

**CHARACTERIZATION OF THE UNFOLDING, REFOLDING, AND
AGGREGATION PATHWAYS OF TWO PROTEINS IMPLICATED IN
CATARACTOGENESIS: HUMAN GAMMA D AND HUMAN
GAMMA S CRYSTALLIN**

by

MELISSA SUE KOSINSKI-COLLINS

B.S. Biochemistry and Molecular Biology
University of Massachusetts at Amherst, 2000

Submitted to the Department of Biology
in partial fulfillment of the requirements for the degree of

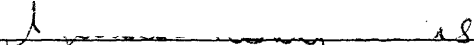
DOCTOR OF PHILOSOPHY
in Biochemistry
at the

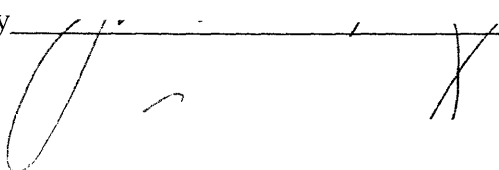
Massachusetts Institute of Technology

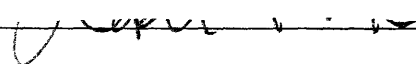
February 2005

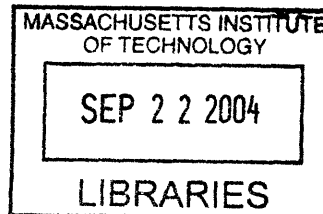
© 2005 Melissa S. Kosinski-Collins
All rights reserved

The author hereby grants MIT permission to reproduce and to distribute publicly paper
and electronic copies of this thesis document in whole or in part.

Signature of Author  _____
Department of Biology
September 2004

Certified by  _____
Jonathan King
Thesis Supervisor

Accepted by  _____
Stephen Bell
Chairman, Graduate Committee



ARCHIVES

**CHARACTERIZATION OF THE UNFOLDING, REFOLDING, AND
AGGREGATION PATHWAYS OF TWO PROTEINS IMPLICATED IN
CATARACTOGENESIS: HUMAN GAMMA D AND HUMAN
GAMMA S CRYSTALLIN**

by

MELISSA SUE KOSINSKI-COLLINS

Submitted to the department of Biology of the Massachusetts Institute of Technology on September 17, 2004 in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biochemistry

ABSTRACT

Human γ D crystallin (H γ D-Crys) and human γ S crystallin (H γ S-Crys), are major proteins of the human eye lens and are components of cataracts. H γ D-Crys is expressed early in life in the lens cortex while H γ S-Crys is expressed throughout life in the lens epithelial cells. Both are primarily β -sheet proteins made up of four Greek keys separated into two domains and display 69% sequence similarity.

The unfolding and refolding of H γ D-Crys and H γ S-Crys have been characterized as a function of guanidinium hydrochloride (GdnHCl) concentration at neutral pH and 37°C, using intrinsic tryptophan fluorescence to monitor *in vitro* folding. Equilibrium unfolding and refolding experiments with GdnHCl showed unfolded protein is more fluorescent than its native counter-part despite the absence of metal or ion-tryptophan interactions in both of these proteins. This fluorescence quenching may influence the lens response to ultraviolet light radiation or the protection of the retina from ambient ultraviolet damage.

Wild-type H γ D-Crys exhibited reversible refolding above 1.0 M GdnHCl. Aggregation of refolding intermediates of H γ D-Crys was observed in both equilibrium and kinetic refolding processes. The aggregation pathway competed with productive refolding at denaturant concentrations below 1.0 M GdnHCl, beyond the major conformational transition region. H γ S-Crys, however, exhibited a two-state reversible unfolding and refolding with no evidence of aggregation. Atomic force microscopy of H γ D-Crys samples under aggregating conditions revealed ordered fiber structures that could recruit H γ S-Crys to the aggregate.

To provide fluorescence reporters for each quadrant of H γ D-Crys, triple mutants each containing three tryptophan to phenylalanine substitutions and one native tryptophan have been constructed and expressed. Trp68-only and Trp156-only retained the quenching pattern of wild-type H γ D-Crys.

During equilibrium refolding/unfolding, the tryptophan fluorescence signals indicated that domain I (W42-only and W68-only) unfolded at lower concentrations of GdnHCl than domain II (W130-only and W156-only). Kinetic analysis of both the unfolding and refolding of the triple mutant tryptophan proteins identified an intermediate along the H γ D-Crys folding pathway with domain I unfolded and domain II

intact. This species is a candidate for the partially folded intermediate in the *in vitro* aggregation pathway of H γ D-Crys.

An N143D deamination post-translational modification has recently been identified in H γ S-Crys that is present in high concentrations in insoluble protein removed from cataractous lenses. The presence of the N143D mutation did not significantly affect the equilibrium or kinetic properties of H γ S-Crys indicating that this mutation is unlikely to be involved in protein destabilization during cataract formation *in vivo*.

The method in which H γ D-Crys aggregates on its own and engages neighboring molecules in the polymerization process *in vitro* may provide insight into the process of cataractogenesis *in vivo*.

Thesis Supervisor: Dr. Jonathan King, Professor of Biology

2
copies of
a leases look

ACKNOWLEDGEMENTS

In general, I find these things hard to write. How do you say thank you to all those who helped you over the course of four years? The truth is, you can never truly thank people by putting down a name on a piece of paper and saying, "There it is. Consider yourself thanked." I understand I will never actually be able to express the gratitude I feel towards everyone that has been a part of my life over the last four years, but please take it as a consolation that I recognize your presence in my life and in my work and thank you for it. Even the fact that you are reading this now means that I have not been forgotten and the countless hours I spent working on this thesis were not in vain.

I would like to thank Jonathan King, my thesis advisor, for his help and support. Thanks for making me see the "alternative hypothesis" and for forcing me to recognize that you cannot throw a ΔG and m -value on every protein in the world. You took me on in a time when you didn't have any other graduate students and let me be an active and important part of the lab despite my youth and data-amassing mind-set. Thanks for supporting my alternative career goals and being an advisor, teacher, and a friend. You have shown me that it is okay to go against the grain and be yourself no matter what the rest of the world thinks.

Shannon. Major Non-Flaugh. Xenia Onatopp. Auntie Sha-sha. Where do I begin? You taught me some of the most important lessons I have learned in graduate school. How to share; how to be a scientist; how to teach; how to recycle; how to help; how to learn; how to listen; and, most importantly, how to be myself. You are truly the reason I am still here and the reason that I am following this non-conventional path through science. Thanks for being my Sha-sha, my little sister. I think you taught me more than I taught you on so many levels beyond this scientific box in which we live. You amaze me and are a person so incredible that I thought only existed in stories. You are a model in which I will forever try to mold myself and I will always harbor a little jealousy that I can never achieve. You will change the world someday, Shannon, I have no doubt in that. If you get discouraged at times, remember how much you changed my life and keep in mind it was for the better. Long live the green tape ball...

Veronica Zepeda, you are the best undergrad in the world. You were never a UROP, you were always the part-time grad student. Thank for all of your help with the γS studies and the single alanine mutants. Good luck in grad school, VZ.

Ishara Mills, I would have never passed my prelim without you. Thanks for being a sounding board for the past four years for all of my problems and complaints. I am psyched that you have joined the lab and "Team Crystallin." I couldn't think of a better person to take over this project. Go forth and clone, my dear, go forth and clone...

I would like to thank Dan Solis, Dr. Davide Marini, and Dr. Peter Weigele for their help with the AFM and TEM work and for spending hour upon hour with me in small windowless rooms. You all have the patience of saints.

I would also like to thank Dr. Stephen Raso for his guidance during the early stages of this project. You taught kinetic curve-fitting to someone who looked at you with a blank expression most of the time and you spent A LOT of time with me showing me how to be successful scientist. The irony is, you haven't been in the lab for 3 years, but I still feel I can count on you to help me out and you still "got my back."

I would like to thank all the past and present members of the King lab, Cammie Haase-Pettingell, Ajay Pande, Claire Ting, Welkin Pope, Ryan Simkovsky, Cindy Woolley, Jacqueline Piret, Patricia Clark, Mary Kate McBreyer, Perry Chou, and Kristen Cook, for their help and support. We certainly have a unique lab with a very exciting dynamic, and it made every day fun and interesting.

I would like to acknowledge the laboratories of Larry Takemoto and George Benedek for providing the protein material used in several of the experiments described in this thesis.

I would also like to thank the members of my thesis committee, Bob Sauer, Amy Keating, Vernon Ingram, and Jack Liang. I know your time is valuable and I appreciate willingness to assist and guide me through the final stages of my graduate school career.

The last but certainly not least thanks goes to my family. My parents have supported me unconditionally throughout my life. Mom, you are the most incredible teacher I have ever known and I hope to make you proud as I follow in your footsteps in my future career. John, you are my anchor. You pushed me to finish grad school even when I was on the edge of quitting. How can you be so strong? You amaze me and I love and thank you for it. Morgan, you are the most prized accomplishment of my grad school days. You light up my life and make every day a joy to live. I can't imagine life without you, your blue eyes, and puppy.

This research was supported by NIH grant GM17980 awarded to Jonathan King and an NIH training grant awarded to Melissa Kosinski-Collins.

"It is always useful to know where a friend and relation is, whether you want him or whether you don't."

A. A. Milne's Rabbit

"If possible, try to find a way to come downstairs that doesn't involve going bump, bump, bump on the back of your head."

Pooh's Little Instruction Manual

BIOGRAPHICAL NOTE
Melissa Sue Kosinski-Collins

Education and Training

Massachusetts Institute of Technology. Ph.D., 2004, Biochemistry.

University of Massachusetts, Amherst, B.S., 2000, Biochemistry and Molecular Biology;
Summa Cum Laude

Research and Professional Experience

2001-2004 Massachusetts Institute of Technology Department of Biology, Cambridge MA, Graduate Research Assistant in the laboratory of Prof. Jonathan King, Ph.D.

1999-2000 University of Massachusetts Department of Biochemistry and Molecular Biology, Amherst, MA, Undergraduate Research Assistant in the laboratory of Prof. Lila M. Gierasch, Ph.D.

1997-2000 University of Massachusetts Department of Chemistry, Amherst, MA, Undergraduate Research Assistant in the laboratory of Prof. William J. Vining, Ph.D.

1999-2000 University of Massachusetts Department of Plant and Soil Sciences, Amherst, MA, Undergraduate Research Assistant in the laboratory of Prof. Susan Han, Ph.D.

1996-1997 University of Massachusetts Department of Veterinary and Animal Sciences, Amherst, MA, Undergraduate Research Technician in the laboratory of Prof. Abel Ponce de Leon, Ph.D.

Publications

Kosinski-Collins, M.S., Flaugh, S.L., & King, J.A. 2004. "Probing folding and fluorescence quenching in human γ D crystallin Greek key domains using triple tryptophan mutant proteins." *Protein Sci.* **13**: 2223-2235

Kosinski-Collins, M.S. & King, J.A. 2003. "In vitro unfolding, refolding, and polymerization of human γ D crystallin, a protein involved in cataract formation." *Protein Sci.* **12**: 480-490.

Chemistry and Biochemistry Educational Software

Saunders General Biochemistry Interactive CD ROM, Saunders College Publishing & Archipelago Productions (2000).

Saunders Organic Chemistry Interactive CD ROM, Saunders College Publishing (2000).

Exploring the World of Plastics CD ROM, National Plastics Museum and Center (1999).

TABLE OF CONTENTS

Title Page.....	1
Abstract.....	2
Acknowledgements.....	4
Biographical Note.....	6
Table of Contents.....	7
List of Figures.....	12
List of Tables.....	13
List of Abbreviations.....	14
I. Introduction.....	16
A. The protein folding problem.....	16
i. The models of protein folding: framework versus hydrophobic collapse...17	
ii. Folding of α -helical proteins.....	17
iii. Folding of β -sheet proteins.....	18
iv. The parallel β -helix.....	19
v. Model β -sheet proteins.....	19
a. Central questions of this thesis.....	19
b. Two-state β -sheet protein folding.....	20
c. Intermediate stabilization by pH: IFABP.....	21
d. Using tryptophan as a conformational probe: CRABPI.....	22
e. Hydrogen bond network formation: IL-1 β	24
f. Aggregation-prone intermediates: IL-1 β	25
B. β -sheet proteins inside the living lens.....	27
i. The chaperone: α -crystallin.....	27
ii. The $\beta\gamma$ -crystallins.....	28
iii. Folding of the $\beta\gamma$ -crystallins:	
small β -sheet proteins important in cataract.....	31
a. Stability differences between β - and γ -crystallins.....	31
b. Assessing the importance of interface interactions.....	31
c. Oligomerization of the $\beta\gamma$ -crystallins.....	33

C. Cataracts as a protein deposition disease.....	34
D. Mechanisms of protein aggregation.....	35
i. Disordered protein polymerization.....	35
ii. Ordered protein polymerization.....	35
a. Aggregation caused by mutation.....	36
b. Domain swapping.....	40
c. Loop-sheet insertion.....	44
d. Amyloidosis.....	47
e. Prion diseases.....	53
II. Characterizing the Unfolding, Refolding, and Aggregation of Human γ D Crystallin, an Eye Lens Protein Implicated in Cataract Formation.....	56
A. Introduction.....	56
B. Materials and Methods.....	61
i. Expression and purification.....	61
ii. Circular dichroism.....	61
iii. Equilibrium refolding and unfolding.....	61
iv. Unfolding fluorescence kinetics.....	62
v. Refolding fluorescence kinetics.....	62
vi. Refolding solution turbidity kinetics.....	62
vii. Atomic force microscopy.....	63
viii. BisANS binding assay.....	63
C. Results.....	63
i. Purification.....	63
ii. Equilibrium unfolding and refolding.....	64
iii. Kinetic unfolding through a partially unfolded intermediate.....	70
iv. Competing productive and aggregation-prone refolding pathways	75
v. The self-associated aggregate has a fibril structure.....	76
vi. The aggregate contains exposed hydrophobic pockets.....	83
D. Discussion.....	83
i. Crystallin refolding and aggregation <i>in vitro</i>	83

ii.	Formation of filamentous aggregates.....	88
iii.	Cataract formation in the lens.....	90
III.	Assignment of Tryptophan Fluorescence of Human γ D Crystallin Using Triple Tryptophan to Phenylalanine Mutant Proteins.....	93
A.	Introduction.....	93
B.	Materials and Methods.....	94
i.	Cloning and site-directed mutagenesis.....	94
ii.	Expression and purification of triple mutants.....	97
iii.	Circular dichroism.....	97
iv.	Ultraviolet light absorbance.....	97
v.	Fluorescence emission spectra.....	98
C.	Results.....	98
i.	Purification of his-tagged crystallin.....	98
ii.	Structure assignment.....	99
iii.	Fluorescence spectra.....	100
D.	Discussion.....	108
i.	Expression and structure of crystallin proteins.....	108
ii.	The quenching of tryptophan emission in the native state.....	109
iii.	Tryptophan quenching and protection from UV radiation.....	109
IV.	Identification of an Intermediate During Unfolding and Refolding of Human γ D Crystallin Using Triple Tryptophan to Phenylalanine Mutant Proteins.....	113
A.	Introduction.....	113
B.	Materials and Methods.....	114
i.	Equilibrium refolding and unfolding.....	114
ii.	Unfolding fluorescence kinetics.....	114
iii.	Refolding fluorescence kinetics.....	115
C.	Results.....	116
i.	Equilibrium Refolding and Unfolding.....	116
ii.	Unfolding kinetics.....	120

iii.	Refolding kinetics.....	124
D.	Discussion.....	128
i.	Stability of crystallin triple tryptophan mutants.....	128
ii.	Unfolding and refolding intermediates.....	131
V.	Characterizing the Unfolding and Refolding Pathway of Human γ S Crystallin and N143D, a Deamidated Mutant Identified in Age-Onset Cataracts.....	132
A.	Introduction.....	132
B.	Materials and Methods.....	134
i.	Expression and purification.....	134
ii.	Circular dichroism.....	137
iii.	Ultraviolet light absorbance.....	137
iv.	Fluorescence emission spectroscopy.....	138
v.	Equilibrium unfolding and refolding.....	138
vi.	Unfolding fluorescence kinetics.....	139
vii.	Refolding fluorescence kinetics.....	139
viii.	Refolding native gel electrophoresis.....	140
C.	Results.....	140
i.	Purification and structure assignment of H γ S-Crys.....	140
ii.	Stability of human γ S crystallin.....	143
iii.	Unfolding kinetics.....	149
iv.	Refolding kinetics.....	150
E.	Discussion.....	156
i.	Aggregation and refolding in human γ S crystallin.....	156
ii.	Deamidation at position 143 in human γ S crystallin.....	156
iii.	Role of chaperonin recognition.....	162
VI.	Recruitment of Human γ S Crystallin to the Human γ D Crystallin Aggregation Pathway.....	163
A.	Introduction.....	163
B.	Materials and Methods.....	163

C. Results.....	164
D. Discussion.....	166
VII. Concluding Discussion.....	174
A. Aromatic-aromatic ring interactions in human γ D crystallin.....	174
B. Hysteresis and protein folding.....	178
C. Domain swapping as a mechanism of aggregation for human γ D crystallin.....	185
D. Concluding remarks.....	186
VIII. References.....	188
IX. Appendix.....	206
A. Protein parameters.....	207
B. Mutagenic primer table.....	207
C. Unfolding and refolding equilibrium data analysis.....	208
D. Unfolding and refolding kinetic curve fit calculations.....	211
E. Oligomeric ring structures of human γ D crystallin.....	213

LIST OF FIGURES

1-1 Schematic diagram of the lens.....	29
1-2 Proposed mechanism of disordered aggregation.....	37
1-3 Diagram of hemoglobin aggregation.....	37
1-4 Proposed mechanism of domain swapping.....	41
1-5 Ribbon structure of β -crystallin in domain swapped conformation.....	41
1-6 Proposed mechanism of loop-sheet insertion.....	45
1-7 Ribbon diagram of cleaved α 1-antitrypsin polymer.....	45
1-8 Proposed mechanism of amyloid formation.....	49
1-9 Proposed mechanism of transthyretin Oligomerization.....	49
2-1 Hypothesized threading structure of human γ D crystallin.....	59
2-2 Circular dichroism of human γ D crystallin.....	65
2-3 Equilibrium unfolding and refolding of human γ D crystallin.....	67
2-4 Equilibrium refolding as a function of incubation time.....	71
2-5 Fluorescence of kinetic unfolding of human γ D crystallin.....	73
2-6 Solution turbidity of kinetic refolding of human γ D crystallin.....	77
2-7 Fluorescence of kinetic refolding of human γ D crystallin.....	79
2-8 Atomic force microscopy of refolding human γ D crystallin.....	81
2-9 Bis-ANS binding of refolding human γ D crystallin.....	85
2-10 Model of human γ D crystallin refolding and aggregation.....	91
3-1 Ribbon diagram of human γ D crystallin.....	95
3-2 Far-UV circular dichroism of triple tryptophan human γ D crystallin.....	101
3-3 Ultraviolet light absorbance of triple tryptophan human γ D crystallin.....	103
3-4 Fluorescence spectra of triple tryptophan human γ D crystallin.....	105
3-5 Ribbon structures of “cage” residues in human γ D crystallin.....	111
4-1 Equilibrium unfolding and refolding of triple tryptophan human γ D crystallin.....	117
4-2 Unfolding kinetics triple tryptophan human γ D crystallin.....	121
4-3 Refolding kinetics of triple tryptophan human γ D crystallin.....	125
4-4 Schematic diagram of folding intermediates of human γ D crystallin.....	129
5-1 Ribbon diagram of human γ S crystallin.....	135

