D: fiber cells removing their organelles denoted by the black dots, E: fiber cells with their characteristic hexagonal shape and F: nuclear region (7).

The human eye lens is primarily made up of fiber cells. These cells arise upon the differentiation and elongation of a single layer of epithelial cells. During differentiation, they lose their nuclei and organelles which helps to prevent light

scattering. Lens fiber cells are highly rich in proteins called crystallins. Crystallins constitute 90% of the total lens protein content, 35% of the total lens mass and are comprised of two types: α and βγ. Of these, two members of α (αA and αB) constitute 30% of the total lens crystallin content, 7 of β [four acidic members (βA1, βA2, βA3 and βA4), three basic members (βB1, βB2 and βB3)] account for 35% and three members of γ (γC, γD and γS) contribute to 25% of the total lens crystallin content in the human lens. The molecular weight of each of these crystallins in monomeric form is between 19-27 kDa. The α-crystallins belong to the small heat shock protein family and form multimers up to 60 units and form aggregates 800-1200 kDa in size. They can exchange subunits dynamically and show chaperone-like function by binding to partially opened and denatured β, γ-crystallins thus preventing them from forming aggregates or light scattering particles (28, 29). β-Crystallins also exist as oligomers while γ-crystallins occur as monomers. The main function of these proteins in the lens is to provide transparency and necessary refractive index gradient so as to focus the light onto the retina.

Distribution of these three types of crystallins in the lens is asymmetric and biphasic (4). The central portion of the lens is rich in the β and γ-crystallins (See Figure 5.1) and the embryonic nuclear region is particularly rich in γ-crystallins (6). These proteins are present at high concentrations and are

arranged compactly in a short range spatial order as in dense liquids or glasses to provide transparency and necessary refractive index (10). These two features are achieved by their being folded in a compact and highly stable manner using a series of super-secondary structural motifs called the Greek key fold. Each domain of the β and γ-crystallins consists of two intercalated antiparallel β-sheet Greek key motifs. Each motif consists of four sequential antiparallel β-strands arranged with -1, -1, +3 topology leaving C-terminal strand as an outer strand by generating $(3,1)_c$ type of Greek key motif in βγ-crystallins as shown in Figure 5.2 (40). This kind of Greek key motif super-secondary folding, intra-molecular interactions between domains, inter-molecular interactions including 3D domain swapping and local interactions involved in the close packing of co-operative units offers them stress resistance and structural stability (41).

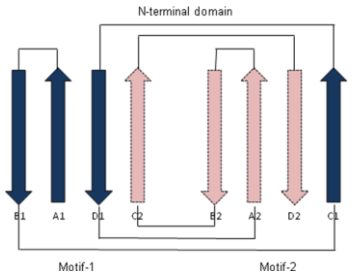


Figure 5.2: Cartoon depicting the topology of the characteristic domain structure found in all βγ-crystallins showing how two Greek key motifs, each comprising four β-strands (A, B, C and D) associate to form two motifs.

The role of the Greek key motif, based on its unique topology, has helped us understand the structural perspective, but how this Greek key-derived dense

packing of the β and γ-crystallins in the eye lens translates into transparency is an issue of functional interest. It is with this in mind that we have chosen congenital cataracts associated with mutations in β and γ-crystallins to study the functional role of the Greek key. The main reason for selecting congenital cataract mutations to study this issue is twofold. One is that congenital cataracts are essentially due to genetic mutations in origin and two is the fact that the 52 mutations that have been reported so far in human β and γ-crystallins are associated with an interesting phenotypic dichotomy (Table 1.3 and 1.4). About half of them generate nuclear cataract, blocking the central visual axis (Figure 5.3A) and hence causing complications which can disrupt the development of the visual pathways leading to conditions such as amblyopia (visual stimulation either fails to transmit or is poorly transmitted through the optic nerve to the brain), nystagmus (involuntary movement of the eye often called as dancing eye) and possibly alter association deficits in the growing infant, if not treated at the earliest (173-177). The rest are associated with peripheral cataracts which do not demand early action, since they do not block the visual axis (Figure 5.3B).

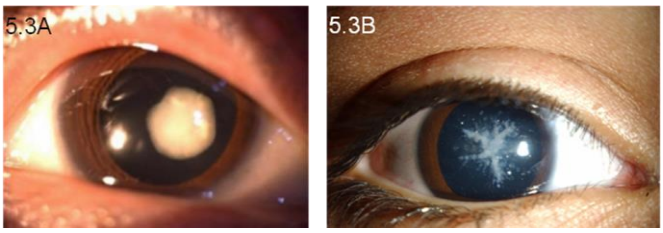


Figure 5.3A: Slit lamp photograph of the eye of a nuclear cataract patient (153). Figure 5.3B: Coralliform cataract is similarly shown (Coralliform: round or elongated processes radiating out of the center of the lens) (138).

In an effort to understand the molecular structural basis behind the phenotypic dichotomy (nuclear cataract-peripheral cataract), we have chosen to study representative mutations in human γD-crystallin (HGDC) and human γS-crystallin (HGSC). The reasons for selecting to study these two are: (a) they are monomers; the crystal and solution structures are well known for HGDC and the crystal structure of the C-terminal domain of HGSC has been studied in detail and the solution structure of the full length murine γS-crystallin has been resolved by NMR, (b) structural analysis of several mutants (R14C, P24T, R36S, R58H, G61C, E107A and W157X of HGDC and G18V of HGSC are available to extend our analysis and (c) these cataracts caused by respective mutations are Mendelian monogenic disorders.

It thus becomes possible to go beyond reporting the mutations in crystallins associated with congenital cataract and to study the functional genetics or attempt a protein structural rationale. We have chosen three mutants, two from HGDC (P24T and R77S) and one from HGSC (D26G) which are associated with peripheral cataracts, and three mutants (A36P, R140X and G165fs) of HGDC and the mutant V42M of HGSC associated with nuclear cataract. In addition, we have prepared two full length chain mutants of HGDC, not reported naturally: Y134A, which is a mutation in the fourth Greek key motif, but which still keeps all the four motifs intact and the double mutant L45PL54P

which disrupts the second Greek key motif, both of human γD-crystallin. We have then compared their properties with the other mutants and their respective WT by cloning, expressing and isolating the WT and mutant proteins (P24T, R77S, A36P, R140X, G165fs and L45PL54P, Y134A of HGDC and D26G and V42M of HGSC).