5-2	Fluorescence emission spectra of human γ S crystallin.....	141
5-3	Equilibrium unfolding and refolding of human γ S crystallin.....	145
5-4	Temperature denaturation of human γ S crystallin with CD.....	147
5-5	Unfolding kinetics of human γ S crystallin.....	151
5-6	Refolding kinetics of human γ S crystallin.....	153
5-7	Refolding native polyacrylamide gel of wild-type human γ S crystallin.....	157
5-8	Refolding native polyacrylamide gel of N143D human γ S crystallin.....	159
6-1	Coaggregation of human γ S crystallin and human γ D crystallin.....	167
6-2	Time dependent recruitment of human γ S crystallin to the γ D aggregate.....	169
6-3	Quantitation of γ S crystallin time-dependent recruitment to γ D aggregate.....	171
7-1	Ribbon diagram of the aromatic network of human γ D crystallin.....	175
7-2	Ribbon diagram of proline 154 and 162 of human γ S crystallin.....	181
9-1	Atomic force microscopy of aggregate fibers of human γ D crystallin.....	213
9-2	Atomic force microscopy of oligomeric pores of human γ D crystallin.....	215

LIST OF TABLES

2-1	Kinetic rate constants for unfolding and refolding wild-type H γ D-Crys.....	75
3-1	Fluorescence emission constants for triple tryptophan mutant H γ D-Crys.....	107
4-1	Equilibrium characterization of triple tryptophan mutant H γ D-Crys.....	119
4-2	Unfolding kinetic rate constants for triple tryptophan mutant H γ D-Crys.....	124
4-3	Refolding kinetic rate constants for triple tryptophan mutant H γ D-Crys.....	128
5-1	Equilibrium characterization of H γ S-Crys.....	144
5-2	Unfolding kinetic rate constants for H γ S-Crys.....	149
5-3	Refolding kinetic rate constants for H γ S-Crys.....	150
8-1	Comparison of residue differences between H γ S-Crys and H γ D-Crys.....	183

LIST OF ABBREVIATION

A β	Amyloid- β protein
AD	Alzheimer's disease
AFM	Atomic force microscopy
ANS	1-Anilinonaphthalene-8-sulfonate
APP	Amyloid Precursor Protein
bisANS	4,4' Dianilio-1,1'-binaphthyl-5,5'-disulfonic acid
CD	Circular dichroism
<i>C. elegans</i>	<i>Caenorhabditis elegans</i>
CRABPI	Cellular retinoic acid binding protein
DNA	Deoxyribonucleic acid
DTT	Dithioreitol
<i>E. Coli</i>	<i>Escherichia coli</i>
EDTA	Ethylenediaminetetraacetate
FALS	Familial lateral sclerosis
Far-UV	Far ultraviolet light
FPLC	Fast phase liquid chromatography
$\Delta G_{(H_2O)}$	Gibb's free energy extrapolated to 0 M denaturant
GdnHCl	Guanidine hydrochloride
H-D	Hydrogen-deuterium exchange
H γ D-Crys	Human γ D crystallin
H γ S-Crys	Human γ S crystallin
HMW	High molecular weight
IFABPI	Intestinal fatty acid binding protein I
IL-1 β	Interleukin-1 β
iLBP	Intracellular lipid binding protein
IPTG	Isopropyl- β -D-thiogalactoside
LB	Luria-Bertani broth
LMW	Low molecular weight
m-value	Resistance to solvent denaturation

NaPO ₄	Sodium phosphate
Near-UV	Near ultraviolet light
Ni-NTA	Nickel nitrilotriacetic acid
NMR	Nuclear Magnetic Resonance
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PD	Parkinson's disease
PrP	Prion cellular protein
PrP ^{Sc}	Scrapie form of the prion protein
RNA	Ribonucleic acid
<i>S. Cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SDS	Sodium dodecyl sulfate
SF-flu	Stopped-flow fluorescence
SOD1	Cu/Zn superoxide dismutase 1
T _m	Melting temperature
TEM	Transmission electron microscopy
TRIS	Tris(hydroxymethyl)-amino methane
TSE	Transmissible spongiform encephalopathies
TTR	Transthyretin
UV	Ultraviolet light

CHAPTER I: INTRODUCTION

A. THE PROTEIN FOLDING PROBLEM

The biomedical research community has been inundated with new technologies and discoveries that seem to be multiplying at an exponential rate. Things that seemed completely unattainable just five or ten years ago have now been transformed into routine procedures that are taught to introductory biology majors and sold as ready-made kits from mass manufacturing companies. For example, RNA interference loops and microarray gene chips are customizable at a touch of a button, and a few minutes at a computer keyboard will align a newly discovered gene with hundred of possible functional cousins.

Perhaps the most striking advance in the scientific world has been in DNA manipulation and sequencing. DNA sequencing has become so accessible at large throughput facilities that the average incoming biochemistry student has never even seen the massive glass plates that were used historically for these experiments, never mind actually run a sequencing gel. The biomedical community has successfully sequenced entire genomes including many animal model organisms such as *C. elegans*, *S. Cerevisiae*, and the mouse. The majority of the human genome was successfully completed in 2001 by two competing groups (Lander et al. 2001 and Venter et al. 2001). We now possess every chromosomal coding sequence in the human cell. In essence, that means we know the sequence of every RNA and, perhaps most shockingly, a possible sequence for every protein that sustains human life! Unfortunately, unlike the sequencing of the human genome, there is no shot-gun method of folding a protein when faced with simply its amino acid sequence.

Protein sequence dictates structure and polypeptide structure controls protein function. We now have half of the code needed to break the mystery of how proteins fold. We have the sequence, we need to be able to understand and predict the formation of structure from this information. As early as 1960, Christian Anfinsen set out upon this lofty goal and subsequently won the Nobel Prize for his insights into protein folding as a sequence guided process. Numerous scientists have spent their entire careers trying to unravel the “second-half” of the genetic code. We are flying at light speed in other

scientific arenas, but in the world of protein folding, we are moving relatively at a snail's pace. How does a linear sequence of amino acids actually control and maintain a complex three-dimensional protein fold? Moreover, how does this complex fold become altered and manipulated in protein deposition disorders?

i. The models of protein folding: framework versus hydrophobic collapse

There are two conceptual models most frequently used to describe the events that occur in protein folding. In the framework model, isolated pieces of secondary structure form independently in regions of the polypeptide chain upon folding (Ptitsyn 1998, Goldberg 1969, Wetlaufer 1973, Jaenicke 1999). These isolated α -helices, β -sheets, and loops then come together forming the native tertiary contacts. Alternatively, in the hydrophobic collapse model, upon folding, the polypeptide chain undergoes a rapid compaction in which the hydrophobic residues come together and form a rigid core (Dill and Chan 1997, Onuchic et al. 1997, Eaton et al. 1997). The hydrophobic core provides a stabilized center around which the native secondary structure forms.

Some combination of both of these mechanisms is probably applicable for most proteins containing α -helices and β -sheets. One reasonable approach to tackling the grand problem of protein folding is to understand the kinetic and thermodynamic properties of each class of secondary structure separately.

ii. Folding of α -helical proteins

Some headway has been made in understanding the propensity of an amino acid sequence to adopt an α -helical conformation (Muñoz and Serrano 1995). There are 3.6 residues per turn in an α -helix, and this means that within an isolated helix, one amino acid will interact only with other amino acids within a four residue linear distance of itself. In cases in which short-range interactions of i to $i+4$ between residues are dominant, estimations of secondary structure may be determined from knowledge of linear sequence interactions. These local interactions allow a rapid backbone hydrogen-

bonding network to be established for α -helices that results in complete folding within the millisecond time scale (Capaldi and Radford 1998).

For sequences of amino acids within a polypeptide chain in which local interactions dominate, we can generally assess the ability and likelihood that the sequence in question will form an α -helix based on the type of amino acids involved and the “capping” residues at the end of the sequence. Specifically, our ability to predict coiled-coil domains or the collagen triple helix fold is making progress (Hu et al. 1990, Brodsky and Ramshaw 1997).

In the cases of coiled-coils or four helix bundles, interactions between helices are reasonably well understood as well. The heptad repeat is a predictable motif observed in these types of interacting helices controlling individual helix stability and interaction energies (Cohen and Parry 1990, O’Shea et al. 1991). The stabilities and conformations of four-helix bundles are controlled by local and capping interactions as described above, together with the “knobs into holes” packing between helices (Dunker and Jones 1978). Successful *de novo* design of four-helix bundles has been achieved based on knowledge of hydrophobic packing, conformational entropy, interhelical turn residues, and helix-dipole interactions (Kamtekar and Hecht 1995). Specifically, the structure and stability of hydrophobic point mutants for homodimeric and heterodimeric coiled-coils may be calculated accurately based on side-chain packing interactions (Keating et al. 2001).

ii. Folding of β -sheet proteins

With the progress in understanding α -helical conformations, many laboratories have now turned their attention away to a more elusive problem; β -sheet folding. To understand β -sheet protein folding, we cannot simply look at the primary sequence of amino acids in a protein and predict what interaction will be present in the final β -fold. The global sequence control of residues involved in β -sheets could perhaps extend from residue i to residue $i+4$ or from residue i to residue $i+204$. In other words, one amino acid may be conformationally stabilized into a β -sheet structure by interaction with others residues throughout the entire rest of the protein molecule.

iii. The parallel β -helix

Accurate structure prediction of all β proteins has been most successful for parallel β -helices, structures in which the β -strands wind around upon each other in coil-like fashion similar to the processive nature of an α -helix (Jenkins and Pickersgill 2001). Although the distance between β -strands in the primary sequence may be variable with the insertion of large domains that project out from the stacked β -helix, many β -helical rungs in any given protein will be found in succession in the primary sequence. Many of these structures, like the P22 tailspike protein and pertactin from *Bordella Pertussis* contain a hydrophobic core of successive stacked nonpolar residues that runs throughout the length of the β -helix central cavity (Betts et al. 2004, Emsley et al. 1996). An accurate computational algorithm for β -helix prediction has been developed that takes advantage of the processive nature of the helical rungs and the regular placement of the hydrophobic stacked core (Bradley et al. 2001).

iv. Model β -sheet proteins

This question of how and why β -sheet proteins fold and aggregate has been addressed by several research laboratories through analysis of small primarily β -sheet proteins adopting relatively simple folds and by observing kinetic and thermodynamic trends between members of small molecule β -sheet protein families that demonstrate high levels of structural homology. Techniques such as circular dichroism (CD), nuclear magnetic resonance imaging (NMR), fluorescence imaging, hydrogen deuterium exchange (H-D), and stopped-flow fluorescence (SF-flu) have provided the scientific community with some insight into the folding pathways of β -sheet proteins, aggregation intermediates, and thermodynamic stability (Capaldi and Radford 1998).

b. Central questions of this thesis

In this thesis, I have investigated why the folded state of some proteins exhibit much greater stability under physiological conditions than others. In addition, I have

studied the process of β -sheet protein aggregation in a small model Greek key protein that is involved in the protein deposition disorder of cataract. The protein that is the primary focus of this thesis, human γ D crystallin (H γ D-Crys), remains in solution at high concentrations for more than fifty years in most adults. In addition to deciphering how its amino acid sequence drives it into a β -sheet fold, we hope to understand which aspects of the sequence are important for the maintenance of this long term stability.

Despite the stability of the crystallins, as well as many other proteins, aggregated or polymerized states form under physiological conditions. For the eye lens proteins this aggregated state causes cataract, a major disease of the elderly. As discussed below, considerable evidence points to partially folded or partially unfolded conformations, as the species responsible for the self-association reactions. A second focus of this thesis has been the identification of partially folded or partially unfolded conformers of H γ D-Crys that might be involved in aggregation reactions related to cataract.

In the sections following, I review selected aspects of β -sheet protein folding and misfolding and the lens crystallin literature to set the context for the experimental results described for the human γ -crystallins.

b. Two-state β -sheet protein folding

A central experimental question in the investigation of β -sheet folding, unfolding, and misfolding, has been the nature of the partially folded intermediate in these processes. Some of the first advances in the field of β -sheet protein folding came with the determination that small molecule β -sheet proteins were capable of undergoing reversible transition from the native folded to denatured, unfolded form. A reversible transition is two-state if does not exhibit populated folding intermediates in its conversion from the unfolded to the folded form. When the differential fluorescence or CD of a protein is monitored versus varying concentrations of urea or guanidinium hydrochloride, two-state protein denaturation/renaturation transitions only show one true folded and one true unfolded baseline.

Initially, it was believed that many proteins and their homologous family members exhibited two-state folding. Among β -sheet proteins, these included the Src

homology domains from α -spectrin, Fyn tyrosine kinase Src, bovine γ B crystallin, and phosphatidylinositol-3-kinase; the fibronectin type III domains including human fibronectin and the third domain of human tenascin; tendamistat; some members of the interleukin family; and certain proteins classified in the intracellular lipid binding protein family (Capaldi and Radford 1998, Jaenicke 1999).

Further work beyond simple equilibrium studies has revealed that many of these proteins experience transient folding intermediates that are not visible in conventional equilibrium denaturation. However, these types of two-state studies can be of importance in determination of the Gibbs free energy of folding (Fersht 1999). Equilibrium transitions of β -sheet molecules have been subsequently used in the dissection of the importance of interactions in β -sheet molecules using double mutant cycles and as a primary means of folding classification (Horovitz and Fersht 1990).

In order to successfully analyze the kinetics of folding in a model β -sheet molecule, it is necessary to extend one step beyond two-state equilibrium and to characterize the intermediates that form in the unfolding and refolding pathways of these molecules. Unfortunately, many intermediates are unstable, occur only transiently, and/or are seen in only minute concentrations. We will discuss several specific examples of model β -sheet proteins whose folding intermediates have been observed and well characterized.

c. Intermediate stabilization by pH: IFABP

One model family system that has been extensively analyzed in β -sheet protein studies is known as the intracellular lipid binding protein (iLBP) family. All of these proteins are β -barrel proteins containing solvent filled cavities that bind fatty acid (Banaszak et al. 1993). This particular family of proteins is quite unique in that it retains high levels of structural homology but relatively low levels of sequence homology. This system has proved to be very important for identifying different folding pathways that ultimately led to the same quaternary fold (Gunasekaran et al. 2001).

Dalessio and Ropson studied one member of the iLBP family, intestinal fatty acid binding protein (IFABP), to elucidate the folding dependence of this molecule on pH

(Dalessio and Ropson 1998). IFABP had an apparent two-state equilibrium-folding pathway as monitored by CD, fluorescence, and absorbance studies. SF-flu studies predicted that IFABP goes through one molten-globule-like intermediate structure in its unfolding pathway, but a much more denatured-like folding intermediate retaining little or no secondary structure as determined by CD in its refolding pathway. The unfolding and refolding of IFABP were the same at all studied pH ranges except for pH 4.0. IFABP subjected to pH 4.0 aggregated rapidly as the concentration of protein was increased. Soluble protein formed in these conditions produced an equilibrium folding intermediate with a fluorescence and CD signal that did not look like the signals characteristic of the molten globule folding intermediate characterized at pH 7.0. This transient folding structure was probably not a folding intermediate found in the normal physiological pathway. This result is in contrast to results reviewed by Ptitsyn stating that intermediates stabilized by low solution pH conditions are generally productive pathway intermediates (Ptitsyn 1998).

The pH effects on IFABP refolding and unfolding by SF-Flu were then characterized in detail (Dalessio and Ropson 1998). Kinetic folding experiments performed at pH 6-9 showed an initial burst phase. The authors postulated this burst was due to a global hydrophobic collapse that occurs only when the denaturant concentration was low. Refolding experiments performed at pH 10 showed that IFABP experienced triphasic kinetics alluding to the presence of an additional folding intermediate under these conditions. This new intermediate may be a structure involved in the productive folding pathway that was simply not present long enough for an accurate spectroscopic signal to be detected at lower pH levels. The authors tentatively concluded that changing the solution conditions in terms of pH, could in fact stabilize the transient intermediates in β -sheet protein folding pathways normally invisible to current spectroscopic technology.

d. Using tryptophan as a conformational probe: CRABPI

Another well characterized member of the iLBPs is cellular retinoic acid binding protein I (CRABPI). CRABPI was initially characterized as exhibiting a two-state

transition using urea denaturation and equilibrium analysis. CRABPI has three intrinsic tryptophans at positions 7, 87, and 109 that provided fluorescence reporters if excited at 280 nm with emission monitored at 350 nm. These tryptophans have been used to study the stability and folding of the molecule.

Double tryptophan mutants containing only one of the three native tryptophans demonstrated that Trp7 and Trp87 showed significant changes in fluorescence from the native to the unfolded state (Clark et al. 1996). Furthermore, Trp109 experienced a significant quenching when placed in its native environment. Later experiments performed by Steve Eyles showed that this quenching could be at least partially reversed by removal of Cys95 (Eyles and Gierasch 2000).

Further analysis of the tryptophan single mutants showed that wild-type CRABPI has three definable folding phases that can be monitored using SF-flu (Clark et al. 1996). Initially, a global hydrophobic collapse occurs characterized by all three tryptophans experiencing a more hydrophobic environment and the formation of the only α -helix present in the molecule. The helical structure was confirmed by NMR. A second phase then occurs during which the environments of all of the tryptophans change indicative of the formation of stable β -sheet structure. It is at this point at which the hydrogen-bonding network between the β -strands is established. Finally, ten percent of the CRABPI molecules were shown to exhibit a refolding alternate pathway resulting from of cis/trans isomerization of the Leu84-Pro85 bond (Eyles and Gierasch 2000).

Proline 85 is located in what is considered the hydrophobic core region of CRABPI. Angle restriction in the trans configuration somehow allows for the correct alignment of surrounding residues for a global hydrophobic collapse. Free rotation of the residue 85 ϕ angle observed with residues like alanine, severely disturbs the folding pathway and structural intermediates of CRABPI. These studies have now raised the question as to whether or not hydrophobic collapse can be a contributing factor in the folding of β -sheet proteins that have a hydrophobic cavity. NMR studies using peptides encoding the sequence of the loops of CRABPI have shown that they retain structure in solution without the stabilizing effect of the entire molecule (Rotondi and Gierasch 2003). Together these data suggest that β -sheet proteins require both a global hydrophobic collapse as well as local framework structure for proper folding.

e. Hydrogen bond network formation: Interleukin 1- β

To further analyze the mechanism of β -sheet formation, interleukin 1- β (IL-1 β) was studied in hydrogen-deuterium exchange (H-D) experiments (Varley et al. 1993). IL-1 β is a small β -sheet protein with a high-resolution structure that has been characterized by NMR. IL-1 β appears to be kinetically two-state in unfolding experiments, but refolding studies reveal a stable intermediate molten-globule structure may be formed that has 80-90% of the native state secondary structure (Varley et al. 1993).