We then compared their solubility, structural details using circular dichroism (CD) and fluorescence spectroscopy, structural stability towards chemical denaturation (using the commonly used chemical denaturing agent guanidine hydrochloride, or GuHCl) and thermal denaturation, their tendency to form light scattering aggregate particles *in vitro* using spectral probes, and *in cellulo* by transfecting their cDNAs into a lens cell line, as well as molecular modeling *in silico* by standard methods

Table 5.1 summarizes the relevant properties of the total of 10 mutant proteins we cloned, expressed, isolated, purified and studied their conformational and aggregational properties namely the mutants P24T, A36P, L45PL54P, R77S,Y134A, R140X and G165fs of HGDC and the mutants D26G and V42M of HGSC, as well as the WT proteins. As the Table point out, these mutants fall into two broad groups, one group contains those in which the mutation has not significantly affected the properties, and at best causes local micro-environmental changes leading to mild aggregation with time. The other group

Table 5.1: Summary of some of the properties of representative mutants involved in this study

S.No	Property	Group 1 mutants	Group 2 mutants
1	Solubility	Soluble, though reduced in comparison to the WT	Mostly insoluble, had to be rescued from inclusion bodies
2	Micro-environment around tryptophan in the native state	Tryptophan is buried (λ_{max} 327 nm, I_{em} comparable to WT)	Tends to be exposed (λ_{max} 333-350 nm, I_{em} >> WT)
3	Exposure of nonpolar sidechains, monitored by extrinsic fluorophores	Little exposure to the solvent (e.g., P24T; I_f 6.8, 0.6, 1.8 with bis-ANS, Nile Red and Thioflavin-T respectively, using 80 µM of each dye)	Greater exposure to the solvent (e.g., R140X; I_f 262.2, 4.1, 27.7 with bis-ANS, Nile Red and Thioflavin-T respectively)
4	Structural stability monitored by chemical denaturation	Two-state denaturation profile, occurring earlier than WT and ΔG° lower than WT	Three state denaturation with an intermediate
5	Thermal denaturation	T_m value about the same or lower than WT by a few degrees (e.g., WT HGDC 81.5°C and 79°C for P24T°C)	T_m value much lower than WT (e.g., 49°C for A36Pof HGDC and 60°C for V42M cf. 77°C of HGSC)
6	Time dependent light scattering at 61°C	Comparable to WT. Very little scattering after 900 s	Occurs far sooner (e.g., at 200 s V42M of γS)
7	*In situ* aggregation when transfected in cell line	Does not occur	Occurs in cells
8	Cataract phenotype	Essentially peripheral	Invariably nuclear

Group 1 mutants: P24T, R77S, Y134A, WT (HGDC) and D26G (HGSC)

Group 2 mutants: A36P, R140X, G165fs, L45PL54P (HGDC) and V42M (HGSC)

comprises mutations which have distorted the Greek key folding in a fashion that leads to major changes in the properties and ready aggregation almost like amyloid-type particles. It is also to be noted that mutants belonging to Group 1

are associated with peripheral cataracts, while those in Group 2 are associated with nuclear cataract.

Based on these results, we have extended our discussion to the reported mutations in other γ and β-crystallins present in the human infant lens. The results of such comparison and analysis of all mutations in the βγ-crystallins reported so far, suggest that when the mutation distorts even a single Greek key motif sequence in the molecule, the resultant protein (a) becomes sparingly soluble in water, (b) loses its structural compactness and weakens its stability (c) exposes several buried residues to the surface causing self-aggregates, (d) tends to generate amyloidogenic aggregates *in vitro*, and (e) displays light scattering particles when transfected in cell line, all of which are consistent with the phenotype of nuclear cataract. These mutations lead to what has been termed as a 'protein disorder disease'. In contrast, mutations that do not affect the Greek key fold do not display most of the above properties, but only affect the local environment around the mutation site, generating what has been termed as 'native state aggregates' or 'protein condensation disease', consistent with the observed progressive peripheral cataract phenotype. Our work thus highlights the connection between the Greek key super-secondary structural motif in the βγ-crystallins and lens transparency.

In summary, we show here that the Greek key supersecondary structural motif found in human βγ-crystallins (which are rich in the nuclear region of the human lens) plays an essential functional role in generating and maintaining transparency in the mammalian eye lens.

References

1. Resnikoff S, Pascolini D, Etya'ale D, Kocur I, Pararajasegaram R, et al. (2004) Global data on visual impairment in the year 2002. Bull World Health Organ 82: 844-851.
2. Foster A, Gilbert C, Rahi J (1997) Epidemiology of cataract in childhood: a global perspective. J Cataract Refract Surg 23 Suppl 1: 601-604.
3. Eckstein M, Vijayalakshmi P, Killedar M, Gilbert C, Foster A (1996) Aetiology of childhood cataract in south India. Br J Ophthalmol 80: 628-632.
4. Augusteyn RC (2010) On the growth and internal structure of the human lens. Exp Eye Res 90: 643-654.
5. Kuszak JR (1990) Embryology and anotomy of the lens. Clinical Ophthalmology, JB Lippincott, Philadelphia: p. 1-9.
6. McAvoy JW (1978) Cell division, cell elongation and distribution of alpha-, beta- and gamma-crystallins in the rat lens. J Embryol Exp Morphol 44: 149-165.
7. Perng MD, Zhang Q, Quinlan RA (2007) Insights into the beaded filament of the eye lens. Exp Cell Res 313: 2180-2188.
8. Bassnett S (2002) Lens organelle degradation. Exp Eye Res 74: 1-6.
9. Fagerholm PP, Philipson BT, Lindstrom B (1981) Normal human lens - the distribution of protein. Exp Eye Res 33: 615-620.
10. Delaye M, Tardieu A (1983) Short-range order of crystallin proteins accounts for eye lens transparency. Nature 302: 415-417.
11. Swamy MS, Abraham EC (1991) Reverse-phase HPLC analysis of human alpha crystallin. Curr Eye Res 10: 213-220.

12. Kato K, Shinohara H, Kurobe N, Goto S, Inaguma Y, et al. (1991) Immunoreactive alpha A crystallin in rat non-lenticular tissues detected with a sensitive immunoassay method. Biochim Biophys Acta 1080: 173-180.
13. Bhat SP, Nagineni CN (1989) alpha B subunit of lens-specific protein alpha-crystallin is present in other ocular and non-ocular tissues. Biochem Biophys Res Commun 158: 319-325.
14. Dubin RA, Wawrousek EF, Piatigorsky J (1989) Expression of the murine alpha B-crystallin gene is not restricted to the lens. Mol Cell Biol 9: 1083-1091.
15. Iwaki T, Kume-Iwaki A, Liem RK, Goldman JE (1989) Alpha B-crystallin is expressed in non-lenticular tissues and accumulates in Alexander's disease brain. Cell 57: 71-78.
16. Lowe J, Landon M, Pike I, Spendlove I, McDermott H, et al. (1990) Dementia with beta-amyloid deposition: involvement of alpha B-crystallin supports two main diseases. Lancet 336: 515-516.
17. Renkawek K, de Jong WW, Merck KB, Frenken CW, van Workum FP, et al. (1992) alpha B-crystallin is present in reactive glia in Creutzfeldt-Jakob disease. Acta Neuropathol 83: 324-327.
18. Duguid JR, Rohwer RG, Seed B (1988) Isolation of cDNAs of scrapie-modulated RNAs by subtractive hybridization of a cDNA library. Proc Natl Acad Sci U S A 85: 5738-5742.
19. Bova MP, Ding LL, Horwitz J, Fung BK (1997) Subunit exchange of alphaA-crystallin. J Biol Chem 272: 29511-29517.

20. Haley DA, Horwitz J, Stewart PL (1998) The small heat-shock protein, alphaB-crystallin, has a variable quaternary structure. J Mol Biol 277: 27-35.
21. Horwitz J (2003) Alpha-crystallin. Exp Eye Res 76: 145-153.
22. Peschek J, Braun N, Franzmann TM, Georgalis Y, Haslbeck M, et al. (2009) The eye lens chaperone alpha-crystallin forms defined globular assemblies. Proc Natl Acad Sci U S A 106: 13272-13277.
23. Haslbeck M, Kastenmuller A, Buchner J, Weinkauf S, Braun N (2008) Structural dynamics of archaeal small heat shock proteins. J Mol Biol 378: 362-374.
24. Kim KK, Kim R, Kim SH (1998) Crystal structure of a small heat-shock protein. Nature 394: 595-599.
25. Bagneris C, Bateman OA, Naylor CE, Cronin N, Boelens WC, et al. (2009) Crystal structures of alpha-crystallin domain dimers of alphaB-crystallin and Hsp20. J Mol Biol 392: 1242-1252.
26. Jehle S, van Rossum B, Stout JR, Noguchi SM, Falber K, et al. (2009) alphaB-crystallin: a hybrid solid-state/solution-state NMR investigation reveals structural aspects of the heterogeneous oligomer. J Mol Biol 385: 1481-1497.
27. Laganowsky A, Benesch JL, Landau M, Ding L, Sawaya MR, et al. (2010) Crystal structures of truncated alphaA and alphaB crystallins reveal structural mechanisms of polydispersity important for eye lens function. Protein Sci 19: 1031-1043.
28. Horwitz J (1992) Alpha-crystallin can function as a molecular chaperone. Proc Natl Acad Sci U S A 89: 10449-10453.

29. Horwitz J (2000) The function of alpha-crystallin in vision. Semin Cell Dev Biol 11: 53-60.
30. Raman B, Rao CM (1994) Chaperone-like activity and quaternary structure of alpha-crystallin. J Biol Chem 269: 27264-27268
31. Raman B, Rao CM (1997) Chaperone-like activity and temperature-induced structural changes of alpha-crystallin. J Biol Chem 272: 23559-23564.
32. Rajaraman K, Raman B, Rao CM (1996) Molten-globule state of carbonic anhydrase binds to the chaperone-like alpha-crystallin. J Biol Chem 271: 27595-27600.
33. Kumar LV, Rao CM (2000) Domain swapping in human alpha A and alpha B crystallins affects oligomerization and enhances chaperone-like activity. J Biol Chem 275: 22009-22013.
34. Thomson JA, Augusteyn RC (1985) Ontogeny of human lens crystallins. Exp Eye Res 40: 393-410.
35. Werten PJ, Lindner RA, Carver JA, de Jong WW (1999) Formation of betaA3/betaB2-crystallin mixed complexes: involvement of N- and C-terminal extensions. Biochim Biophys Acta 1432: 286-292.
36. Bateman OA, Lubsen NH, Slingsby C (2001) Association behaviour of human betaB1-crystallin and its truncated forms. Exp Eye Res 73: 321-331.
37. Bateman OA, Sarra R, van Genesen ST, Kappe G, Lubsen NH, et al. (2003) The stability of human acidic beta-crystallin oligomers and hetero-oligomers. Exp Eye Res 77: 409-422.

38. Bloemendal H, de Jong W, Jaenicke R, Lubsen NH, Slingsby C, et al. (2004) Ageing and vision: structure, stability and function of lens crystallins. Prog Biophys Mol Biol 86: 407-485.
39. Basak A, Bateman O, Slingsby C, Pande A, Asherie N, et al. (2003) High-resolution X-ray crystal structures of human gammaD crystallin (1.25 A) and the R58H mutant (1.15 A) associated with aculeiform cataract. J Mol Biol 328: 1137-1147.
40. Hutchinson EG, Thornton JM (1993) The Greek key motif: extraction, classification and analysis. Protein Eng 6: 233-245.
41. Jaenicke R, Slingsby C (2001) Lens crystallins and their microbial homologs: structure, stability, and function. Crit Rev Biochem Mol Biol 36: 435-499.
42. Hemmingsen JM, Gernert KM, Richardson JS, Richardson DC (1994) The tyrosine corner: a feature of most Greek key beta-barrel proteins. Protein Sci 3: 1927-1937.
43. Kong F, King J (2011) Contributions of aromatic pairs to the folding and stability of long-lived human gammaD-crystallin. Protein Sci 20: 513-528.
44. Slingsby C, Clout NJ (1999) Structure of the crystallins. Eye (Lond) 13 (Pt 3b): 395-402.
45. Kosinski-Collins MS, Flaugh SL, King J (2004) Probing folding and fluorescence quenching in human gammaD crystallin Greek key domains using triple tryptophan mutant proteins. Protein Sci 13: 2223-2235.

46. Flaugh SL, Kosinski-Collins MS, King J (2005) Contributions of hydrophobic domain interface interactions to the folding and stability of human gammaD-crystallin. Protein Sci 14: 569-581.
47. Flaugh SL, Kosinski-Collins MS, King J (2005) Interdomain side-chain interactions in human gammaD crystallin influencing folding and stability. Protein Sci 14: 2030-2043.
48. Das P, King JA, (2010) Zhou R beta-Strand interactions at the domain interface critical for the stability of human lens gammaD-crystallin. Protein Sci 19: 131-140.
49. Mills IA, Flaugh SL, Kosinski-Collins MS, King JA (2007) Folding and stability of the isolated Greek key domains of the long-lived human lens proteins gammaD-crystallin and gammaS-crystallin. Protein Sci 16: 2427-2444.
50. Lubsen NH, Aarts HJ, Schoenmakers JG (1988) The evolution of lenticular proteins: the beta- and gamma-crystallin super gene family. Prog Biophys Mol Biol 51: 47-76.
51. Bax B, Lapatto R, Nalini V, Driessen H, Lindley PF, et al. (1990) X-ray analysis of beta B2-crystallin and evolution of oligomeric lens proteins. Nature 347: 776-780.
52. Ramaekers FC, Dunia I, Dodemont HJ, Benedetti EL, Bloemendal H (1982) Lenticular intermediate-sized filaments: biosynthesis and interaction with plasma membrane. Proc Natl Acad Sci U S A 79: 3208-3212
53. Nicholl ID, Quinlan RA (1994) Chaperone activity of alpha-crystallins modulates intermediate filament assembly. EMBO J 13: 945-953.

54. Sandilands A, Prescott AR, Hutcheson AM, Quinlan RA, Casselman JT, et al. (1995) Filensin is proteolytically processed during lens fiber cell differentiation by multiple independent pathways. Eur J Cell Biol 67: 238-253.
55. Georgatos SD, Gounari F, Goulielmos G, Aebi U (1997) To bead or not to bead? Lens-specific intermediate filaments revisited. J Cell Sci 110 (Pt 21): 2629-2634.
56. Chepelinsky AB (2009) Structural function of MIP/aquaporin 0 in the eye lens; genetic defects lead to congenital inherited cataracts. Handb Exp Pharmacol: 265-297.
57. Harries WE, Akhavan D, Miercke LJ, Khademi S, Stroud RM (2004) The channel architecture of aquaporin 0 at a 2.2-A resolution. Proc Natl Acad Sci U S A 101: 14045-14050.
58. Sohl G, Willecke K (2004) Gap junctions and the connexin protein family. Cardiovasc Res 62: 228-232.
59. Simon AM, Goodenough DA (1998) Diverse functions of vertebrate gap junctions. Trends Cell Biol 8: 477-483.
60. Louis CF, Hur KC, Galvan AC, TenBroek EM, Jarvis LJ, et al. (1989) Identification of an 18,000-dalton protein in mammalian lens fiber cell membranes. J Biol Chem 264: 19967-19973
61. Tenbroek E, Arneson M, Jarvis L, Louis C (1992) The distribution of the fiber cell intrinsic membrane proteins MP20 and connexin46 in the bovine lens. J Cell Sci 103 (Pt 1): 245-257.