Using stopped-flow far-UV CD and SF-flu of ANS binding and tryptophan fluorescence, the refolding kinetics of IL-1 β were characterized. In combination, by reporting on different phenomenon, these two techniques showed that three refolding intermediates of IL-1 β exist.

CD experiments showed that all secondary structure was fully formed at 6 seconds. The fluorescence data, however, demonstrated an initial burst phase faster than 6 seconds corresponding to a hydrophobic collapse and a subsequent slow folding process that continued for over 20 minutes. The discrepancies in the two data sets suggest that secondary structure formation is separable from initial folding collapse in IL-1 β and the final hydrophobic environment of tryptophan is not fully achieved until the slow orientation of small pieces of β -sheet occurs.

The kinetics of refolding of IL-1 β were then characterized with quenched-flow H-D exchange experiments with two-dimensional ^1H - ^{15}N correlation spectroscopy to monitor the rate of stable backbone amide hydrogen bonding of the protein (Varley et al. 1993). By observing amide peak protection versus incubation time, they were able to see that no stable hydrogen bonds representative of β -sheet formation were present until 1 second of refolding had passed. A preliminary hydrogen-bonding system was observed at 1 second although the hydrogen-bonding network for the protein did not fully form until 25 seconds. It was further demonstrated that all the early forming β -sheets were in areas of the molecule that interact with α -helices or other types of early folding structures independent of other β -sheets.

The initial folding states seen with fluorescence and CD were, therefore, not a result of the formation of true β -sheets, because these structures require the formation of hydrogen bonds that would be observable in the H-D exchange experiments. Fluorescence and CD data perhaps allude to an early forming molten globule folding form of IL-1 β that forms isolated unstable β -strands or β -like structures. These data demonstrate true β -sheet formation and inter-strand alignment of these types of secondary structures are perhaps the last features that form in the small molecule β -sheet protein folding pathways.

f. Aggregation prone intermediates: IL-1 β

Scientific attention has moved from mechanistic studies of folding and refolding to determining the mechanisms of off-pathway aggregation processes. In the past few years, studies have determined that the incorrect folding of some naturally occurring proteins normally present in native α -helical conformations to primarily β -sheet molecules may be implicated in the formation of harmful aggregation intermediates (Prusiner 2000). Some off-pathway β -sheet intermediates self-associate and result in the formation of amyloidogenic plaques and ultimately untimely death in higher organisms. For this reason, it is useful to study the aggregation process of known small β -sheet molecules to try and establish folding trends of β -sheet proteins.

The effects of single residue mutations on aggregation of IL-1 β during the folding process were studied by Finke et al. (2000). Ninety percent of the wild-type IL-1 β produced in *E. Coli* expression systems was found in the soluble portion of the centrifuged cellular lysate. Several different point mutations have been isolated and characterized in this protein that cause over 90% of the protein expressed in *E. Coli* to be in the form of inclusion bodies. Mutation of a charged lysine residue present in the native protein in a flexible loop at position 97 to a hydrophobic isoleucine resulted in the production of a folded protein that was thermodynamically more stable than the wild-type monomer form, but that was expressed primarily in inclusion bodies. This suggested that the mutation was acting on a folding intermediate or at the junction of productive folding and off-pathway aggregation. K97I appears to have the same slow folding step, the same

synthesis rate on the ribosome, and the same level of chaperone interaction and assistance as wild-type IL-1 β . Protein extracted from *in vitro* aggregates of K97I and *in vivo* inclusion bodies have demonstrated both systems produce insoluble proteins with nearly identical secondary structures as observed with Fourier transfer infrared spectroscopy.

Using SF-flu to measure kinetic refolding characteristics of IL-1 β showed that high protein concentrations changed the early events in refolding of the molecule (Finke 2000). As the concentration of K97I was increased, the transition from the unfolded to intermediate state was found to be the only folding step dependent on concentration. Similarly, stopped-flow light scattering demonstrated that whereas wild-type IL-1 β experienced no scattering no matter what concentration of protein was tested, K97I light scattering was present in samples of greater than 18 μ M. It was, therefore, hypothesized that aggregation was occurring in the unfolded to intermediate transition.

The self-association process of K97I most likely occurred between unfolded molecules under native conditions first experienced when initial refolding buffer was injected in the GdnHCl solution. The authors proposed a scheme for the possible aggregation mechanism of K97I. The proposed model relies on a nucleation mechanism for aggregation based on the fact that a critical protein concentration is needed for protein self-association. The hypothesis states that the flexible loop encompassing residues 86-99 represented an isolated "hydrophobic micro-domain" in the early peptide sampling process. The substitution of the hydrophobic residue isoleucine for the charged lysine, caused an increase in the hydrophobic surface area of the micro-domain. The hydrophobic area could more efficiently form and thus can more effectively interact with other exposed hydrophobic loops. The continued association of the hydrophobic regions over time could eventually lead to aggregation. Lysine 97 present in wild-type IL-1 β had a strong charge that destabilized the hydrophobic packing of the overall molecule preventing aggregation.

Such aggregation intermediates in β -sheet proteins stabilized by micro-domain packing may represent possible a therapeutic target for prevention of self-association protein folding diseases.

B. β -SHEET PROTEINS INSIDE THE LIVING LENS

The human lens contains a unique class of proteins that are important for both generalized protein folding studies and for their involvement in aggregation and cataract. The protein in the lens may be present at concentrations as high as 700 mg/mL and maintains a short-range order that allows light to be correctly diffracted and focused on the retina for proper vision (Figure 1-1) (Delaye and Tardieu 1983, Jaenicke 1999). The proteins in the lens must maintain high stability over the human lifetime as there is no protein turnover in this organ. The ubiquitous crystallins may be separated into two classes, α -crystallins and $\beta\gamma$ -crystallins.

i. The chaperone: α -crystallin

The two identified human α -crystallins are α A and α B crystallin, both of which are thought to perform chaperone-like functions in the human lens. The native conformation of lens α -crystallin appears to have a 3:1 stoichiometric ratio of α A to α B crystallin (Horwitz et al. 1999). Although the crystal structures of neither of the α -crystallins have been solved directly, light scattering and mass spectroscopy studies have shown that these proteins self-associate *in vitro* to form multimeric species composed of anywhere from 24 to 33 protein subunits (Aquilina et al. 2003).

Electron microscopy has shown that the multimers are spherical with hollow centers that are thought to bind destabilized or partially unfolded protein species nonspecifically (Haley et al. 2000). *In vitro* studies have shown that the presence of α -crystallin may inhibit aggregation of not only the $\beta\gamma$ -crystallins, but other aggregation-prone proteins such as the serpins and abrin (Reddy et al. 2002, Devlin et al. 2003).

The association of the α -crystallin spherical complex with a destabilized protein may prevent further self-association and aggregation. It is unclear whether the α -crystallins are present in the high molecular weight protein removed from cataractous lenses simply because they have associated with the aggregating material in a chaperone-related role or whether destabilization and/or mutation of the α -crystallin itself can initiate cataract formation (Clark and Muchowski 2000, Horwitz 2000, MacRae 2000).

Most recently, studies of normal and cataractous lenses showed that certain covalent damage and specific residue racemization in the α -crystallins are the same in both the HMW and LMW fractions in both lens types indicating that covalent damages in α -crystallin are unlikely the causative agent of cataract (Takemoto and Boyle 1998, Fujii et al. 2003).

ii. *The $\beta\gamma$ -crystallins*

The $\beta\gamma$ -crystallins are primarily structural proteins of the lens (Jaenicke 1999). The $\beta\gamma$ -crystallins found in the human lens are β A1 crystallin, β A3 crystallin, β A4 crystallin, β B1 crystallin, β B2 crystallin, γ C crystallin, γ D crystallin, and γ S crystallin. Sequence alignment of both DNA and amino acid sequence shows that the $\beta\gamma$ -crystallin family includes the two-domain protein, protein S of *Myxococcus xanthus* (Jaenicke 1999). The only known single domain $\beta\gamma$ -crystallin homologue is the dormant protein spherulin S3a from *Physarum polycephalum* (Jaenicke 1999).

The $\beta\gamma$ -crystallins are made up of two highly homologous domains that appear to be the result of gene duplication and fusion during evolution (Lubsen et al. 1988). Each domain is primarily β -sheet and is composed of two Greek keys separated by a linker of variable length. The β -crystallins have approximately 30% sequence similarity to the γ -crystallins. The β -crystallins are oligomeric in solution while the γ -crystallins are monomeric and the β -crystallins contain N- and C-terminal extensions absent in the γ -crystallins (Jaenicke et al. 1999).

The β -crystallins have a connecting peptide between the N- and C-terminal domains that is found in an extended conformation in the crystal structure, whereas the linker peptide in the γ -crystallins has a central glycine that adopts a V-like conformation (Blundell et al. 1993, Bax et al. 1990, Najmundin et al. 1993). The $\beta\gamma$ -crystallins have been implicated in the development of juvenile onset cataracts and these proteins are primary components of mature onset cataracts.

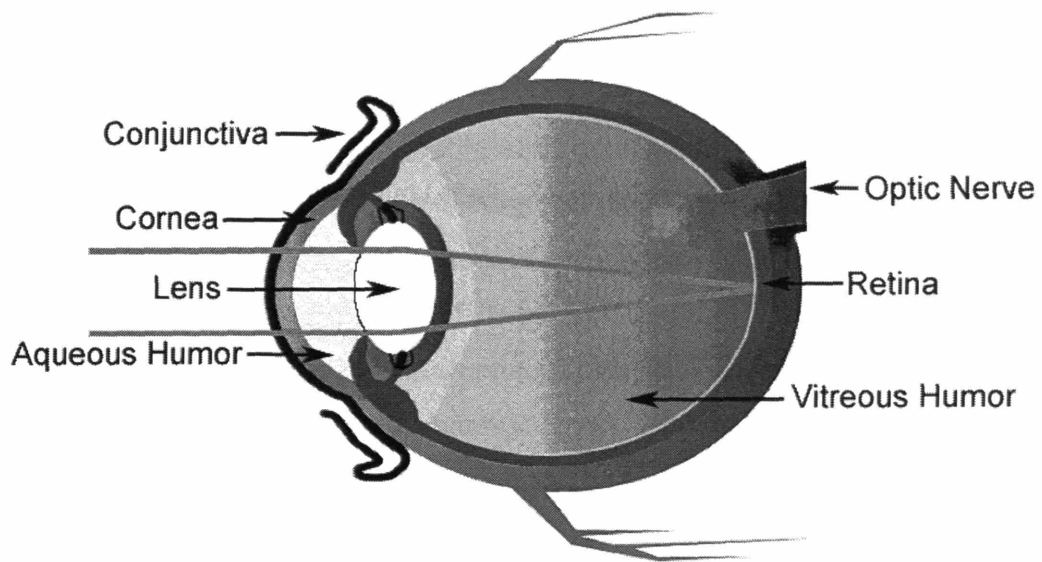


Figure 1-1

Schematic diagram of the human eye. The purple lines represent the passage of light through the lens and its refraction onto the retina. The different regions of the eye are labeled.

iii. Folding of the $\beta\gamma$ -crystallins: small β -sheet proteins important in cataract

The $\beta\gamma$ -crystallins have been studied for their role in cataract formation and for their importance as small model β -sheet proteins. The Greek key fold has a compact hydrophobic core at the center with highly organized, structural loops at either end. The $\beta\gamma$ -crystallins contain four Greek key motifs with two in each domain. These proteins have remarkable stability presumably reflecting their functional requirement to remain soluble at high concentration for long periods of time in the eye lens (Jaenicke 1999). The β -sheet and Greek key propensities of the $\beta\gamma$ -crystallins have been analyzed in an attempt to determine how and why β -sheet proteins fold and how the Greek key structure of these proteins is involved in protein stability. I will discuss several thorough studies of specific $\beta\gamma$ -crystallins not directly covered in later chapters.

a. Stability differences between β - and γ -crystallins

To determine the overall stability of the $\beta\gamma$ -crystallins, the difference in resistance to GdnHCl denaturation by the $\beta\gamma$ -crystallins was assessed spectroscopically by comparing the unfolding characteristics of human β B2 crystallin to human γ C crystallin (Fu and Liang 2002). Size exclusion chromatography indicated that intramolecular contacts between β B2 crystallin dimers were disrupted at very low concentration of GdnHCl. The thermodynamic stability of β B2 crystallin monomer was higher than that of γ C crystallin monomer when unfolding was measured with absorbance at 235 nm, fluorescence at 320 nm, and far-UV at 223 nm. Both proteins showed equilibrium transitions that were best fit to a three-state model demonstrating that both β B2 and γ C crystallin unfolded via a partially denatured intermediate.

b. Assessing the importance of interface interactions

The intermediate observed in these studies is likely a structure with one domain folded and one domain in native-like conformation. Given the two-domain structure of the $\beta\gamma$ -crystallins, it seems equally likely that the interface interactions between the

individual domains are important in $\beta\gamma$ -crystallin stability. Several research laboratories have attempted to resolve the structure of the partially unfolded intermediate and assess the contribution of the domain interface to the overall stability of the $\beta\gamma$ -crystallins (Jaenicke 1999).

Isolated N- and C-terminal domains of Protein S were constructed and studied to determine whether the domain interface was important to maintaining the stability of this $\beta\gamma$ -crystallin as well (Wenk et al. 1999). Protein S is a calcium binding two-domain $\beta\gamma$ -crystallin that has 172 amino acids (Wistow et al. 1985). Native Protein S showed independent domain folding where the N-terminal domain is less stable than the C-terminal domain. The stability and cooperativity of the intact molecule increased significantly in the presence of calcium although the stability of the N-terminal isolated domain was unaffected by the cation. Protein S showed a biphasic transition when denatured in urea at pH 2.0 while the isolated domains each show two-state transitions (Wenk and Mayr 1998, Wenk et al. 1999). Furthermore the midpoint of the denaturation transition for the isolated N- and C-terminal domains were located at lower concentrations of urea than wild type indicating a destabilization of the truncated proteins. The isolated domains did not exhibit any propensity to form pseudo-native molecules when mixed together in solution. From these studies, domain interactions appear to be important to the stability of Protein S (Wenk et al. 1999).

The importance of the domain interaction in γ -crystallin stabilization was further assessed in γ B crystallin from calf eye lens. When denatured in urea at pH 2.0, an unfolding intermediate was stabilized as shown by a biphasic equilibrium transition (Rudolph et al. 1990). Kinetic experiments revealed that the intermediate retained a similar structure during both unfolding and refolding analysis in which the C-terminal domain was unstructured and the N-terminal domain was in a rigid conformation. Construction of the isolated domains of γ B revealed that the N-terminal domain was more resistant to GdnHCl denaturation than the C-terminal domain (Mayr et al. 1997). Furthermore, similar to Protein S, the individual domain constructs of bovine γ B crystallin did not associate to form a pseudo-native molecule in solution. Formation of the interface was necessary to maintain native-like structure in the C-terminal domain.

c. Oligomerization of $\beta\gamma$ -crystallins

The $\beta\gamma$ -crystallins have also been studied as model β -sheet proteins prone to multimerization. The stability and aggregation characteristics of the β -crystallins were studied using temperature and urea induced unfolding experiments (Bateman et al. 2001). Oligomerization and precipitation of human β -crystallins were observed at high temperature. When hetero-oligomers were constructed by subunit exchange, β A1 was stabilized when it oligomerized with β B1 crystallin. These results suggest that the formation of hetero-oligomers of different subunits may add a further level of stabilization to the β -crystallins.

The N- and C-terminal extensions of the oligomeric β -crystallins are thought to be involved in protein multimerization (Norledge et al. 1996). The energetics of the domain to domain interactions were studied in recombinant mouse β A3 and β B2 crystallin by Sergeev et al. (2004). Both proteins were found in reversible monomer-dimer equilibrium from 5-35°C. The amount of time the two β -crystallins spent in the dimer state increased significantly at elevated temperatures. The overall entropy and enthalpy of binding in recombinant proteins with truncated N-termini increased as compared to wild-type. It is possible that post-translational truncations of β -crystallins in the eye lens make the protein more susceptible to association, aggregation, and cataract.

C. CATARACT AS A PROTEIN DEPOSITION DISEASE

Mature onset cataracts affect nearly 15% of the US population over 40 years of age and are the leading cause of blindness worldwide (NEI 2002). Pathological studies of cataractous lenses have revealed that cataracts are composed of protein aggregates that precipitate or polymerize in the lens region of the eye (Oyster 1999).

The aggregates removed from patients with mature onset cataracts are quite amorphous by light microscopy and do not appear to have distinct structures or sizes. The precipitation may occur anywhere in the lens from the nucleus to the dividing epithelial cells and can begin at any point in life although the prevalence of mature onset cataracts in the population increases sharply with age. The aggregated material extracted from aged lenses is composed of a variety of proteins most of which belong to the crystallin family rather than one dominant species (Hanson et al. 2000).

There are many covalent modifications in the high molecular weight (HMW) insoluble protein fraction of aged lenses including deamidations, methylations, acetylations, carbamylations, glycosylations, truncations, cysteine and methionine oxidations, and tryptophan ring cleavages (Hanson et al. 1998, Lin et al. 1998, Ma et al. 1998, Hanson et al. 2000, Clovis and Garland 2002, Lapko et al. 2002, Srivastava and Srivastava 2003, Srivastava et al. 2004). The prevalence of these modifications is markedly increased in the HMW material with respect to the low molecular weight (LMW) fractions. It is still unclear whether these modifications are the causative agents of the cataracts or occur after the protein aggregates. It is possible, however, that changes such as these locally or globally destabilize the polypeptide structure making it more susceptible to aggregation. For all these reasons, the mechanism of mature onset cataractogenesis is still unknown.

D. MECHANISMS OF PROTEIN AGGREGATION

i. Disordered protein polymerization

The insoluble, nonfibrous nature of the protein removed from aged cataractous lenses, makes cataract a possible candidate for non-specific protein aggregation (Figure 1-2). In this model, regions of several molecules would associate at random and eventually precipitate from solution.

Aggregation is rarely a random process in situations not affected by heat destabilization. Oftentimes, protein self association occurs via specific native or non-native interactions. Lysozyme, in particular, may undergo non-specific protein aggregation under certain conditions, but it may also form specific amyloidogenic interactions under similar conditions (Fischer et al. 1993, Cao et al. 2004).

ii. Ordered protein polymerization

Many diseases are now known to be caused by protein misfolding and aggregation. Afflictions such as Alzheimer's disease (AD), Creutzfeldt-Jakob's Disease, and Parkinson's disease (PD) are thought to be directly caused by the misfolding and aggregation of proteins in the brain (Caughey and Lansbury 2003). Although the exact mechanism of axon death and dementia are unknown, all of these diseases have distinct ordered protein aggregates that accumulate in the brain of afflicted patients. Despite the unstructured nature of the mature onset cataract observed in the light microscope, it seems reasonable that the species that initiates aggregate formation is composed of ordered inter-protein associations. Furthermore, the eye lens is packed with crystallins in their native states, so cataract is likely to form from native-like or partially unfolded proteins. Such species are likely to be sufficiently structured to yield ordered polymerization reactions.

Several studies have been performed on crystallins demonstrating these proteins are capable of ordered aggregation. Amyloid fibers can be produced from β - and γ -crystallins when they are refolded out of acidic conditions (Meehan et al. 2004). The

crystallins also bind thioflavin T and Congo red and show birefringence (Fredrickse 2000, Sandilands et al. 2002). These dyes are used to identify ordered β -sheet structures and amyloid. Atomic force microscopy of human γ D crystallin shows formation of distinct twisted fibrils that collapse to form fiber-like aggregates (Kosinski-Collins and King 2003). It is likely that aggregation of a few protein molecules in an organized self-association process may seed the subsequent aggregation of other polypeptides thus facilitating cataract formation.

Ordered aggregation can occur through native-like or nonnative interactions. Native-like aggregation mechanisms include domain swapping and loop-sheet insertion. Nonnative interactions that cause aggregation include the prion diseases, amyloid formation, and self-association caused by mutation.

a. Aggregation caused by mutation and association of native-like states

Sickle cell anemia is caused by a specific glutamine to valine mutation at position six of hemoglobin (Ingram 1956). This mutation substitutes a surface exposed polar amino acid with a hydrophobic residue. The mutated valine associates with an adjacent hemoglobin molecule at Phe85 and Leu88 (Bihoreau 1992). Because the native form of hemoglobin has four subunits, there are two valines that provide points of contact for other proteins on each side of the molecule. If nearby mutated proteins associate in a staggered network, multiple molecules can continue to associate via this valine-hydrophobic connection forming long extended, fiber-like threads (Figure 1-3). Under conditions of low oxygen, the long thin fibers grow across the red blood cell and distort the cell morphology to a “sickle-like” shape.

Aggregation of the native-like state occurring due to mutation of the crystallin proteins has been implicated in juvenile onset cataracts. Numerous heritable mutations in chromosome 22 in the human γ D crystallin and human γ C crystallin genes are involved in the pathogenesis of juvenile cataract (Heon et al. 1999; Kmoch et al. 2000, Ren et al. 2000, Santhiya et al. 2002, Nandrot, et al. 2003, Shentu et al. 2004).

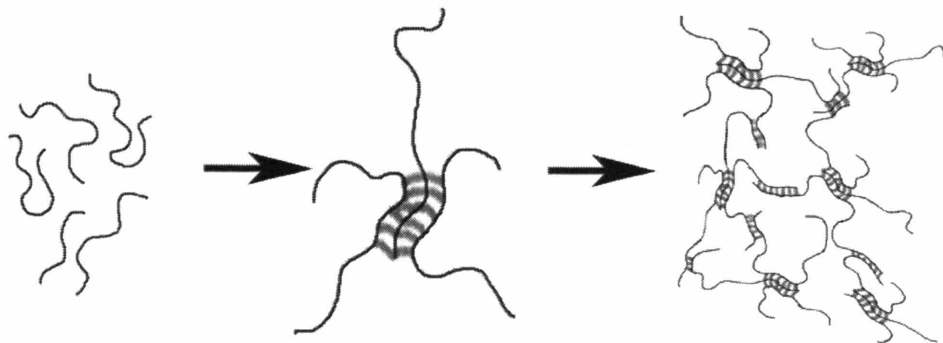


Figure 1-2

Schematic representation of one possible mechanism of disordered aggregation. Unfolded proteins close to one another in solution associate via non-native contacts (gray lines). The contacts may continue throughout many monomers creating an polymerized “mat” of protein.

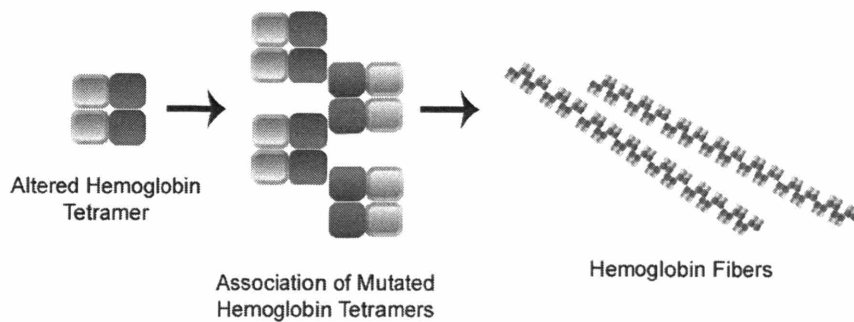


Figure 1-3

Schematic representation of hemoglobin tetramer association in sickle-cell anemia. The hemoglobin tetramers associate via a non-native valine which forms a nonpolar interaction with a hydrophobic pocket in another region of the molecule. The polymerization continues through multiple tetramer eventually forming a long, thin fiber.

An arginine to cysteine mutation at position 114 in HyD-Crys exposes a cysteine sulfur on the edge of domain II of the molecule. *In vitro* studies on the mutated protein showed that an intermolecular disulfide bond formed between Cys 114 of one molecule and the natively solvent exposed Cys 110 of an adjacent molecule (Stephan et al. 1999, Pande et al. 2000). High molecular weight oligomers were formed that retained the molecule's overall native state, but were covalently bonded through this non-native disulfide bonding network.

Two additional unique mutations in HyD-Crys are the R38H mutation that results in aculeiform juvenile cataracts and the R36S mutation that causes congenital juvenile cataracts (Heon et al. 1999). In the human eye, these mutations cause the protein to crystallize and fall out of solution. Studies performed on the R38H and R36S mutants showed crystals formed that retained native-like structure *in vitro* but formed strong intermolecular association tendencies not observed in the wild-type protein (Pande et al. 2001).

Although the aforementioned amino acid substitutions are caused by hereditary DNA mutations, the underlying mechanism of ordered protein aggregation in these processes may be applicable to mature onset cataracts as well. Crystallin proteins removed from old cataractous lenses contain many covalent modifications. It is still unclear whether these modifications destabilize the protein making it more prone to cataract or if covalent damage is accumulated after the precipitates form. It is possible, however, that deamidation, peptide bond cleavage, methylation, carbamylation, aromatic ring opening, and even cysteine or methionine oxidation caused by one of the many oxidative stresses present in the lens may accumulate in the crystallins disrupting protein structure. As demonstrated in the both sickle cell anemia and the juvenile onset cataracts, one exposed hydrophobic or cysteine residue can act as an interaction interface for oligomer formation.

b. Domain Swapping

A candidate mechanism for aggregation of many proteins is domain swapping (Liu and Eisenberg 2002). In these polymerization reactions, native inter-domain interactions in a monomer are replaced with interactions between folded conformational regions of multiple polypeptide chains (Figure 1-4). In most cases the swapped structure had been identified unexpectedly in the crystal form or the protein. The kinetic process has only been observed and studied experimentally in a limited number of cases.

One protein “swaps out” an independently folded unit of its structure for the same folding unit of another protein of identical sequence. Domain swapping may involve isolated pieces of secondary structure such as one β -strand or one α -helix or may involve the swapping of an entire domain of the protein. Because domain swapping involves proteins of the same linear sequence with little or no 3-dimensional remodeling of the folded secondary structure, the intermolecular interactions formed are native-like, often as strong as the monomer interactions, and cannot be visualized using techniques that monitor loss of secondary structure like circular dichroism.

Domain swapping can occur in closed conformation in which a specific number of molecules swap conformational units resulting in a circularized protein multimer with no free unpartnered domains. Alternatively, domain swapping has been hypothesized to occur in infinitum producing an open-ended conformation with unpartnered domains on either end of the growing chain. It is in this open-conformation that polymerization and aggregation of proteins is thought to occur (Liu and Eisenberg 2002, Rousseau et al. 2003).

The first protein to be crystallized in a domain swapped conformation was diphtheria toxin (Bennett et al. 1994). Domain swapping produces either dimers or trimers in RNase A that involve swapping of either the N- or C-terminal portions of the molecule, indicating that one protein may have multiple domain swapped conformations (Liu et al. 2001, Liu et al. 2002).

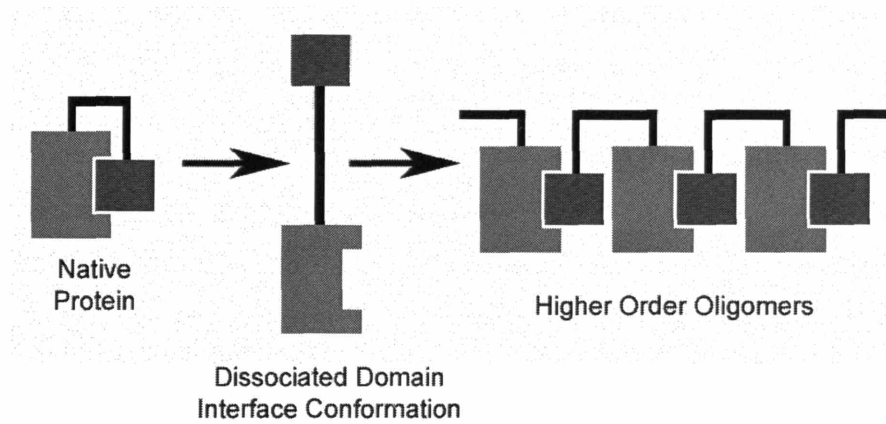


Figure 1-4

Schematic representation of domain swapping. Intramolecular contacts are broken between two individual domains in the protein and are replaced by native-like intermolecular domain contacts. The domain swapping reaction may occur between several protein molecules and may form higher order oligomers.



Figure 1-5

Ribbon diagram of β B2 crystallin in its crystallized domain swapped conformation (Bax et al. 1990).

A functional role for domain swapping was first identified in Glyoxalase I. This protein is enzymatically active in the domain swapped dimeric state, but its activity is reduced when it is found in a less stable monomeric state (SaintJean et al. 1998).

In vitro studies have allowed domain swapping to become appreciated as a viable mechanism for aggregation. The rate at which suc1 aggregated during heating was correlated with the rate at which the protein was able to domain swap (Rousseau et al. 2001). In addition, during refolding of several proteins including suc1, the maltose binding protein, and *E. Coli* tryptophanase, domain swapping is thought to occur as part of either on- or off-pathway conformational states (Silow et al. 1999, Ganesh et al. 2001, London et al. 1974). Finally, and perhaps most strikingly, protein engineering has produced a three-helix bundle protein that was capable of producing not only domain swapped dimers, but higher order oligomeric structures as well (Liu et al. 2001).

Because the $\beta\gamma$ -crystallins are made up of two highly homologous domains, they are a likely candidate for domain swapping. In fact, whereas the γ -crystallins are monomeric in solution, the β -crystallins are dimeric and are composed of two domain swapped chains (Figure 1-5) (Bax et al. 1990). Deamidation and truncation in β B1 crystallin cause elongation of the dimer and modifications of the β -crystallins tend to cause association of β B2 crystallin with the modified protein (Lampi et al. 2001, Zhang et al. 2003). Perhaps a destabilization in the domain interface causes the native intramolecular interface to dissociate and become more likely to domain swap intermolecularly. A domain swapped core of the crystallins could act as a nucleus for further aggregation and, ultimately, cataract formation.

b. Loop-sheet insertion

The serine protease inhibitors or serpins are a class of proteins with distinct structural characteristics that control function. Serpins are composed of a central β -sheet structure with a highly flexible, unstructured, sixteen residue loop. During serpin function, the reactive loop is cleaved in the metastable state which has a melting temperature (T_m) of 60°C and the cleaved strand may become incorporated into the β -sheet forming a highly stable structure with a T_m of greater than 100°C (Carrell and Gooptu 1998).

One well studied serpin is known as α 1-antitrypsin. In the active state, α 1-antitrypsin has a β -sheet made up of five antiparallel β -strands with a surrounding flexible loop. The loop is considered extremely flexible, because its 3-dimensional structure varies based on the pH at which it is crystallized (Ryu et al. 1996, Kim et al. 2001). When the loop is cleaved by trypsin, it inserts itself into the β -sheet forming a six stranded highly stable, but inactive protein.

Normal, functional, loop-sheet insertion occurs intramolecularly, however, it is possible that the cleaved strand of one protein may insert into an adjacent sheet of another protein molecule (Figure 1-6). The polymerization process may continue between several protein molecules and eventually the chain becomes long enough causing the polymer to fall out of solution (Banzon and Kelly 1992, Carrell et al. 1994, Huntingtin et al. 1999).

Intermolecular loop-sheet insertion has been observed with many serpins and the subsequent serpin deficiency has been identified as the cause of many diseases (Figure 1-7). For example, α 1-antichymotrypsin deficiency results in liver damage, α 1-antitrypsin deficiency causes cirrhosis and emphysema, antithrombin deficiency results in thrombosis, and C1-inhibitor deficiency results in angioedema (Bruce et al. 1994, Sifers 1995).

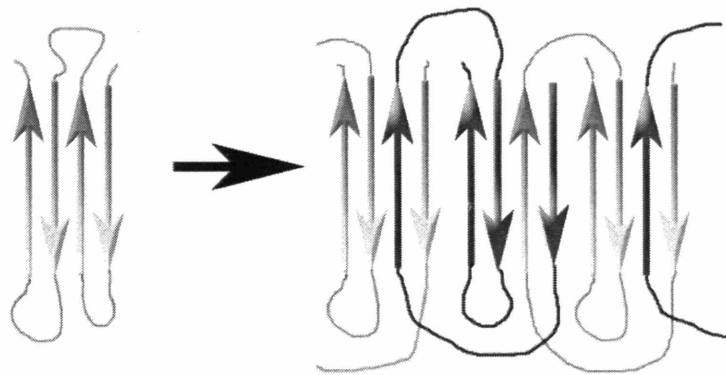


Figure 1-6

Schematic representation of loop-sheet insertion. In this mechanism, a β -strand from one protein is disrupted from its intramolecular native state and becomes incorporated into a nearby protein via native-like interactions.

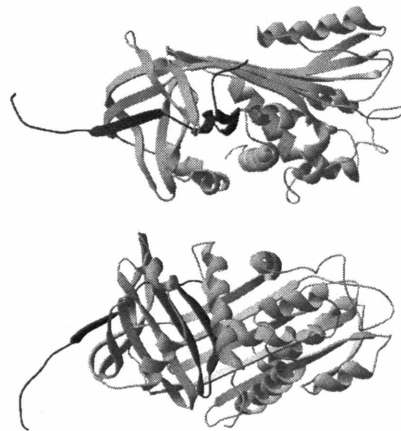


Figure 1-7

Ribbon structure of a cleaved α 1-antitrypsin polymer. The intermolecular loop-sheet insertions of β -strands are shown in black (Huntington et al. 1999).

The structure of the $\beta\gamma$ -crystallins is quite high in β -sheet content. Each domain is composed of two Greek key motifs. In cataracts removed from aged lenses, one major covalent modification is peptide bond cleavage. For human γ D crystallin, a premature truncation at Trp156 has recently been identified as a causative agent of juvenile onset cataracts (Santhiya et al. 2002). One could imagine that peptide backbone cleavage or simply tertiary structure destabilization could produce a free strand with the potential of becoming incorporated into another domain. If this inter-molecular β -sheet is stable, it has the potential to be propagated over several molecules thus producing a large molecular weight species prone to aggregation.

In addition, α -crystallin is known to bind the $\beta\gamma$ -crystallins both *in vitro* and *in vivo*. Recently, Devlin et al. demonstrated that α -crystallin was able to prevent the nucleation-dependent aggregation of α 1-antitrypsin and α 1-antichymotrypsin (2003). Perhaps the loop-sheet insertion self-association reaction of these two serpins presents a binding epitope for α -crystallin that is similar to the binding region formed in the polymerization of the $\beta\gamma$ -crystallins.

c. Amyloidosis

Many human proteins are capable of forming amyloid-like fibrils *in vivo* and causing a variety of diseases many of which are characterized by neurodegeneration (Kelly 1996). Some of the most well understood amyloid forming polypeptides include transthyretin (TTR) which causes senile systemic amyloidosis, β 2-microglobulin which causes dialysis and hereditary renal amyloidosis, lysozyme which causes autosomal hereditary amyloidosis, and the A β -peptide which may cause AD (Saraiva 2001, Inuoe et al. 1997, Pepys et al. 1993). In the amyloidogenic diseases, native protein conformation is altered resulting in the formation of a protein structure capable of self-associating into a primarily β -sheet fiber structure in which the β -strands are perpendicular to the long axis of the fibril (Figure 1-8) (Sipe 1994).

Amyloid proteins bind thioflavin T and show Congo red birefringence. The fiber-like structure of amyloid may be anywhere from 60-100 Å in width and may be branched or unbranched. Oftentimes, the fibrils are twisted and demonstrate regular periodicity visible in TEM and AFM images (Caughey and Lansbury 2003).

Transthyretin is a protein that undergoes an acid-induced conformational change *in vitro*. Under these conditions, an intermediate is formed that is capable of polymerizing into amyloid fibrils (Figure 1-9) (Colon and Kelly 1991, Colon and Kelly 1992). TTR is normally found in the tetrameric state at pH 7.5. When the pH is reduced to 5, the tetramer dissociates into monomers that have both secondary and tertiary structure, but do not have exposed hydrophobic surfaces. The pH-induced monomers are thought to be the intermediates in amyloid formation (Lai et al. 1996). Interestingly, these monomers are not detected in a mutant of TTR, T119M, known to be resistant to amyloid fibril formation (McCutchsen et al. 1995).

Perhaps the most prominent of the amyloid-forming polypeptides is the Aβ-peptide because of its involvement in AD. The amyloid precursor protein (APP), a membrane protein of unknown function, is cleaved in the cell by presenilins and a peptide of 40 or 42 amino acids is released. When released from the membrane, the Aβ-40 or Aβ-42 peptides are converted to amyloidogenic intermediates with strong propensities to form β-sheets. When the concentration of the Aβ-40 or Aβ-42 peptides is high enough, atomic force microscopy has revealed a nucleation-dependent polymerization mechanism involving a protofibril intermediate which then converts to a twisted amyloid fiber (Harper et al. 1997b, Harper et al. 1999). The protofibrils are approximately 40% of the width of the fibers and are usually unbranched. The protofibril to fiber transition may be accelerated by addition of fiber nuclei seeding units (Harper et al. 1997a).

Recent studies have identified other protein structures in the amyloidogenic pathway (Kayed et al. 2003). Many fiber-forming proteins may adopt a conformation consisting of a “ring” of monomers linked together via an unknown association mechanism.