62. Shiels A, Hejtmancik JF (2013) Genetics of human cataract. Clin Genet 84: 120-127.
63. Willecke K, Jungbluth S, Dahl E, Hennemann H, Heynkes R, et al. (1990) Six genes of the human connexin gene family coding for gap junctional proteins are assigned to four different human chromosomes. Eur J Cell Biol 53: 275-280.
64. Ponnam SP, Ramesha K, Matalia J, Tejwani S, Ramamurthy B, et al. (2013) Mutational screening of Indian families with hereditary congenital cataract. Mol Vis 19: 1141-1148.
65. Ponnam SP, Ramesha K, Tejwani S, Ramamurthy B, Kannabiran C (2007) Mutation of the gap junction protein alpha 8 (GJA8) gene causes autosomal recessive cataract. J Med Genet 44: e85.
66. Somaraju Chalasani ML, Muppirala M, SP GP, Kannabiran C, Swarup G A (2013) cataract-causing connexin 50 mutant is mislocalized to the ER due to loss of the fourth transmembrane domain and cytoplasmic domain. FEBS Open Bio 3: 22-29.
67. Graw J (2009) Genetics of crystallins: cataract and beyond. Exp Eye Res 88: 173-189.
68. Brakenhoff RH, Aarts HJ, Reek FH, Lubsen NH, Schoenmakers JG (1990) Human gamma-crystallin genes. A gene family on its way to extinction. J Mol Biol 216: 519-532.

69. Pras E, Frydman M, Levy-Nissenbaum E, Bakhan T, Raz J, et al. (2000) A nonsense mutation (W9X) in CRYAA causes autosomal recessive cataract in an inbred Jewish Persian family. Invest Ophthalmol Vis Sci 41: 3511-3515.
70. Hansen L, Yao W, Eiberg H, Kjaer KW, Baggesen K, et al. (2007) Genetic heterogeneity in microcornea-cataract: five novel mutations in CRYAA, CRYGD, and GJA8. Invest Ophthalmol Vis Sci 48: 3937-3944.
71. Graw J, Klopp N, Illig T, Preising MN, Lorenz B (2006) Congenital cataract and macular hypoplasia in humans associated with a de novo mutation in CRYAA and compound heterozygous mutations in P. Graefes Arch Clin Exp Ophthalmol 244: 912-919.
72. Laurie KJ, Dave A, Straga T, Souzeau E, Chataway T, et al. (2013) Identification of a novel oligomerization disrupting mutation in CRYAlphaA associated with congenital cataract in a South Australian family. Hum Mutat 34: 435-438.
73. Mackay DS 14Mackay DS, Andley UP, Shiels A (2003) Cell death triggered by a novel mutation in the alphaA-crystallin gene underlies autosomal dominant cataract linked to chromosome 21q. Eur J Hum Genet 11: 784-793.
74. Khan AO, Aldahmesh MA, Meyer B (2007) Recessive congenital total cataract with microcornea and heterozygote carrier signs caused by a novel missense CRYAA mutation (R54C). Am J Ophthalmol 144: 949-952.
75. Su D, Guo Y, Li Q, Guan L, Zhu S, et al. (2012) A novel mutation in CRYAA is associated with autosomal dominant suture cataracts in a Chinese family. Mol Vis 18: 3057-3063.

76. Santhiya ST, Soker T, Klopp N, Illig T, Prakash MV, et al. (2006) Identification of a novel, putative cataract-causing allele in CRYAA (G98R) in an Indian family. Mol Vis 12: 768-773.
77. Litt M, Kramer P, LaMorticella DM, Murphey W, Lovrien EW, et al. (1998) Autosomal dominant congenital cataract associated with a missense mutation in the human alpha crystallin gene CRYAA. Hum Mol Genet 7: 471-474.
78. Gu F, Luo W, Li X, Wang Z, Lu S, et al. (2008) A novel mutation in AlphaA-crystallin (CRYAA) caused autosomal dominant congenital cataract in a large Chinese family. Hum Mutat 29: 769.
79. Chen Q, Ma J, Yan M, Mothobi ME, Liu Y, et al. (2009) A novel mutation in CRYAB associated with autosomal dominant congenital nuclear cataract in a Chinese family. Mol Vis 15: 1359-1365.
80. Liu M, Ke T, Wang Z, Yang Q, Chang W, et al. (2006) Identification of a CRYAB mutation associated with autosomal dominant posterior polar cataract in a Chinese family. Invest Ophthalmol Vis Sci 47: 3461-3466.
81. Safieh LA, Khan AO, Alkuraya FS (2009) Identification of a novel CRYAB mutation associated with autosomal recessive juvenile cataract in a Saudi family. Mol Vis 15: 980-984.
82. Sacconi S, Feasson L, Antoine JC, Pecheux C, Bernard R, et al. (2012) A novel CRYAB mutation resulting in multisystemic disease. Neuromuscul Disord 22: 66-72.

83. Vicart P, Caron A, Guicheney P, Li Z, Prevost MC, et al. (1998) A missense mutation in the alphaB-crystallin chaperone gene causes a desmin-related myopathy. Nat Genet 20: 92-95.
84. Liu Y, Zhang X, Luo L, Wu M, Zeng R, et al. (2006) A novel alphaB-crystallin mutation associated with autosomal dominant congenital lamellar cataract. Invest Ophthalmol Vis Sci 47: 1069-1075.
85. Berry V, Francis P, Reddy MA, Collyer D, Vithana E, et al. (2001) Alpha-B crystallin gene (CRYAB) mutation causes dominant congenital posterior polar cataract in humans. Am J Hum Genet 69: 1141-1145.
86. Selcen D, Engel AG (2003) Myofibrillar myopathy caused by novel dominant negative alpha B-crystallin mutations. Ann Neurol 54: 804-810.
87. Pilotto A, Marziliano N, Pasotti M, Grasso M, Costante AM, et al. (2006) alphaB-crystallin mutation in dilated cardiomyopathies: low prevalence in a consecutive series of 200 unrelated probands. Biochem Biophys Res Commun 346: 1115-1117.
88. Inagaki N, Hayashi T, Arimura T, Koga Y, Takahashi M, et al. (2006) Alpha B-crystallin mutation in dilated cardiomyopathy. Biochem Biophys Res Commun 342: 379-386.
89. Devi RR, Yao W, Vijayalakshmi P, Sergeev YV, Sundaresan P, et al. (2008) Crystallin gene mutations in Indian families with inherited pediatric cataract. Mol Vis 14: 1157-1170.

90. van Veen T, van Winsen L, Crusius JB, Kalkers NF, Barkhof F, et al. (2003) [Alpha]B-crystallin genotype has impact on the multiple sclerosis phenotype. Neurology 61: 1245-1249.
91. Stoevring B, Frederiksen JL, Christiansen M (2007) CRYAB promoter polymorphisms: influence on multiple sclerosis susceptibility and clinical presentation. Clin Chim Acta 375: 57-62.
92. Meyer E, Rahman F, Owens J, Pasha S, Morgan NV, et al. (2009) Initiation codon mutation in betaB1-crystallin (CRYBB1) associated with autosomal recessive nuclear pulverulent cataract. Mol Vis 15: 1014-1019.
93. Cohen D, Bar-Yosef U, Levy J, Gradstein L, Belfair N, et al. (2007) Homozygous CRYBB1 deletion mutation underlies autosomal recessive congenital cataract. Invest Ophthalmol Vis Sci 48: 2208-2213.
94. Wang KJ, Wang S, Cao NQ, Yan YB, Zhu SQ (2011) A novel mutation in CRYBB1 associated with congenital cataract-microcornea syndrome: the p.Ser129Arg mutation destabilizes the betaB1/betaA3-crystallin heteromer but not the betaB1-crystallin homomer. Hum Mutat 32: E2050-2060.
95. Wang S, Zhao WJ, Liu H, Gong H, Yan YB (2013) Increasing betaB1-crystallin sensitivity to proteolysis caused by the congenital cataract-microcornea syndrome mutation S129R. Biochim Biophys Acta 1832: 302-311.
96. Mackay DS, Boskovska OB, Knopf HL, Lampi KJ, Shiels A (2002) A nonsense mutation in CRYBB1 associated with autosomal dominant cataract linked to human chromosome 22q. Am J Hum Genet 71: 1216-1221.

97. Yang J, Zhu Y, Gu F, He X, Cao Z, et al. (2008) A novel nonsense mutation in CRYBB1 associated with autosomal dominant congenital cataract. Mol Vis 14: 727-731.
98. Wang J, Ma X, Gu F, Liu NP, Hao XL, et al. (2007) A missense mutation S228P in the CRYBB1 gene causes autosomal dominant congenital cataract. Chin Med J (Engl) 120: 820-824.
99. Wang KJ, Wang BB, Zhang F, Zhao Y, Ma X, et al. (2011) Novel beta-crystallin gene mutations in Chinese families with nuclear cataracts. Arch Ophthalmol 129: 337-343.
100. Willoughby CE, Shafiq A, Ferrini W, Chan LL, Billingsley G, et al. (2005) CRYBB1 mutation associated with congenital cataract and microcornea. Mol Vis 11: 587-593.
101. Yao K, Li J, Jin C, Wang W, Zhu Y, et al. (2011) Characterization of a novel mutation in the CRYBB2 gene associated with autosomal dominant congenital posterior subcapsular cataract in a Chinese family. Mol Vis 17: 144-152.
102. Lou D, Tong JP, Zhang LY, Chiang SW, Lam DS, et al. (2009) A novel mutation in CRYBB2 responsible for inherited coronary cataract. Eye (Lond) 23: 1213-1220.
103. Santhiya ST, Kumar GS, Sudhakar P, Gupta N, Klopp N, et al. (2010) Molecular analysis of cataract families in India: new mutations in the CRYBB2 and GJA3 genes and rare polymorphisms. Mol Vis 16: 1837-1847.

104. Pauli S, Soker T, Klopp N, Illig T, Engel W, et al. (2007) Mutation analysis in a German family identified a new cataract-causing allele in the CRYBB2 gene. Mol Vis 13: 962-967.

105. Roshan M, Vijaya PH, Lavanya GR, Shama PK, Santhiya ST, et al. (2010) A novel human CRYGD mutation in a juvenile autosomal dominant cataract. Mol Vis 16: 887-896.

106. Litt M, Carrero-Valenzuela R, LaMorticella DM, Schultz DW, Mitchell TN, et al. (1997) Autosomal dominant cerulean cataract is associated with a chain termination mutation in the human beta-crystallin gene CRYBB2. Hum Mol Genet 6: 665-668.

107. Gill D, Klose R, Munier FL, McFadden M, Priston M, et al. (2000) Genetic heterogeneity of the Coppock-like cataract: a mutation in CRYBB2 on chromosome 22q11.2. Invest Ophthalmol Vis Sci 41: 159-165.

108. Vanita, Sarhadi V, Reis A, Jung M, Singh D, et al. (2001) A unique form of autosomal dominant cataract explained by gene conversion between beta-crystallin B2 and its pseudogene. J Med Genet 38: 392-396.

109. Yao K, Tang X, Shentu X, Wang K, Rao H, et al. (2005) Progressive polymorphic congenital cataract caused by a CRYBB2 mutation in a Chinese family. Mol Vis 11: 758-763.

110. Bateman JB, von-Bischhoffshaunsen FR, Richter L, Flodman P, Burch D, et al. (2007) Gene conversion mutation in crystallin, beta-B2 (CRYBB2) in a Chilean family with autosomal dominant cataract. Ophthalmology 114: 425-432.

111. Li FF, Zhu SQ, Wang SZ, Gao C, Huang SZ, et al. (2008) Nonsense mutation in the CRYBB2 gene causing autosomal dominant progressive polymorphic congenital coronary cataracts. Mol Vis 14: 750-755.
112. Liu BF, Liang JJ (2005) Interaction and biophysical properties of human lens Q155* betaB2-crystallin mutant. Mol Vis 11: 321-327.
113. Wang L, Lin H, Gu J, Su H, Huang S, et al. (2009) Autosomal-dominant cerulean cataract in a chinese family associated with gene conversion mutation in beta-B2-crystallin. Ophthalmic Res 41: 148-153.
114. Hansen L, Mikkelsen A, Nurnberg P, Nurnberg G, Anjum I, et al. (2009) Comprehensive mutational screening in a cohort of Danish families with hereditary congenital cataract. Invest Ophthalmol Vis Sci 50: 3291-3303.
115. Mothobi ME, Guo S, Liu Y, Chen Q, Yussuf AS, et al. (2009) Mutation analysis of congenital cataract in a Basotho family identified a new missense allele in CRYBB2. Mol Vis 15: 1470-1475.
116. Riazuddin SA, Yasmeen A, Yao W, Sergeev YV, Zhang Q, et al. (2005) Mutations in betaB3-crystallin associated with autosomal recessive cataract in two Pakistani families. Invest Ophthalmol Vis Sci 46: 2100-2106.
117. Santhiya ST, Shyam Manohar M, Rawlley D, Vijayalakshmi P, Namperumalsamy P, et al. (2002) Novel mutations in the gamma-crystallin genes cause autosomal dominant congenital cataracts. J Med Genet 39: 352-358.