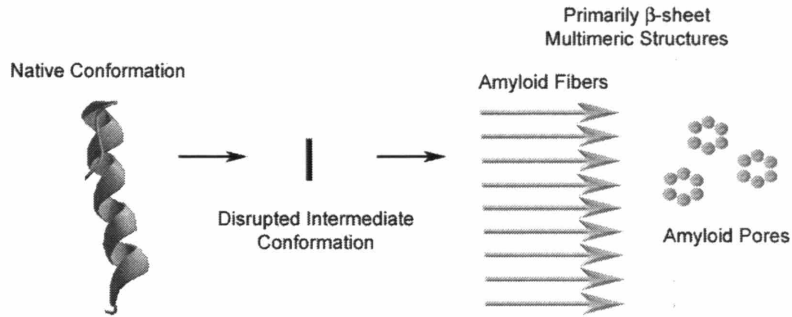


Figure 1-8

Schematic representation of disordered amyloid formation. In amyloidosis, the native protein changes conformation to an intermediate conformation. The intermediate can then polymerize into amyloid fibers or pores.

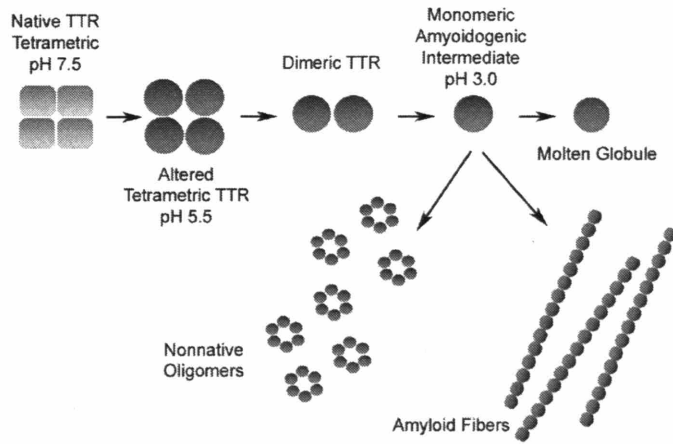


Figure 1-9

Schematic representation of TTR aggregation (Colon and Kelly 1991). As the pH of the solution is lowered an amyloidogenic intermediate forms that is capable of becoming fibers or pores.

The pore-like species appear independently from the amyloid fibers and are thought to be an off-pathway polymerized state. The oligomeric, amyloid pores have structures that are similar to bacterial cytolytic β -barrel pore-forming toxins and thus are believed to increase membrane permeability (Hotze et al. 2002). Historically, the amyloidogenic fiber aggregates were thought to be the toxic species in amyloidogenic disease. However, protofibrils and small oligomeric aggregates of transthyretin (TTR), Cu/Zn superoxide dismutase 1 (SOD1), amyloid- β protein (A β) and α -synuclein all cause cellular toxicity reminiscent of their respective diseases, systemic amyloidosis, familial amyotrophic lateral sclerosis (FALS), Alzheimer's disease (AD), and Parkinson's disease (PD) (Caughey and Lansbury 2003).

Experiments performed on TTR have shown that small oligomeric species cause apoptosis of IMR-32 cells in culture (Reixach et al. 2004). Aggregates of TTR were filtered using size exclusion chromatography and only species of less than 100 KDa induced cell death. The cytotoxic species is thought to be in either a monomeric or a nonnative hexameric state.

The cellular conformation of α -synuclein is natively unfolded and the protein has an undetermined physiological function (Jensen et al. 1998, Lotharius and Brundin 2002). This protein undergoes a polymerization pathway that proceeds from small spherical protofibrils to long fiber filaments as observed in TEM images (Caughey and Lansbury 2002). The rate of formation of the pore-like oligomers is accelerated by mutations in α -synuclein known to be associated with PD. Furthermore the PD oligomer has significant membrane-binding capability that may be related to the disease (Conway et al. 1998). These circular oligomers of α -synuclein have the ability to disrupt synthetic vesicles and are now considered the cytotoxic cellular species (Volles et al. 2001, Volles and Lansbury 2002).

The mechanism of neuronal cell death in AD is thought to be caused by a change in membrane ion permeability by A β protofibrils (Kawahara and Kuroda 2000). Mixtures of wild-type and pathogenic forms of A β -40 form protofibrils and amyloid pores *in vitro* (Lashuel et al. 2003). TEM of the aggregating protein revealed that several different A β polymerized structures formed including small compact circular particles of 4-5 nm in size, pore-like protofibrils, large spherical aggregates of 18-25 nm in diameter,

as well as fibril and fibrous species. The circular oligomers may create amyloid pores that destabilize membrane permeability (Caughey and Lansbury 2003).

SOD1 undergoes a polymerization process that is slightly different to those described for A β and α -synuclein. Mutants of SOD1 implicated in familial amyotrophic ALS aggregate under oxidizing conditions, but the aggregates are not amyloid in either their appearance or ability to bind Congo red or thioflavin T (Rakhit et al. 2002). Post-mortem studies transgenic mice expressing SOD1 variants observed in ALS have identified fragmented Golgi apparatus reminiscent of insufficient energy production by cellular mitochondria (Stieber et al. 2004). Pore-like oligomeric species have recently been identified during *in vitro* studies that may interact with and disrupt the mitochondrial membranes in ALS neuronal cells (Chung et al. 2003).

Dobson and his coworkers have shown that many proteins, including the crystallins, are capable of undergoing a conformational change that produces amyloid *in vitro* (Meehan et al. 2004). This suggests that under the right conditions, the $\beta\gamma$ -crystallins could undergo a structural rearrangement in the lens to form an intermediate capable of polymerizing into an amyloid fiber.

Specifically, H γ D-Crys has been shown to form both amyloid and non-amyloid fibers *in vitro* (Meehan et al. 2004, Kosinski-Collins and King 2003). Preliminary AFM studies have demonstrated that small circular structures may form during aggregation *in vitro* as well (M. S. Kosinski-Collins, D. J. Solis, A. Belcher, and J. A. King, unpublished results) (Appendix E). We do not know whether any of these species may form in the lens or if a change in lens cell permeability is important in cataract. In addition, we have no idea whether or not amyloid fiber or pore formation actually happens in the eye, but an amyloid fiber or oligomeric seed may very well be the core of the cataractous material removed from aged lenses.

d. Prion diseases

The prion diseases are a distinct class of amyloid-forming molecules. Prion diseases are characterized by a conformational change from one form of a protein to another and result in the accumulation of fiber-like structure that can bind Thioflavin T and show Congo red birefringence. In the transmissible spongiform encephalopathies (TSE) such as kuru and Creutzfeldt's Jakob's disease, the cellular prion protein (PrP) changes from an α -helical state to a primarily β -sheet scrapie species (PrP^{Sc}) (Pan et al. 1993, Safar et al. 1993, Baldwin et al. 1994, Lansbury 1995, Nguyen et al. 1995, Zhang et al. 1995). The cellular function for PrP is not currently known. PrP^{Sc} is a protein-only infectious particle that may convert PrP to the scrapie state and thus makes the disease transmissible when PrP^{Sc} protein is transferred from an infected to a healthy individual. There is a distinct species barrier in the infectivity of prion diseases.

Cohen and Prusiner have hypothesized that the amyloid structure of the aggregated PrP^{Sc} is a trimer composed of three right-handed parallel β -helices. Folding studies on the P22 tailspike protein, a polypeptide adopting a parallel β -helix fold, have shown that this structure is prone to self-association and aggregation during folding and that the resulting aggregate is resistant to SDS and temperature denaturation (King et al. 1996). The native structure of the $\beta\gamma$ -crystallins is composed of antiparallel β -sheets. One could imagine a 3-dimensional rearrangement of the $\beta\gamma$ -crystallins that formed a β -helical structure similar to that observed with PrP.

Although it is unlikely that cataract is a self-propagating prion-like disorder, it is important to remember that once an aggregating nucleus forms, it provides a sink for any other proteins that become unfolded or mutated due to environmental stresses. Because of the lack of protein turnover in the eye, at some rate, the naturally present proteins are being modified through oxidation by atmospheric oxygen, heat, and electromagnetic light (Sloney 2002). If a polymerized "seeding" unit is formed in the eye that contains exposed hydrophobics, it may recruit any other nearby proteins as they become disrupted before

the natural degradation and chaperoning machinery of the lens can stop or impede the aggregation process.

One way in which amyloid may form is via a domain swapping mechanism (Liu and Eisenberg 2002). Crystallography and NMR reveal that two proteins capable of amyloid formation, the human prion protein and human cystatin C, may domain swap and form fibers *in vitro* (Knaus et al. 2002, Janowski et al. 2001). It is possible that the domain-swapped species may either directly assemble into amyloid. Alternatively, polymerization of PrP and cystatin C may involve an intermediate common to the domain swapping and amyloid formation pathways.

CHAPTER II: CHARACTERIZING THE UNFOLDING, REFOLDING, AND AGGREGATION OF HUMAN γ D CRYSTALLIN, AN EYE LENS PROTEIN IMPLICATED IN CATARACT FORMATION

A. INTRODUCTION

Human age-onset cataracts affect nearly 50% of the world's population and are the leading cause of blindness worldwide (Clark 1994). Although cataracts are treatable, this treatment is invasive, expensive, and is performed only if the cataract has reached a sufficient level of severity.

Pathological studies of cataractous lenses have revealed that cataracts are composed of protein aggregates that precipitate in lens cells of the eye. The insoluble protein species obstructs the passage of light through the lens, thereby, blocking light from reaching the photoreceptors in the retina (Benedek 1997).

The human eye lens, as a tissue, is composed of layers of fibrous cells that continuously grow with age. Crystallins comprise 90% of the total protein content of the lens (Oyster 1999). The ubiquitous crystallins are expressed primarily early in life and they, therefore, must remain stable throughout a person's lifetime despite the high protein concentration in the lens and the continued presence of oxidative stress from atmospheric oxygen, UV, and visible light to maintain transparency.

In addition to the unique cellular structure of the lens, the overall protein concentration within these cells is extremely high (approaching 70% g/g wet volume). A short-range order exists between the tightly packed crystallin proteins in the lens that provides minimal solution turbidity and a high degree of light transparency (Delaye and Tardieu 1983).

Cataracts removed from the human eye lens are composed of different species of aggregated crystallins. The cataractous crystallin proteins may be divided into two categories: α -crystallins and $\beta\gamma$ -crystallins. α -crystallin is a member of a class of small heat-shock proteins thought to bind to unfolded polypeptide chains during times of stress and is thus crucial to preventing protein aggregation (Horwitz 2000; Clark and Muchowski 2000; MacRae 2000).

$\beta\gamma$ -crystallins are small 20-30 KDa proteins primarily composed of anti-parallel β -sheets. β - and γ -crystallins are structurally similar. They are both comprised of four Greek-key motifs separated into two domains (Wistow et al.). The domains are very similar and appear to be the result of gene duplication during evolution. The β -crystallins form domain-swapped dimers in solution due to their flexible linker sequence, whereas the γ -crystallins are monomeric in solution (Jaenicke 1999). In addition, the γ -crystallins are the only known crystallins having attractive forces between molecules (Tardieu et al. 1992; Clark 1994). Extensive biophysical studies have been performed on members of the $\beta\gamma$ -crystallin family proteins *in vitro* (Pande et al. 1991; Wenk et al. 2000; Norledge et al. 1997; Rudolph et al. 1990; Mayr et al. 1997; Slingsby et al. 1997).

Human γ D crystallin (H γ D-Crys) is a 173 amino acid protein. H γ D-Crys has a high sequence similarity to its bovine homologues, bovine γ D crystallin (~86%) and bovine γ B crystallin (~75%) (Slingsby et al. 1997). The human structure modeled using the bovine γ -crystallin 3-D structures is shown in Figure 2-1 (Guex and Peitsch 1997; Peitsch 1995 and 1996). The structure of the human protein, has been determined from crystals grown from the human recombinant protein and is fully homologous with the bovine structure (Christine Slingsby, personal communication). H γ D-Crys has four intrinsic tryptophans that may be used to probe the unfolding/refolding progression of the molecule with fluorescence spectroscopy.

Human genetic studies of families exhibiting juvenile-onset cataracts identified a set of surface amino acids of H γ D-Crys, including R114C, R38H, and R36S (Heon et al. 1999; Kmoch et al. 2000). Site-specific mutagenesis of recombinant H γ D-Crys in collaboration with the laboratory of George Benedek (MIT Physics Department), confirmed that these substitutions influence the protein's phase transition characteristics *in vitro* (Pande et al. 2000 and 2001). The mutation R114C resulted in the presence of an extra solvent-exposed cysteine that formed a disulfide bond with the endogenous solvent-exposed cysteine, Cys 110 (Pande et al. 2000). The disulfide linkage caused the formation of high molecular weight oligomers that precipitated out-of solution *in vitro*. The mutations R38H and R36S are associated with aculeiform and congenital juvenile-onset cataracts, respectively. These juvenile cataracts have distinctive morphologies that

suggest protein crystallization occurs within the lens. The purified mutant proteins exhibited decreased solubility *in vitro* (Pande et al. 2001). This alteration in the phase separation characteristics was due to an increased crystal nucleation rate of H γ D-Crys, consistent with the pathology of the inherited amino acid substitutions. Based on the importance of H γ D-Crys in juvenile-onset cataracts, the protein is likely to be an important substrate in the formation of age-onset cataracts as well.

Numerous studies using polypeptides, including many conferring disease phenotypes, such as α -synuclein, transthyretin, and the A β -42 peptide have demonstrated that protein aggregation is often an ordered polymerization process that proceeds via a distinct mechanistic pathway involving a series of specific non-native interactions (Betts, et al. 1999; Harper et al. 1997; Li et al. 2001; Lashuel et al. 1998, Speed et al. 1995). The aggregation pathway of the β -sheet P22 tailspike protein is known to proceed from a destabilized folding intermediate (Haase-Pettingell and King 1988; Speed et al. 1995). We were interested in the possibility that the mechanism of aggregation of H γ D-Crys in the aging lens differed from juvenile onset genetic cataracts and proceeded from a partially unfolded conformer of the wild type protein (Mitraki and King 1989; Wetzel 1994).

The unfolding and refolding of bovine γ B crystallin has been carefully studied by Jaenicke and his colleagues (Rudolph et al. 1990; Mayr et al. 1997; Jaenicke 1999). They described a complex unfolding transition suggesting the presence of a partially folded intermediate with one of the domains ordered and the other disordered.

To explore the possible relationship of protein folding to cataract formation, we have performed a series of unfolding/refolding studies on H γ D-Crys at or near physiological pH and temperature (pH 7.0, 37°C). The experiments identify an aggregation-prone state of H γ D-Crys populated in low concentration of guanidine hydrochloride (GdnHCl). We present direct evidence from atomic force microscopy that the aggregation pathway of human γ D crystallin is ordered and the protein forms fibrils *in vitro*.

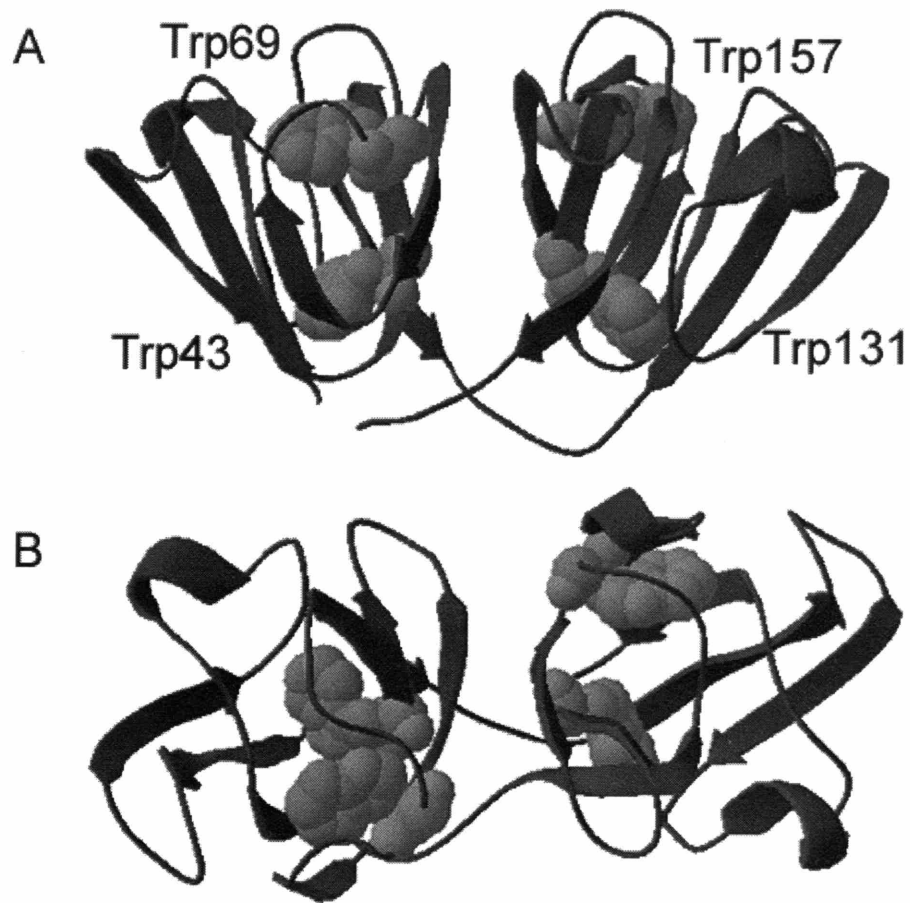


Figure 2-1

Hypothetical ribbon structure of human γ D crystallin showing the location of the four tryptophans at positions 43, 69, 131, and 157. The structure was determined by threading the H γ D-Crys through the known bovine γ D crystallin structure (Guex, et al. 1997, Peitsch, 1995 & 1996). The structure is shown from the side (A) and the top (B).