118. Kannabiran C, Rogan PK, Olmos L, Basti S, Rao GN, et al. (1998) Autosomal dominant zonular cataract with sutural opacities is associated with a splice mutation in the betaA3/A1-crystallin gene. Mol Vis 4: 21.
119. Burdon KP, Wirth MG, Mackey DA, Russell-Eggitt IM, Craig JE, et al. (2004) Investigation of crystallin genes in familial cataract, and report of two disease associated mutations. Br J Ophthalmol 88: 79-83.
120. Zhu Y, Shentu X, Wang W, Li J, Jin C, et al. (2010) A Chinese family with progressive childhood cataracts and IVS3+1G>A CRYBA3/A1 mutations. Mol Vis 16: 2347-2353.
121. Gu Z, Ji B, Wan C, He G, Zhang J, et al. (2010) A splice site mutation in CRYBA1/A3 causing autosomal dominant posterior polar cataract in a Chinese pedigree. Mol Vis 16: 154-160.
122. Bateman JB, Geyer DD, Flodman P, Johannes M, Sikela J, et al. (2000) A new betaA1-crystallin splice junction mutation in autosomal dominant cataract. Invest Ophthalmol Vis Sci 41: 3278-3285.
123. Reddy MA, Bateman OA, Chakarova C, Ferris J, Berry V, et al. (2004) Characterization of the G91del CRYBA1/3-crystallin protein: a cause of human inherited cataract. Hum Mol Genet 13: 945-953.
124. Ferrini W, Schorderet DF, Othenin-Girard P, Uffer S, Heon E, et al. (2004) CRYBA3/A1 gene mutation associated with suture-sparing autosomal dominant congenital nuclear cataract: a novel phenotype. Invest Ophthalmol Vis Sci 45: 1436-1441.

125. Qi Y, Jia H, Huang S, Lin H, Gu J, et al. (2004) A deletion mutation in the betaA1/A3 crystallin gene (CRYBA1/A3) is associated with autosomal dominant congenital nuclear cataract in a Chinese family. Hum Genet 114: 192-197.

126. Lu S, Zhao C, Jiao H, Kere J, Tang X, et al. (2007) Two Chinese families with pulverulent congenital cataracts and deltaG91 CRYBA1 mutations. Mol Vis 13: 1154-1160.

127. Xu J, Wong C, Tan X, Jing H, Zhou G, et al. (2010) Decreasing the homodimer interaction: a common mechanism shared by the deltaG91 mutation and deamidation in betaA3-crystallin. Mol Vis 16: 438-444.

128. Billingsley G, Santhiya ST, Paterson AD, Ogata K, Wodak S, et al. (2006) CRYBA4, a novel human cataract gene, is also involved in microphthalmia. Am J Hum Genet 79: 702-709.

129. Zhou G, Zhou N, Hu S, Zhao L, Zhang C, et al.(2010) A missense mutation in CRYBA4 associated with congenital cataract and microcornea. Mol Vis 16: 1019-1024.

130. Stephan DA, Gillanders E, Vanderveen D, Freas-Lutz D, Wistow G, et al. (1999) Progressive juvenile-onset punctate cataracts caused by mutation of the gammaD-crystallin gene. Proc Natl Acad Sci U S A 96: 1008-1012.

131. Gu F, Li R, Ma XX, Shi LS, Huang SZ, et al. (2006) A missense mutation in the gammaD-crystallin gene CRYGD associated with autosomal dominant congenital cataract in a Chinese family. Mol Vis 12: 26-31.

132. Zhang LY, Gong B, Tong JP, Fan DS, Chiang SW, et al. (2009) A novel gammaD-crystallin mutation causes mild changes in protein properties but leads to congenital coralliform cataract. Mol Vis 15: 1521-1529.
133. Hilal L, Nandrot E, Belmekki M, Chefchaouni M, El Bacha S, et al. (2002) Evidence of clinical and genetic heterogeneity in autosomal dominant congenital cerulean cataracts. Ophthalmic Genet 23: 199-208.
134. Nandrot E, Slingsby C, Basak A, Cherif-Chefchaouni M, Benazzouz B, et al. (2003) Gamma-D crystallin gene (CRYGD) mutation causes autosomal dominant congenital cerulean cataracts. J Med Genet 40: 262-267.
135. Mackay DS, Andley UP, Shiels A (2004) A missense mutation in the gammaD crystallin gene (CRYGD) associated with autosomal dominant "coral-like" cataract linked to chromosome 2q. Mol Vis 10: 155-162.
136. Xu WZ, Zheng S, Xu SJ, Huang W, Yao K, et al. (2004) Autosomal dominant coralliform cataract related to a missense mutation of the gammaD-crystallin gene. Chin Med J (Engl) 117: 727-732.
137. Shentu X, Yao K, Xu W, Zheng S, Hu S, et al. (2004) Special fasciculiform cataract caused by a mutation in the gammaD-crystallin gene. Mol Vis 10: 233-239.
138. Khan AO, Aldahmesh MA, Ghadhfan FE, Al-Mesfer S, Alkuraya FS (2009) Founder heterozygous P23T CRYGD mutation associated with cerulean (and coralliform) cataract in 2 Saudi families. Mol Vis 15: 1407-1411.

139. Yang G, Xiong C, Li S, Wang Y, Zhao J A (2011) recurrent mutation in CRYGD is associated with autosomal dominant congenital coralliform cataract in two unrelated Chinese families. Mol Vis 17: 1085-1089.
140. Sun W, Xiao X, Li S, Guo X, Zhang Q (2011) Mutation analysis of 12 genes in Chinese families with congenital cataracts. Mol Vis 17: 2197-2206.
141. Wang L, Chen X, Lu Y, Wu J, Yang B, et al. (2011) A novel mutation in gammaD-crystallin associated with autosomal dominant congenital cataract in a Chinese family. Mol Vis 17: 804-809.
142. Kmoch S, Brynda J, Asfaw B, Bezouska K, Novak P, et al. (2000) Link between a novel human gammaD-crystallin allele and a unique cataract phenotype explained by protein crystallography. Hum Mol Genet 9: 1779-1786.
143. Wang B, Yu C, Xi YB, Cai HC, Wang J, et al. (2010) A novel CRYGD mutation (p.Trp43Arg) causing autosomal dominant congenital cataract in a Chinese family. Hum Mutat 32: E1939-1947.
144. Santana A, Waiswol M, Arcieri ES, Cabral de Vasconcellos JP, Barbosa de Melo M (2009) Mutation analysis of CRYAA, CRYGC, and CRYGD associated with autosomal dominant congenital cataract in Brazilian families. Mol Vis 15: 793-800.
145. Heon E, Priston M, Schorderet DF, Billingsley GD, Girard PO, et al. (1999) The gamma-crystallins and human cataracts: a puzzle made clearer. Am J Hum Genet 65: 1261-1267.

146. Li F, Wang S, Gao C, Liu S, Zhao B, et al. (2008) Mutation G61C in the CRYGD gene causing autosomal dominant congenital coralliform cataracts. Mol Vis 14: 378-386.
147. Zhang W, Cai HC, Li FF, Xi YB, Ma X, et al. (2011) The congenital cataract-linked G61C mutation destabilizes gammaD-crystallin and promotes non-native aggregation. PLoS One 6: e20564.
148. Messina-Baas OM, Gonzalez-Huerta LM, Cuevas-Covarrubias SA (2006) Two affected siblings with nuclear cataract associated with a novel missense mutation in the CRYGD gene. Mol Vis 12: 995-1000.
149. Zhang LY, Yam GH, Fan DS, Tam PO, Lam DS, et al. (2007) A novel deletion variant of gammaD-crystallin responsible for congenital nuclear cataract. Mol Vis 13: 2096-2104.
150. Ren Z, Li A, Shastry BS, Padma T, Ayyagari R, et al. (2000) A 5-base insertion in the gammaC-crystallin gene is associated with autosomal dominant variable zonular pulverulent cataract. Hum Genet 106: 531-537.
151. Yao K, Jin C, Zhu N, Wang W, Wu R, et al. (2008) A nonsense mutation in CRYGC associated with autosomal dominant congenital nuclear cataract in a Chinese family. Mol Vis 14: 1272-1276.
152. Gonzalez-Huerta LM, Messina-Baas OM, Cuevas-Covarrubias SA (2007) A family with autosomal dominant primary congenital cataract associated with a CRYGC mutation: evidence of clinical heterogeneity. Mol Vis 13: 1333-1338.

153. Zhang L, Fu S, Ou Y, Zhao T, Su Y, et al. (2009) A novel nonsense mutation in CRYGC is associated with autosomal dominant congenital nuclear cataracts and microcornea. Mol Vis 15: 276-282.
154. Kumar M, Agarwal T, Khokhar S, Kaur P, Roy TS, et al. (2011) Mutation screening and genotype phenotype correlation of alpha-crystallin, gamma-crystallin and GJA8 gene in congenital cataract. Mol Vis 17: 693-707.
155. Sun H, Ma Z, Li Y, Liu B, Li Z, et al. (2005) Gamma-S crystallin gene (CRYGS) mutation causes dominant progressive cortical cataract in humans. J Med Genet 42: 706-710.
156. Vanita V, Singh JR, Singh D, Varon R, Sperling K (2009) Novel mutation in the gamma-S crystallin gene causing autosomal dominant cataract. Mol Vis 15: 476-481.
157. Kumar LV, Ramakrishna T, Rao CM (1999) Structural and functional consequences of the mutation of a conserved arginine residue in alphaA and alphaB crystallins. J Biol Chem 274: 24137-24141.
158. Fu L, Liang JJ (2003) Alteration of protein-protein interactions of congenital cataract crystallin mutants. Invest Ophthalmol Vis Sci 44: 1155-1159.
159. Brown Z, Ponce A, Lampi K, Hancock L, Takemoto L (2007) Differential binding of mutant (R116C) and wildtype alphaA crystallin to actin. Curr Eye Res 32: 1051-1054.
160. Murugesan R, Santhoshkumar P, Sharma KK (2007) Cataract-causing alphaAG98R mutant shows substrate-dependent chaperone activity. Mol Vis 13: 2301-2309.

161. Bova MP, Yaron O, Huang Q, Ding L, Haley DA, et al. (1999) Mutation R120G in alphaB-crystallin, which is linked to a desmin-related myopathy, results in an irregular structure and defective chaperone-like function. Proc Natl Acad Sci U S A 96: 6137-6142.

162. Li H, Li C, Lu Q, Su T, Ke T, et al. (2008) Cataract mutation P20S of alphaB-crystallin impairs chaperone activity of alphaA-crystallin and induces apoptosis of human lens epithelial cells. Biochim Biophys Acta 1782: 303-309.

163. Peterson CA, Piatigorsky J (1986) Preferential conservation of the globular domains of the beta A3/A1-crystallin polypeptide of the chicken eye lens. Gene 45: 139-147.

164. Evans P, Wyatt K, Wistow GJ, Bateman OA, Wallace BA, et al. (2004) The P23T cataract mutation causes loss of solubility of folded gammaD-crystallin. J Mol Biol 343: 435-444.

165. Pande A, Zhang J, Banerjee PR, Puttamadappa SS, Shekhtman A, et al. (2009) NMR study of the cataract-linked P23T mutant of human gammaD-crystallin shows minor changes in hydrophobic patches that reflect its retrograde solubility. Biochem Biophys Res Commun 382: 196-199.

166. Pande A, Pande J, Asherie N, Lomakin A, Ogun O, et al. (2001) Crystal cataracts: human genetic cataract caused by protein crystallization. Proc Natl Acad Sci U S A 98: 6116-6120.

167. Pande A, Pande J, Asherie N, Lomakin A, Ogun O, et al. (2000) Molecular basis of a progressive juvenile-onset hereditary cataract. Proc Natl Acad Sci U S A 97: 1993-1998.

168. Pande A, Gillot D, Pande J (2009) The cataract-associated R14C mutant of human gamma D-crystallin shows a variety of intermolecular disulfide cross-links: a Raman spectroscopic study. Biochemistry 48: 4937-4945.

169. Fu L, Liang JJ (2002) Conformational change and destabilization of cataract gammaC-crystallin T5P mutant. FEBS Lett 513: 213-216.

170. Liu BF, Song S, Hanson M, Liang JJ (2008) Protein-protein interactions involving congenital cataract T5P gammaC-crystallin mutant: a confocal fluorescence microscopy study. Exp Eye Res 87: 515-520.

171. Talla V, Narayanan C, Srinivasan N, Balasubramanian D (2006) Mutation causing self-aggregation in human gammaC-crystallin leading to congenital cataract. Invest Ophthalmol Vis Sci 47: 5212-5217.

172. Talla V, Srinivasan N, Balasubramanian D (2008) Visualization of in situ intracellular aggregation of two cataract-associated human gamma-crystallin mutants: lose a tail, lose transparency. Invest Ophthalmol Vis Sci 49: 3483-3490.

173. Amaya L, Taylor D, Russell-Eggitt I, Nischal KK, Lengyel D (2003) The morphology and natural history of childhood cataracts. Surv Ophthalmol 48: 125-144.

174. Krishnamurthy R, Vanderveen DK (2008) Infantile cataracts. Int Ophthalmol Clin 48: 175-192.

175. Lambert SR, Drack AV (1996) Infantile cataracts. Surv Ophthalmol 40: 427-458.
176. Putzar L, Hotting K, Roder B (2010) Early visual deprivation affects the development of face recognition and of audio-visual speech perception. Restor Neurol Neurosci 28: 251-257.
177. Birch EE, Cheng C, Stager DR, Jr., Weakley DR, Jr., Stager DR, Sr. (2009) The critical period for surgical treatment of dense congenital bilateral cataracts. Journal of American Association for Pediatric Ophthalmology and Strabismus 13: 67-71.
178. Purkiss AG, Bateman OA, Goodfellow JM, Lubsen NH, Slingsby C (2002) The X-ray crystal structure of human gamma S-crystallin C-terminal domain. J Biol Chem 277: 4199-4205.
179. Wu Z, Delaglio F, Wyatt K, Wistow G, Bax A (2005) Solution structure of (gamma)S-crystallin by molecular fragment replacement NMR. Protein Sci 14: 3101-3114.
180. Augusteyn RC, Chandrasekher G, Ghiggino KP, Vassett P (1994) Probing the microenvironments of tryptophan residues in the monomeric crystallins of the bovine lens. Biochim Biophys Acta 1205: 89-96.
181. Rosen CG, Weber G (1969) Dimer formation from 1-amino-8-naphthalenesulfonate catalyzed by bovine serum albumin. A new fluorescent molecule with exceptional binding properties. Biochemistry 8: 3915-3920.

182. Sutter M, Oliveira S, Sanders NN, Lucas B, van Hoek A, et al. (2007) Sensitive spectroscopic detection of large and denatured protein aggregates in solution by use of the fluorescent dye Nile Red. J Fluoresc 17: 181-192.