B. MATERIALS AND METHODS

i. *Expression and purification*

Recombinant human γ D crystallin was prepared from *E. Coli* as described (Pande et al. 2000). Briefly, the protein was purified by fractionating cell lysate on a size exclusion column followed by cation-exchange chromatography as described (Broide et al. 1991).

ii. *Circular dichroism*

CD spectra of native and refolded protein were collected on an Aviv Associates (Lakewood, NJ) model 202 circular dichroism spectrometer. All readings were performed on 0.1 mg/mL H γ D-Crys protein samples.

iii. *Equilibrium refolding and unfolding*

For the unfolding equilibrium titration, purified H γ D-Crys was diluted to 10 μ g/ml in increasing amounts of GdnHCl in S buffer from 0 to 5.5 M. S buffer contained 10 mM NaPO₄, 5 mM DTT, 1 mM EDTA, pH 7.0. The samples were incubated at 37°C until equilibrium was reached (about 6 hours). For the refolding titration, 100 μ g/ml protein was denatured in 5.5 M GdnHCl in S buffer at 37°C for 5 hours. The protein was subsequently refolded by dilution to 10 μ g/ml into decreasing concentrations of GdnHCl from 5.5 to 0.55 M. The fluorescence spectra of the equilibrated samples were determined using a Hitachi F-4500 fluorimeter equipped with a continuous temperature control system with excitation at 295 nm and emission from 310 to 420 nm. The emission intensities at 350 nm were used for data analysis. The excitation and emission slits were both set to 10 nm.

iv. *Unfolding fluorescence kinetics*

Tryptophan environment changes with refolding were monitored using a Hitachi F-4500 fluorimeter equipped with a continuous temperature control system. Native protein (100 $\mu\text{g/ml}$ in S buffer at 37°C) was unfolded by dilution into S buffer to final concentrations of 10 $\mu\text{g/ml}$ HyD-Crys and 5.5 M GdnHCl. Loss of global structure was monitored with continuous excitation at 295 nm and emission at 350 nm at 37°C.

v. *Refolding fluorescence kinetics*

Tryptophan environment changes with refolding were monitored using a Hitachi F-4500 fluorimeter equipped with a continuous temperature control system. Native protein was denatured at 100 $\mu\text{g/ml}$ concentration in 5.5 M GdnHCl in S buffer at 37°C for 5 hours. The unfolded protein was refolded by dilution into S buffer to final concentrations of 10 $\mu\text{g/ml}$ HyD-Crys and 0.55 M GdnHCl or 1.5 M GdnHCl. Increase in global structure while refolding was monitored with continuous excitation at 295 nm and emission at 350 nm. Fluorescence wavelength spectra were obtained after 5 hours of refolding both before and after a 20-minute spin at 12,000 rpm. A background fluorescence correction was made by obtaining spectra of S buffer containing 0.55 M GdnHCl and subtracting it from the refolded sample data.

vi. *Refolding solution turbidity kinetics*

Solution turbidity changes caused by formation of high molecular weight species were monitored with refolding in a Cary 50 Bio UV/Vis spectrophotometer. Samples were prepared as discussed for the refolding fluorescence kinetic experiments. Absorbance spectra from 260 to 350 nm were monitored at various time intervals. Absorbance values at 280 nm were used for data analysis. Samples were stirred throughout the refolding process.

vii. *Atomic force microscopy*

AFM analysis was performed using the tapping method as described (Marini et al. 2002). Ten μl of sample was allowed to nonspecifically bind to a mica surface for a total drying time of 75 seconds. The mica was then washed with 150 μl of milli-Q water and allowed to air-dry before imaging.

viii. *BisANS binding assay*

The character of the H γ D-Crys aggregate was probed by monitoring bisANS fluorescence changes upon binding. H γ D-Crys aggregates were prepared by refolding denatured protein in 0.55 M GdnHCl in S buffer at 10 $\mu\text{g/ml}$ H γ D-Crys at 37°C. Samples were removed at refolding time intervals ranging from 30 seconds to 4 hour. BisANS was added to the resulting sample to reach a final small molecule concentration of 20 nM. The fluorescence spectrum of the sample was determined by excitation at 350 nm and emission from 400 to 600 nm on a Hitachi 4500 fluorimeter. Background fluorescence of bisANS in 0.55 M GdnHCl was subtracted from the resulting spectrum for data analysis purposes.

C. RESULTS

i. *Purification*

Recombinant human γD crystallin was expressed from a pET.16 plasmid in *E. Coli* strain *BL21(DE3)*. The recombinant H γ D-Crys expressed in *E. coli* folded into soluble protein subunits that were stable in the bacterial cytoplasm. Methods developed for the purification of γD crystallin from the bovine lens were used to purify the crystallin subunits from the bacterial lysate (Pande et al. 2000). Recombinant H γ D-Crys contained an N-terminal methionine that is post-translationally cleaved in the native protein. The crystallin purified from *E. Coli* had far-UV CD (Figure 2-2) and absorption spectra (data

not shown) similar to those observed for soluble bovine γ D crystallin extracted from calf lenses (Hay et al. 1994). X-ray diffraction studies of crystals grown from the recombinant protein indicated the same fold as that found for the bovine protein (Basak et al. 2003).

ii. *Equilibrium unfolding and refolding*

Human γ D crystallin did not denature in urea at concentrations in excess of 8 M, but did exhibit a denaturation transition in GdnHCl. Unfolding and refolding transitions were examined using GdnHCl as the denaturant in equilibrium analyses. Buffers were maintained at pH 7 and temperatures of 25°C or 37°C. Under these conditions a cooperative unfolding transition occurred in the range of 2.0 to 3.3 M GdnHCl.

With respect to the overall tryptophan fluorescence, the native-state of the protein was quenched to a greater extent than the denatured state (Figure 2-3A). Based on the hypothesized structure of H γ D-Crys, that native-state fluorescence quenching may occur because each of the four tryptophan side-chains present in the protein are within close proximity to polar residues. The maximum fluorescence intensity of native H γ D-Crys occurred at 325 nm while the maximum fluorescence intensity of denatured protein occurred at 348 nm.

At 37°C, the unfolding equilibrium transition of H γ D-Crys at 350 nm showed that protein unfolding was a two-state process (Figure 2-3B). Control titrations with sodium chloride demonstrated that fluorescence changes observed in high and low denaturant concentrations reflected structural alterations in the protein as opposed to high salt effects from GdnHCl (data not shown). UV/Vis solution turbidity experiments did not reveal the presence of any high molecular weight aggregates.

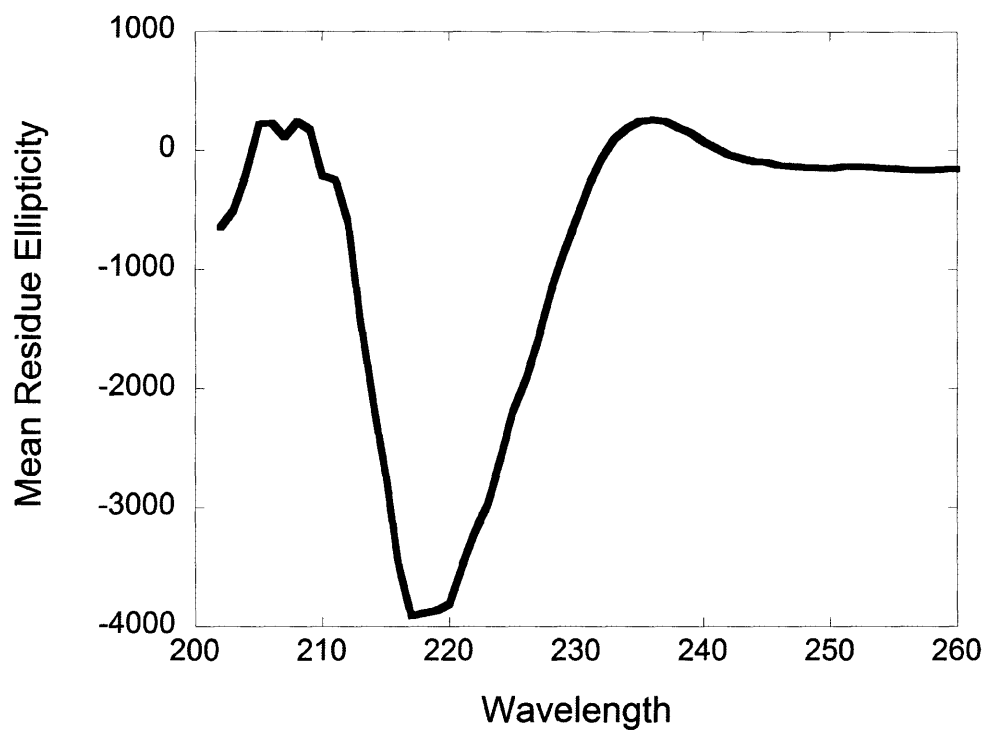


Figure 2-2

Far-UV CD spectrum of 0.1 mg/mL recombinant HyD-Crys in S Buffer purified from *E. Coli* lysates.

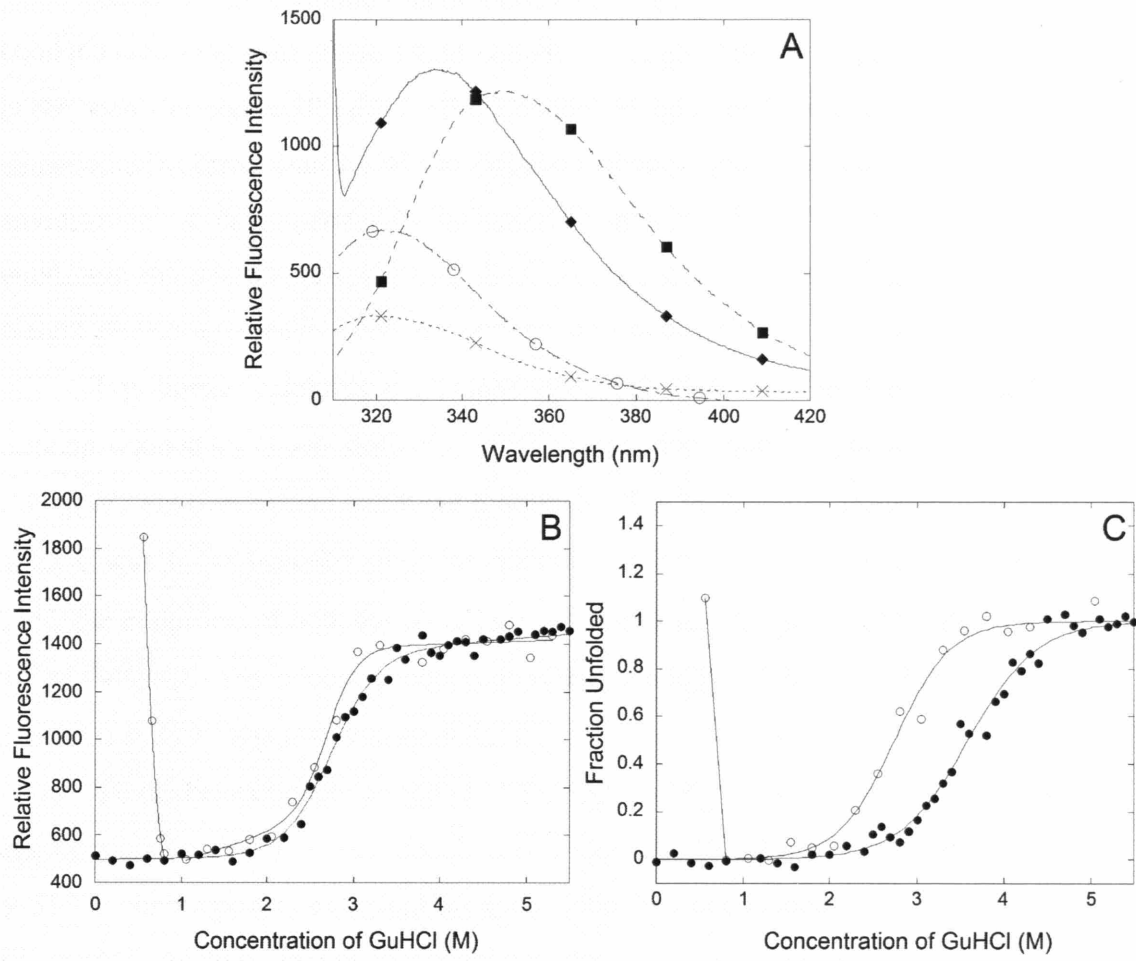


Figure 2-3

Equilibrium unfolding and refolding of HyD-Crys in GdnHCl. Intrinsic tryptophan fluorescence was monitored of HyD-Crys during folding and refolding and all samples were equilibrated in S buffer at a protein concentration of 10 $\mu\text{g/ml}$. (A) Fluorescence wavelength spectra of native protein (○), denatured protein in 5.5 M GdnHCl (■), HyD-Crys refolded at 0.55 M GdnHCl before a 12,000 rev/min centrifugation (◆), and after a 20 minute centrifugation (×) at 37°C. (B) Relative fluorescence intensity of unfolding (●) and refolding (○) at 350 nm at 37°C. (C) Fraction unfolded intensity of unfolding (●) and refolding (○) at 350 nm at 25°C. Fraction unfolded values were calculated from raw fluorescence intensity measurements using the method described by Pace et al. (1989). Samples were allowed to equilibrate for 6 hours in the appropriate conditions.

At 37°C, the refolding transition of HyD-Crys appeared to be reversible in GdnHCl concentrations above 1.0 M GdnHCl, though with the suggestion of a hysteresis in the transition region (Figure 2-3B). Dilution of denatured protein to GdnHCl concentrations lower than 1.0 M gave rise to a protein species with non-native tryptophan environments as demonstrated by increased fluorescence intensity and wavelength maximum shifts to 330 nm (Figure 2-3A). These samples contained high molecular weight protein aggregates as evidenced by solution turbidity measurements.

Equilibrium refolding and unfolding of HyD-Crys as a function of GdnHCl concentration at 25°C was similar to 37°C, but the refolding and unfolding transition curves did not overlay and exhibited hysteresis. The midpoint of the unfolding transition at 25°C was 3.7 M GdnHCl while the midpoint of the refolding transition was 2.7 M. At 37°C the midpoints of both the unfolding and refolding transitions were approximately 2.7 M GdnHCl. The refolding aggregation reaction at low GdnHCl concentration observed at 37°C was also visible at 25°C.

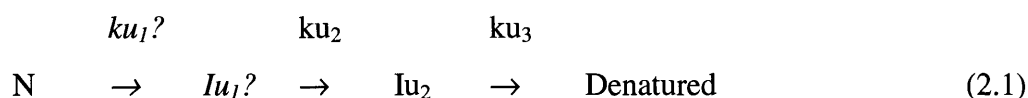
To further examine the nature of the aggregation reaction, we characterized the aggregation reaction as a function of sample incubation time after initiation of refolding at 37°C. Native protein exhibited no appreciable increase in solution turbidity after 40 hours of incubation under these conditions. The location of the aggregation transition was not altered with increasing times of incubation after dilution from 3 to 40 hours. This suggested that the reaction depended on the population of a distinct conformer during refolding in 1 M GdnHCl and did not require a rate-limiting nucleation event (Figure 2-4). Samples containing a final GdnHCl concentration greater than 1.0 M did not exhibit appreciable solution turbidity even after 40 hours. We did not observe aggregation in the major conformational transition occurring from 2 to 3.3 M GdnHCl indicating that intermediates populated in this transition were not the precursors of the aggregated state.

The HyD-Crys protein refolded in aggregation-prone conditions (0.55 M GdnHCl) was centrifuged at 12,000 rev/min. Solution turbidity measurements measured at 280 nm and 350 nm indicated no high molecular weight aggregates remained in the supernatant after centrifugation (data not shown). The fluorescence spectrum of the resulting supernatant revealed that a significant portion of the protein was lost to

precipitation, but the fraction remaining had a maximum intensity wavelength indistinguishable from native (Figure 2-3A). This indicated that there was a native-like species present at the end of the reaction and that the off-pathway aggregation process competed with a productive refolding pathway under aggregation-prone conditions. Integration of the fluorescence wavelength spectra of the resulting supernatant in the experiment shown in figure 2-3 revealed that 50% +/- 2% of the protein was refolded productively and 50% +/- 2% of the protein became incorporated into the aggregate.

iii. *Kinetic unfolding through a partially unfolded intermediate*

The unfolding of HyD-Crys was monitored over time with fluorescence at 37°C (Figure 2-5, Table 2-1). A possible early intermediate (Iu_1) may have been populated within the dead-time of the experiment as indicated by the burst fluorescence intensity at the on-set of the experiment. The only observable intermediate (Iu_2) formed with a $t_{1/2}$ of 55 s. This intermediate was not as quenched as native indicating the polar-tryptophan interaction had been disrupted. In the final unfolding transition ($Iu_2 \rightarrow$ denatured), occurring with a $t_{1/2}$ of 2700 s, a slow re-quenching took place indicative of solvent rearrangement or isolated local restructuring in the unfolded state. The unfolding process was completed in ~2 hours.



Solution turbidity measurements showed no evidence of aggregation during this process. HyD-Crys showed the same kinetic rates and intermediates during unfolding regardless of pH, temperature, starting concentration of protein, and presence or absence of DTT (data not shown).

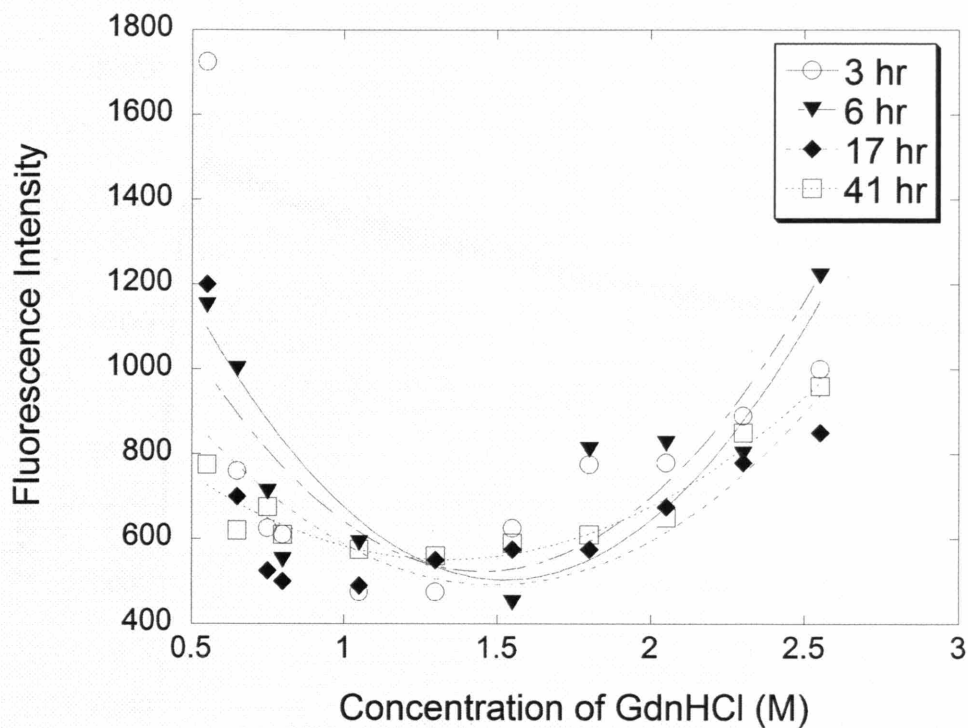


Figure 2-4

Behavior of refolding HyD-Crys species as a function of incubation of the reactions. The intrinsic tryptophan fluorescence of the aggregation-prone region of the refolding equilibrium was monitored with excitation at 295 nm and emission at 350 nm at 37°C in S buffer for 3 hours (○), 6 hours (▼), 17 hours (◆), and 41 hours (□).

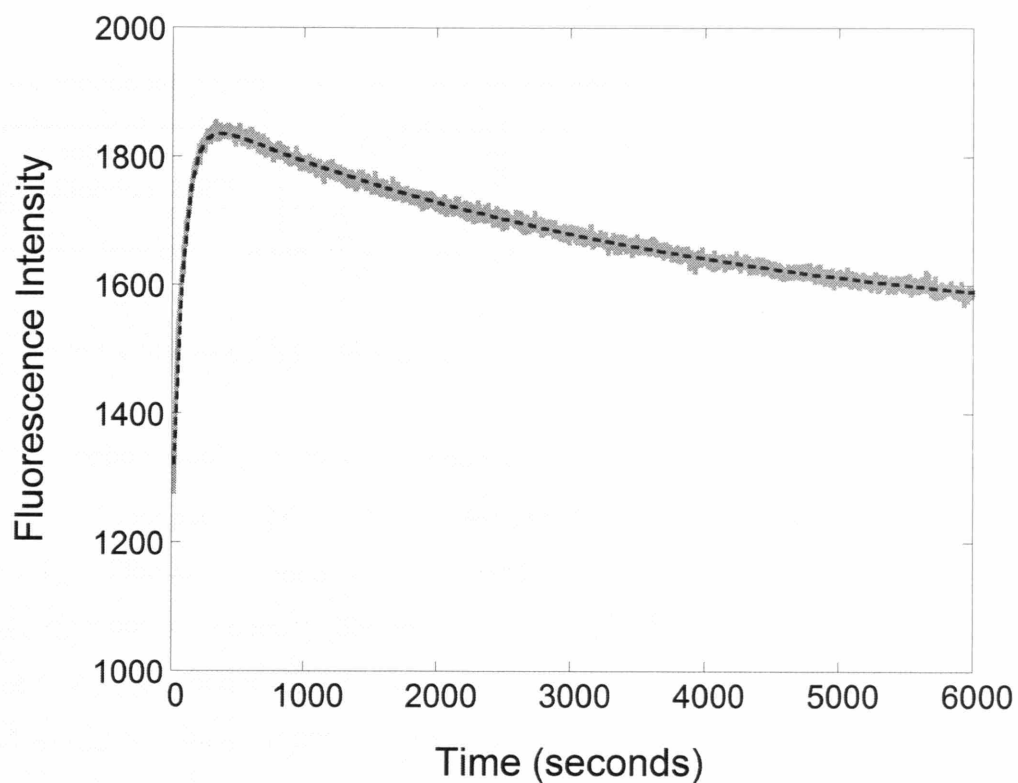


Figure 2-5

Unfolding kinetics of HyD-Crys monitoring intrinsic tryptophan fluorescence with excitation at 295 nm and emission at 350 nm. HyD-Crys was denatured by rapid dilution into 5.5 M GdnHCl at 37°C in S buffer to a final protein concentration of 10 µg/ml. The unfolding protein time course is shown in gray with the resulting three-state curve fit (dashed line).

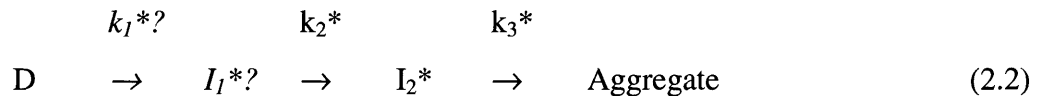
Table 2-1

Kinetic rate constants for unfolding, productive refolding and aggregation of Human γ D Crystallin

	k_1	$t_{1/2}$	k_2	$t_{1/2}$	k_3	$t_{1/2}$
Unfolding	<0.46	<15 s	0.013	55 s	0.00027	2700 s
Refolding/Aggregation	<0.46	<15 s	0.012	60 s	0.00028	2500 s
Productive Refolding	<0.46	<15 s	0.0095	73 s	0.00067	1030 s

iv. *Competing productive and aggregation-prone refolding pathways*

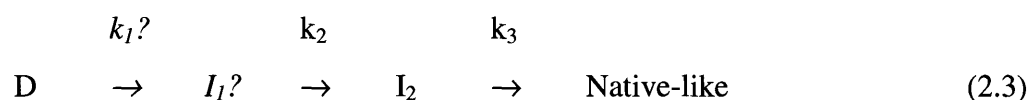
Solution turbidity was monitored at varying time intervals during refolding of HyD-Crys under aggregation prone conditions (Figure 2-6, Table 2-1). The solution turbidity of the first observable species was significantly higher than the native control indicating that a high molecular weight species (I_1^*) had already formed within the dead-time of the experiment. The formation of the first visible high molecular weight species (I_2^*) occurred with an approximate $t_{1/2}= 60$ s while the second state change (I_2^* to aggregate) had an approximate $t_{1/2}=3000$ s. This second transition may have reflected a slow structural transition of the multimeric protein from a highly turbid species to a less turbid species, but also may have corresponded to the progressive precipitation of the aggregate from solution.



This aggregation mechanism was independent of the presence of reducing agents such as DTT and intermediates were confirmed by fluorescence scans (data not shown).

In order to study the kinetic rates and intermediates formed during productive refolding of HyD-Crys without the impact of light scattering, fluorescence experiments were performed on HyD-Crys refolding to a final denaturant concentration of 1.5 M GdnHCl (Figure 2-7, Table 2-1). Under these conditions, fluorescence equilibrium data

revealed the refolding process favored a native-like conformation and no high molecular weight species were detected in solution turbidity measurements. Productive refolding had a possible early intermediate (I_1) that may have formed within the dead-time of the experiment as evidenced by the burst in fluorescence observed at 15 seconds. Spectroscopically, the only visible intermediate (I_2) formed with a $t_{1/2}$ of 73 s and then was converted to native protein with a $t_{1/2}$ of 1030 s. Under these conditions, productively refolded HyD-Crys reached equilibrium in approximately 3 hours whereas refolding HyD-Crys under aggregation-prone conditions did not reach equilibrium until 4 hours had passed.



The final state of the protein under these non-aggregating refolding conditions was native-like in terms of its fluorescence character.

v. The self-associated aggregate has a fibril structure

Atomic force microscopy was used to probe the structure and aggregation mechanism of HyD-Crys. Native crystallin from an S buffer solution was visible upon non-selective binding to the mica surface. The molecules had an apparent height of 0.5 nm, a length of 10 nm, and a width of 10 nm (Figure 2-8A). The discrepancy between this value and the predicted dimensions of the protein (50 Å by 30 Å by 30 Å) were presumably a result of either molecular flattening upon binding, resolution limits of the technique, or dimerization of the protein under the sample preparation conditions. The molecules appeared to contain a region of low height in the very center.

A solution of HyD-Crys was denatured in 5.5 M GdnHCl and subsequently refolded using the aggregation-prone conditions (dilution to 0.55 M GdnHCl). Ten μ L samples were removed and applied to a mica grid at various intervals.

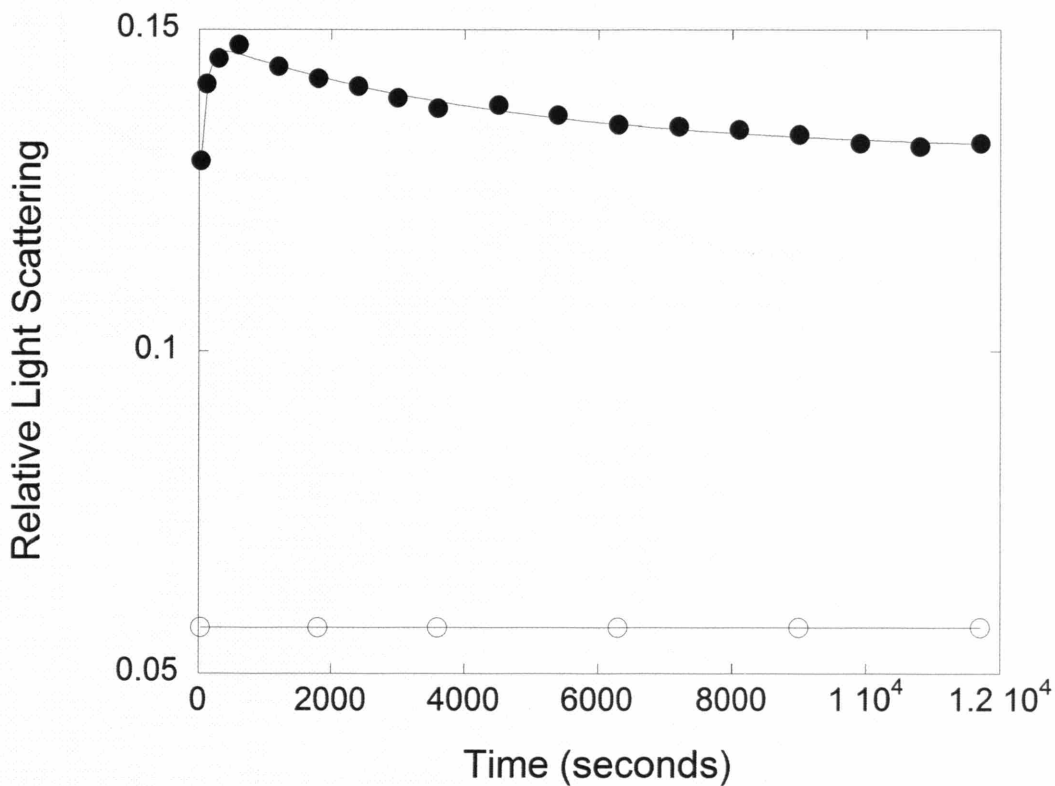


Figure 2-6

Solution turbidity measurements of refolding HyD-Crys over time. HyD-Crys was denatured in 5.5 M GdnHCl at 37°C in S buffer for 5 hours. HyD-Crys was refolded by rapid dilution with S buffer to a final GdnHCl concentration of 0.55 M and a final protein concentration of 10 µg/ml. Relative solution turbidity was determined by measuring refolding HyD-Crys absorbance at 280 nm at various times (●). A background buffer scattering reference is shown for comparison (○).

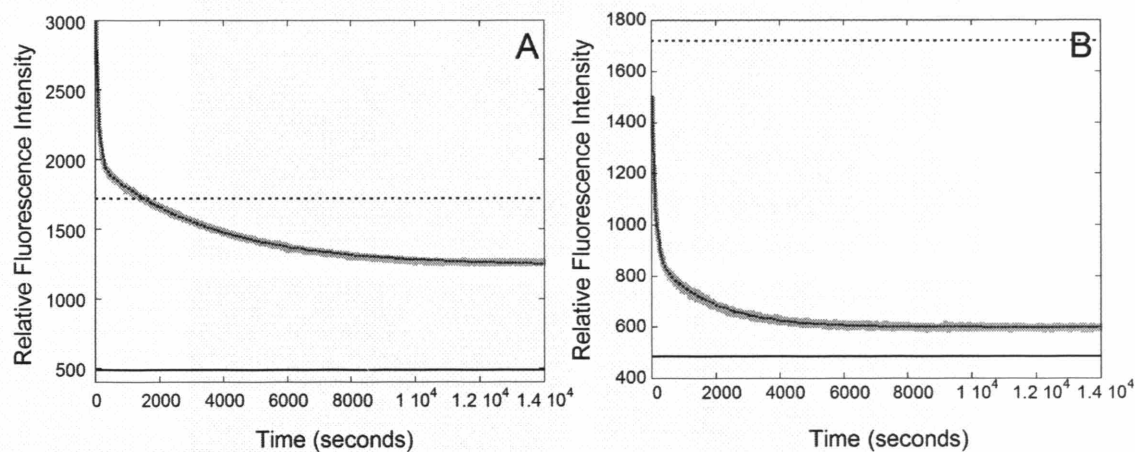


Figure 2-7

Refolding kinetics of HyD-Crys monitoring intrinsic tryptophan fluorescence with excitation at 295 nm and emission at 350 nm. HyD-Crys was denatured in 5.5 M GdnHCl at 37°C in S buffer for 5 hours. HyD-Crys was refolded by rapid dilution with S buffer to a final GdnHCl concentration of 0.55 M (A) and 1.5 M (B) (and a final protein concentration of 10 µg/ml). The refolding protein time course is shown in gray with the resulting three-state curve fit. Reference native (solid line) and denatured (dotted line) intensities are shown for comparison.

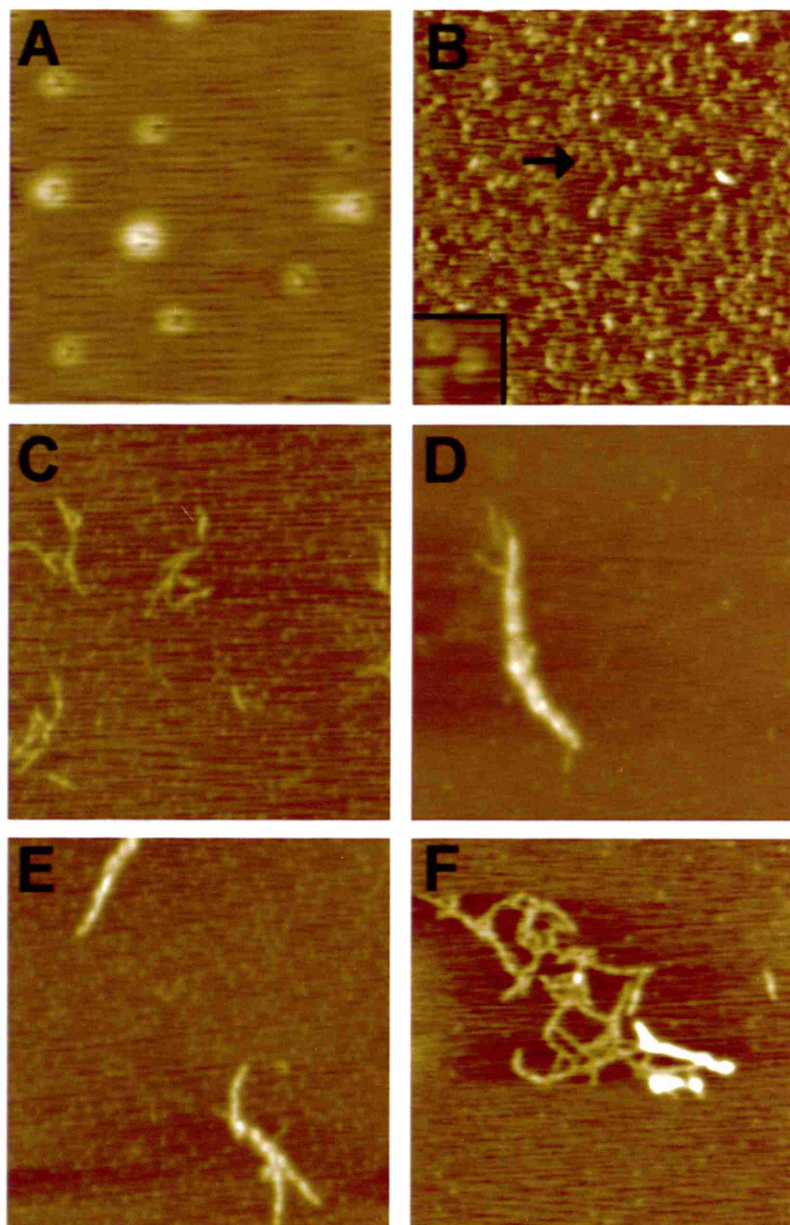
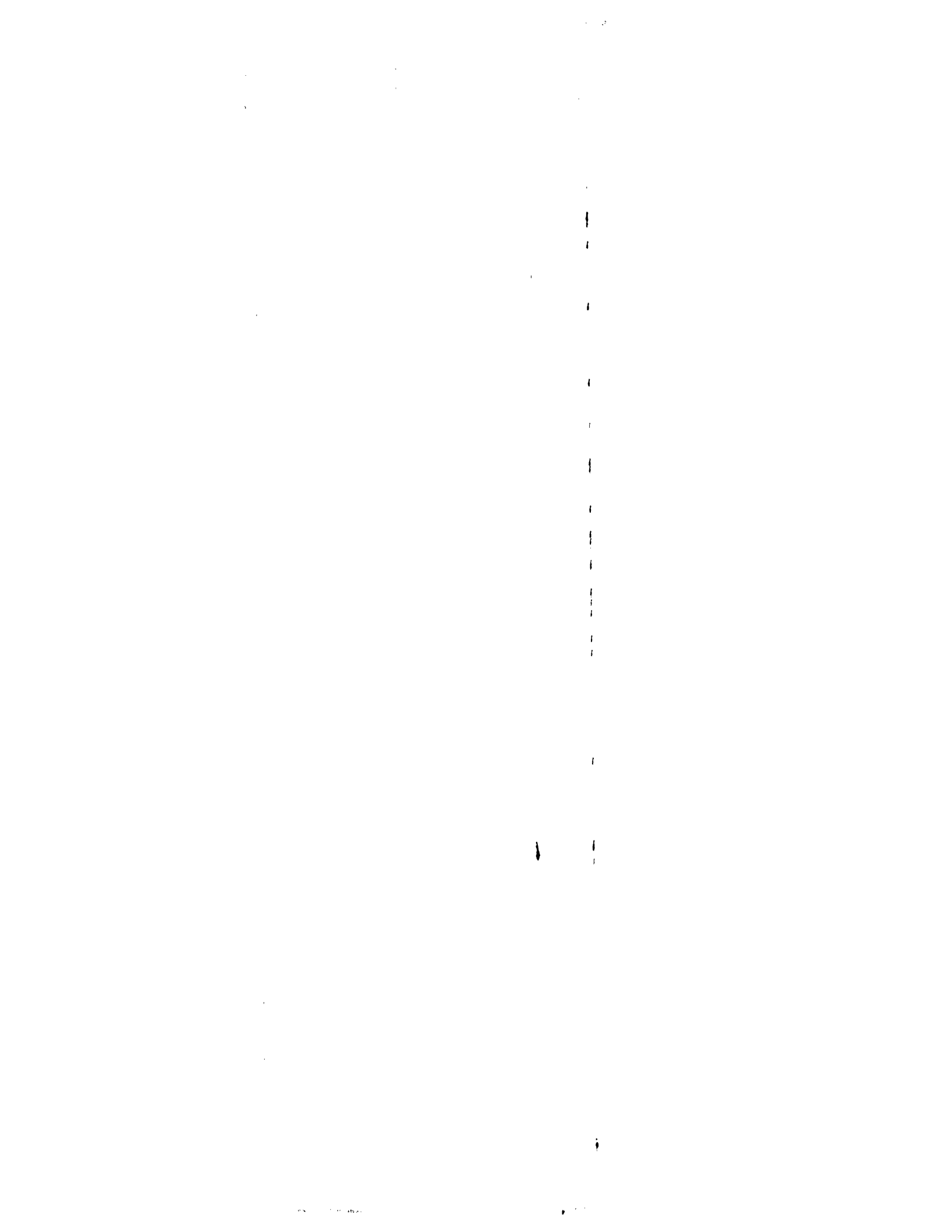


Figure 2-8

AFM height images of refolding HyD-Crys species absorbed to a mica surface. High surfaces are denoted as white (5 nm) and low surfaces are shown in black (0 nm). HyD-Crys was refolded for various times, extracted from solution, absorbed on mica, and observed using tapping mode AFM. Native HyD-Crys (A), and protein refolded for 1 min. (B), 5 min. (C), 24 min. (D), 1 hr. (E), and 2 hr. (F) are shown. Panel A is a 300 nm x 300 nm scanning area while B-F are 1 μm x 1 μm . The inset in B is a close-up of refolded HyD-Crys.