183. Levine III H (1999) Quantification of β-sheet amyloid fibril structures with Thioflavin T. Methods Enzymol 309: 274- 284.

184. Greene Jr. RF, Pace CN (1974) Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme, α-chymotrypsin and β-lactoglobulin. J Biol Chem 249: 5388-5393.

185. Clark AC, Sinclair JF, Baldwin TO (1993) Folding of bacterial luciferase involves a non-native heterodimeric intermediate in equilibrium with the native enzyme and the unfolded subunits. J Biol Chem 268: 10773-10779.

186. Arnold K, Bordoli L, Kopp J, Schwede T (2006) The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. Bioinformatics 22: 195-201.

187. Brooks BR, Brooks 3rd CL, Mackerell AD, Nilsson L, Petrella RJ, et al. (2009) CHARMM: The Biomolecular Simulation Program. J Comp Chem 30: 1545-1615

188. Zhang Y. (2007) Template-based modeling and free modeling by I-TASSER in CASP7. Proteins 69: (Suppl 8): 108–117. doi:10.1002/prot.21702.

189. Karplus M. (2009) HMM-based protein structure prediction. Nucleic Acids Res 37: W492–W497. doi:10.1093/nar/gkp403.

190. Van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC. (2005) GROMACS: Fast, flexible and free. J. Comput. Chem. 26**:** 1701-1718.

191. Ji F, Jung J, Gronenborn AM (2012) Structural and biochemical characterization of the childhood cataract-associated R76S mutant of human gammaD-crystallin. Biochemistry 51: 2588-2596.

192. Saxena VP, Wetlaufert DB (1971) A new basis for interpreting the circular dichroic spectra of proteins. Proc. Nat. Acad. Sci. 68: 969-972.

193. Kosinski-Collins MS, King J (2003) In vitro unfolding, refolding, and polymerization of human gammaD crystallin, a protein involved in cataract formation. Protein Sci 12: 480-490.

194. Chen J, Flaugh SL, Callis PR, King J (2006) Mechanism of the highly efficient quenching of tryptophan fluorescence in human gammaD-crystallin. Biochemistry 45: 11552-11563.

195. Chen J, Toptygin D, Brand L, King J (2008) Mechanism of the efficient tryptophan fluorescence quenching in human gammaD-crystallin studied by time-resolved fluorescence. Biochemistry 47: 10705-10721.

196. Chen J, Callis PR, King J (2009) Mechanism of the very efficient quenching of tryptophan fluorescence in human gamma D- and gamma S-crystallins: the gamma-crystallin fold may have evolved to protect tryptophan residues from ultraviolet photodamage. Biochemistry 48: 3708-3716.

197. Brubaker WD, Freites JA, Golchert KJ, Shapiro RA, Morikis V, et al. (2011) Separating instability from aggregation propensity in gammaS-crystallin variants. Biophys J 100: 498-506.

198. Sandilands A, Hutcheson AM, Long HA, Prescott AR, Vrensen G, et al. (2002) Altered aggregation properties of mutant gamma-crystallins cause inherited cataract. EMBO J 21: 6005-6014.

199. Papanikolopoulou K, Mills-Henry I, Thol SL, Wang Y, Gross AA, et al. (2008) Formation of amyloid fibrils in vitro by human gammaD-crystallin and its isolated domains. Mol Vis 14: 81-89.

200. Wang Y, Petty S, Trojanowski A, Knee K, Goulet D, et al. (2010) Formation of amyloid fibrils in vitro from partially unfolded intermediates of human gammaC-crystallin. Invest Ophthalmol Vis Sci 51: 672-678.

201. Naiki H, Higuchi K, Hosokawa M, Takeda T (1989) Fluorometric determination of amyloid fibrils in vitro using the fluorescent dye, thioflavin T1. Anal Biochem 177: 244-249.

202. Sahin E, Jordan JL, Spatara ML, Naranjo A, Costanzo JA, et al. (2010) Computational design and biophysical characterization of aggregation-resistant point mutations for gammaD crystallin illustrate a balance of conformational stability and intrinsic aggregation propensity. Biochemistry 50: 628-639.

203. Evans SV (1993) SETOR: hardware-lighted three-dimensional solid model representations of macromolecules. J Mol Graph 11: 134-138, 127-138.

204. Hubbard SJ, Thornton JM (1993) NACCESS, Department of Biochemistry and Molecular Biology, University College, London, London, UK.

205. DeLano WL (2002) The PyMOL Molecular Graphics System http://www.pymol.org. Last accessed: 26th December, 2013.

206. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, et al. (2010) A method and server for predicting damaging missense mutations. Nat. Methods 7: 248-249.

207. Ma Z, Piszczek G, Wingfield PT, Sergeev YV, Hejtmancik JF (2009) The G18V CRYGS mutation associated with human cataracts increases gammaS-crystallin sensitivity to thermal and chemical stress. Biochemistry 48: 7334-7341.

208. Ji F, Jung J, Koharudin LM, Gronenborn AM (2013) The human W42R gammaD-crystallin mutant structure provides a link between congenital and age-related cataracts. J Biol Chem 288: 99-109.

209. Xu J, Wang S, Zhao WJ, Xi YB, Yan YB, et al. (2012) The congenital cataract-linked A2V mutation impairs tetramer formation and promotes aggregation of betaB2-crystallin. PLoS One 7: e51200.

210. Banerjee PR, Pande A, Patrosz J, Thurston GM, Pande J (2011) Cataract-associated mutant E107A of human gammaD-crystallin shows increased attraction to alpha-crystallin and enhanced light scattering. Proc Natl Acad Sci U S A 108: 574-579.

211. Graw J (2009) Mouse models of cataract. J Genet 88: 469-486.

212. Pande J (2011) Lecture at the annual meeting of the Association for Research in Vision and Ophthalmology (ARVO), Session 102, Ft. Lauderdale, FL, USA.

213. Benedek GB (1997) Cataract as a protein condensation disease: the Proctor Lecture. Invest Ophthalmol Vis Sci 38: 1911-1921.

214. Gunton JD Shiryayev A, Pagan DL (2007) Protein condensation: kinetic pathways to crystallization and disease. Cambridge University Press, New York.

Publications

1. **Vendra VP**, Garima Agarwal, Chandani S, Venu Talla, Narayanaswamy Srinivasan, Balasubramanian D (2013) Structural Integrity of the Greek Key Motif in βγ-crystallins is Vital for Central Eye Lens Transparency. PLoS One 8: e70336.
2. Srinivasu Karri, Ramesh Babu Kasetti, **Vendra VP**, Sushil Chandani, Dorairajan Balasubramanian (2013) Structural analysisof the mutant protein D26G of human λS crystalline, associated with coppock cataract. Mol Vis 19: 1231-1237. 2013 Jun 5;19:1231-7. Print 201 ; 19:1231-1237
3. **Vendra VP**, Chandani S, Balasubramanian D (2012) The mutation V42M distorts the compact packing of the human gamma-S-crystallin molecule, resulting in congenital cataract. PLoS One 7: e51401
4. **Vendra VP**, Balasubramanian D (2010) Structural and aggregation behavior of the human gammaD-crystallin mutant E107A, associated with congenital nuclear cataract. Mol Vis 16: 2822-2828.