Within the first minute of refolding, small globular structures were visible (Figure 2-8B). In the background of this image, native-like H γ D-Crys monomers are already visible, supporting the competing aggregation and productive pathway model for refolding under these conditions. After approximately 5 minutes of refolding, long, thin, 1 nm high, 15 nm wide, protofibril structures were observed (Figure 2-8C). The protofibrils had an approximate width and height of the soluble crystallin. After about 24 minutes of refolding, protofibrils were visible but were found only in associated masses (Figure 2-8D). The protofibrils had begun to associate from the center producing a species with one thick mid-region and multiple fibril-width tail ends. By one hour of refolding, virtually all of the protofibrils had disappeared, presumably being incorporated into thick 5 nm high, 50 nm wide branched and unbranched fiber bundles (Figure 2-8E). Although the exact height and width of the fiber bundles varied, their overall appearance and approximate size was similar. After several hours of refolding, aggregated masses were visible presumably containing associated fiber bundles (Figure 2-8F).

vi. *The aggregate contains exposed hydrophobic pockets*

H γ D-Crys aggregates were characterized by binding assays with bisANS. All aggregated species, including nuclei presumably present as early as 30 seconds after the start of refolding and fiber bundles present at 4 hours, bound bisANS significantly better than either the unfolded or folded control (Figure 2-9A and 2-9B). This indicated that all H γ D-Crys aggregates under these conditions had non-native patches of exposed hydrophobic residues.

D. DISCUSSION

i. *Crystallin refolding and aggregation in vitro*

Within the eye, the lens crystallins exhibit very long lifetimes at very high concentrations in the presence of ultraviolet and visible radiation. Recombinant human

γ D crystallin was resistant to denaturation by concentrations of urea up to 8 M at neutral pH, but could be denatured upon incubation with high concentrations of GdnHCl. The midpoint of the transition occurred at about 2.7 M GdnHCl, at 37°C.

Within folded H γ D-Crys the fluorescence intensity is representative of quenched tryptophans. Wistow et al. suggested that fluorescence quenching of the native state observed in bovine γ -crystallin proteins is caused by tryptophan-cysteine interactions. The native-state quenching of H γ D-Crys may be due to the interaction of cysteine thiols 19 and 79 with tryptophan 43, cysteine 33 with tryptophans 69, and/or histidine 88 with tryptophan 131. This phenomenon has been observed in other proteins such as cellular retinoic acid binding protein and human serum transferrin N-lobe (Eyles and Gierasch 1999; He et al. 2001). The packing of cysteines against tryptophans may provide a mechanism of protection in the eye from free radical damage initiated by ultraviolet radiation absorption by tryptophans (Davies and Truscott 2001). However, the source of the quenching has not been experimentally determined and may be associated with other residues in the hydrophobic core.

The denatured chains of H γ D-Crys could be refolded by dilution from GdnHCl and the reaction was reversible with one distinct transition in the range of 1-5 M GdnHCl. Bovine γ B crystallin, which is denatured by urea at pH 2.0, but not at pH 7.0, was reversible over all reported concentrations of GdnHCl, but exhibited a three-stage transition in equilibrium unfolding studies, representing sequential denaturation of the C-terminal and N-terminal domains at pH 2.0 (Rudolph et al. 1990; Jaenicke 1999). On the other hand, the closely homologous protein S, a protein also containing two Greek key domains, unfolded and refolded without evidence of separate domain transitions (Wenk et al. 1998). H γ D-Crys may possess differential domain stability that is not visible in the apparent two-state unfolding transition observed during equilibrium unfolding. Future experiments will attempt to characterize the stability of each domain of the human protein.

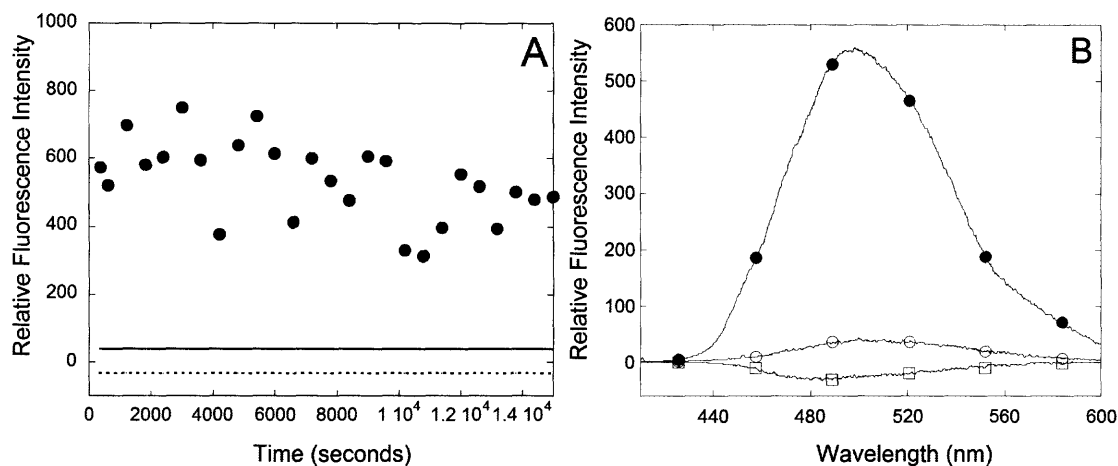


Figure 2-9

Binding of bisANS to refolding H γ D-Crys using bisANS fluorescence with excitation at 360 nm. H γ D-Crys was denatured in 5.5 M GdnHCl at 37°C in S buffer for 5 hours. H γ D-Crys was refolded by rapid dilution with S buffer to a final GdnHCl concentration of 0.55 M and a final protein concentration of 10 μ g/ml. Samples of refolded protein were analyzed at various times, bisANS was added, and small molecule fluorescence was determined at 500 nm (A). Refolded sample (\bullet), denature sample (dashed line) and reference native (solid line) bis-ANS fluorescence are shown. A representative fluorescence scan (B) is shown of aggregated (\bullet), native (\circ), and denatured H γ D-Crys crystallin (\square).

The *in vitro* unfolding and refolding steps of human γ D crystallin were relatively slow compared to some small proteins such as RNase or barstar (Hollien and Marqusee 2002; Nolting et al. 1995), but similar to rates found for some other β -sheet proteins like apo-pseudoazurin (Reader et al. 2001).

During refolding, dilution into concentrations of GdnHCl below 1.0 M at 37°C, resulted in the population of an intermediate that aggregated irreversibly. The aggregating species was populated at low GdnHCl concentrations both at 37°C and at 25°C. At both temperatures, the major conformational transition measured by fluorescence spectroscopy was in the range of 2 to 3.3 M GdnHCl. This was well separated from the GdnHCl concentration in which the aggregation reaction was detected. This suggests that the aggregation-prone intermediate differs from the species populated in the transition or that perhaps similar intermediates are present under both conditions but have increased solubility in higher GdnHCl concentrations.

This behavior differs somewhat from other proteins in which aggregation competes with productive refolding, such as phosphoglycerate kinase, β -galactosidase, interleukin 1- β , and transthyretin (Ghelis and Yon 1982; Wetzel and Chrnyk 1994; Colon and Kelly 1992). For these proteins, dilution to intermediate concentrations of denaturant generated an aggregating intermediate, while dilution to lower concentrations led to the recovery of the native fold. Equilibrium analysis with GdnHCl demonstrated that in HyD-Crys, the aggregating species probably was not generated from a transition intermediate.

Though not well defined at 37°C, a hysteresis between the unfolding and refolding equilibrium curves was very clear at 25°C. The refolding curves had the same midpoint of transition at both temperatures; however, the transition midpoint of unfolding at 37°C occurred at lower GdnHCl concentrations suggesting a thermal destabilization of the native state.

The existence of hysteresis in protein refolding experiments is associated with reactions that are kinetically controlled, and which exhibit high energy barriers between conformational transitions needed for refolding. The folding of a number of β -sheet proteins are kinetically controlled, for example the parallel β -helical tailspike adhesin

(Sturtevant et al., 1989; Chen and King, 1994; Steinbacher et al., 1994). Alkaline phosphatase exhibits a continuous reorganization of the environment around tryptophan 109 to the native state that takes days under refolding conditions (Subramiam et al., 1995). The transformation from the α - to the β -isoform that occurs during refolding of the recombinant mouse prion protein exhibits a high barrier between the native and the denatured state as well (Baskakov et al., 2001). For HyD-Crys, the hysteresis was reduced at higher temperature, consistent with overcoming a kinetic barrier. Correct packing of the hydrophobic cores or formation of proper domain interface alignment are candidates for the slow step. The molecular basis of the hysteresis requires further investigation (Sinclair, et al. 1994; Lai et al. 1997).

ii. *Formation of filamentous aggregates*

AFM images reveal that the *in vitro* aggregate of HyD-Crys was ordered. The aggregate had a filamentous appearance as would be expected from the polymerization of defined subunits. At early times during aggregation, globular species presumably representing soluble crystallin species were present. These were approximately 2-5 times the size of the native crystallin, estimated from the AFM images. These species are candidates for aggregation nuclei. The exposed surfaces at the growing tip may serve as intermolecular interfaces. The multimeric nuclei appeared to polymerize into elongated protofibrils. The thin protofibrils appeared to wind around each other forming fiber bundles. All protofibrils were incorporated into fiber bundles by 4 hours of refolding. Additional exposed hydrophobic surfaces may be the site of association or intermolecular contacts may be made by an unknown mechanism.

The aggregation pathway competed with a productive refolding pathway under the described conditions, suggesting that a partially folded intermediate might be the polymerizing species. Neither equilibrium nor kinetic analyses of HyD-Crys unfolding revealed any evidence for aggregation from the native state of the protein. The native-like species forming during aggregation-prone refolding had a fluorescence spectra and AFM structure similar to native HyD-Crys. Due to the low concentration of protein

present during refolding experiments no other structural evidence was obtainable to confirm this.

Figure 2-10 shows a model for HyD-Crys refolding and aggregation *in vitro*. The initial step on the productive refolding pathway is envisaged to be a global hydrophobic collapse (reviewed in Ptitsyn 1995; Kuwajima 1992) during which large bulky, hydrophobic residues are sequestered away from the polar solvent as evidenced by an increase in the overall fluorescence of the protein. A kinetic partitioning of HyD-Crys chains occurs between the productive refolding pathway and the aggregated complexes. The early collapsed intermediate may be the same for the two pathways. In both aggregation-prone refolding experiments and control productive refolding experiments, a burst phase was visible at 15 seconds that may correspond to the formation of a similar early folding intermediate that is prone to polymerization.

The aggregation/productive refolding breakpoint in the refolding pathway must occur relatively early with respect to the other transition times observed with HyD-Crys as evidenced by the presence of a species that has a high solution turbidity within the dead-time (15 seconds) of the spectroscopic kinetic experiments.

The presence of 1.5 M GdnHCl appeared to shift the equilibrium of the branch point between productive and aggregation-prone folding to favor the productive process as evidenced by equilibrium analysis as well as kinetic experiments in which HyD-Crys was refolded by dilution into 1.5 M GdnHCl.

The HyD-Crys aggregate bound bisANS significantly even as early as the formation of nuclei species, further emphasizing that these aberrant exposed hydrophobic residues exist from the on-set of refolding. The hydrophobic bisANS binding surfaces may correspond to the intermolecular interaction domains that facilitate protofibril and fiber bundle formation. Further work is needed to elucidate which area of the molecule is responsible for these aggregation-facilitating interactions and which area, if any, retains a native-like fold.

The aggregation and refolding pathways were dependent on the concentration of HyD-Crys whereas the unfolding pathway was concentration independent. When the overall concentration of protein was increased by a factor of ten, the presence of 1.5 M GdnHCl was not sufficient to quench the aggregation pathway. This high concentration

aggregation reaction may or may not be the same as the aggregation phenomenon previously described, but these results do suggest a significant divergence of the *in vivo* and *in vitro* pathways. The amount of crystallin found in the eye lens is 70% (g/g wet weight) or approximately 300-500 mg/mL whereas the amount of crystallin used for these experiments was as low as 10 $\mu\text{g/ml}$ (Jaenicke 1999).

iii. *Cataract formation in the lens*

Many human diseases such as glaucoma and Meretoja's syndrome result in the expression of insoluble protein fibers in the eye (Nelson et al. 1999; Kivela et al. 1994). Experiments performed with α -crystallin have demonstrated that a fibrous aggregate forms when this protein is complexed with βL crystallin (Weinreb et al. 2000), and amyloid-like protein species have been identified *in situ* in the mammalian ocular lens using Congo red and thioflavin T binding assays (Fredrikse 2000). The results reported here, suggest that old age cataracts should be re-examined for the presence of fibrillar aggregates containing human γD crystallin.

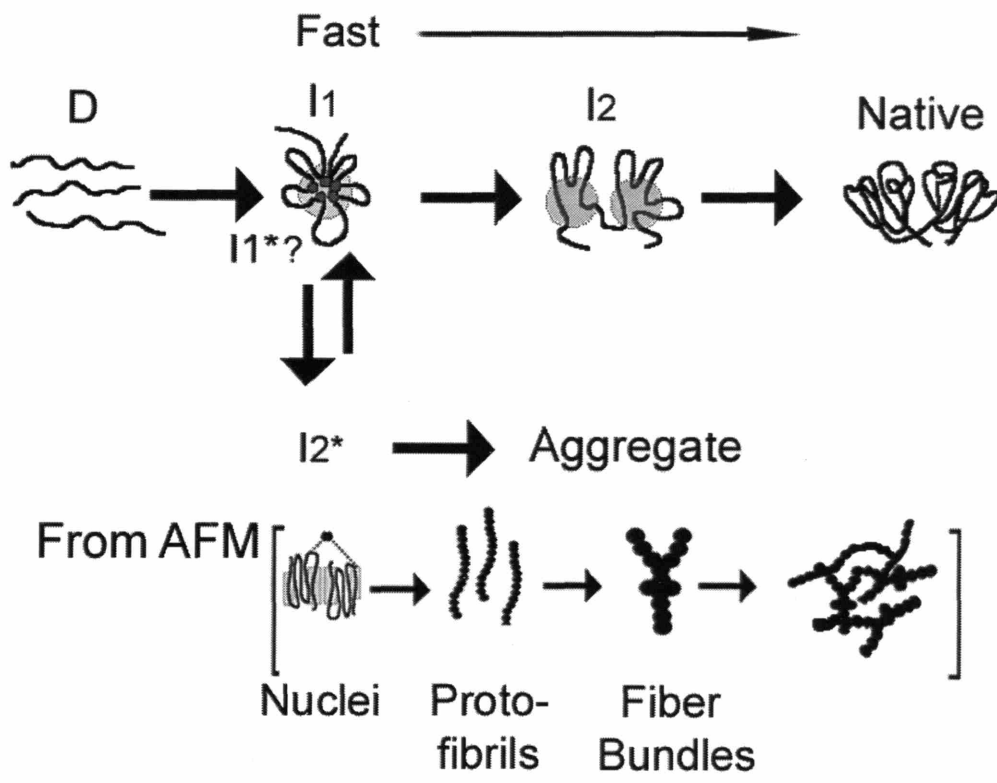


Figure 2-10

Model of HyD-Crys folding and aggregation. Upon rapid dilution into refolding buffer, HyD-Crys undergoes a global hydrophobic collapse. A fraction of the refolding molecules proceed to refold rapidly into a native state through a series of spectroscopically observable intermediates. The remaining fraction of the refolding chains will undergo a specific non-native interaction sequence ultimately resulting in the formation of an elongated aggregate. Spectroscopic techniques reveal three observable aggregation pathway states (I*, and I2*) and AFM images show three distinct aggregation intermediates (nuclei, protofibrils, and fiber bundles). We hypothesize that multiple-fiber bundles can associate, ultimately forming an aggregate mass.

CHAPTER III: ASSIGNMENT OF TRYPTOPHAN FLUORESCENCE OF HUMAN γ D CRYSTALLIN USING TRIPLE TRYPTOPHAN TO PHENYLALANINE MUTANT PROTEINS

A. INTRODUCTION

Human mature onset cataracts affect nearly 15% of the US population over 40 years of age and are the leading cause of blindness worldwide (NEI 2002). Pathological studies of cataractous lenses have revealed that cataracts are composed of protein aggregates that precipitate or polymerize in lens cells of the eye (Oyster 1999).

Human γ D crystallin (H γ D-Crys) is a protein synthesized during embryonic development that must remain soluble in the anucleated cells of the adult human eye lens for proper vision. Covalently modified H γ D-Crys has been recovered in protein aggregates removed from aged, cloudy lenses (Hanson et al. 2000). H γ D-Crys has 173 amino acids and shows high sequence and structural similarity to other γ -crystallins (Basak et al. 2003). Mutations in the gene encoding H γ D-Crys have been found in families exhibiting juvenile onset cataracts further implicating H γ D-Crys in cataractogenesis (Nadrut et al. 2003; Heon et al. 1999; Pande et al. 2001).

H γ D-Crys is composed of anti-parallel β -sheets arranged in four Greek-key motifs separated into two domains (Figure 3-1). The two domains show high levels of structural and sequence conservation and appear to be the result of gene duplication during evolution (Wistow et al. 1983). Like most soluble γ -crystallins, H γ D-Crys is monomeric in solution (Jaenicke 1999). H γ D-Crys has four tryptophans that have been used to probe unfolding and refolding progression with fluorescence spectroscopy (Kosinski-Collins and King 2003).

The fluorescence signal of H γ D-Crys increases when the protein is denatured in high concentrations of guanidine hydrochloride (GdnHCl). This indicates that the tryptophan fluorescence is quenched in the native state.

Although this phenomenon has been observed in other proteins without metal ligands or cofactors (Lakowicz 1999; Eyles and Gierasch 1999; He et al. 2001), it is uncommon. Ultraviolet light has been proposed as one of the etiological agents of cataract formation (McCarty and Taylor 2002; Sasaki et al. 2002). The presence of tryptophan in the lens proteins may function in protecting the retina from ultraviolet light damage (Kurzel et al. 1973). In this case, fluorescence quenching may protect the lens proteins from ultraviolet light absorption. We have investigated which residues are involved in the quenching reaction to provide possible insight into the ability of the lens proteins to maintain stability and transparency over the human lifetime.

We constructed triple mutant tryptophan constructs each containing only one of the four native tryptophans of HyD-Crys to determine the origin of the anomalous quenching signal.

B. MATERIALS AND METHODS

i. *Cloning and site-directed mutagenesis*

The human γ D crystallin coding sequence had been previously cloned as described (Pande et al. 2000). The gene was excised from the pET3a plasmid and ligated into a pQE.1 plasmid (Qiagen) that added an N-terminal 6-His tag to the protein. The integrity of the HyD-Crys gene was confirmed by sequencing at the facilities of Massachusetts General Hospital.

Four triple mutant proteins containing three tryptophan to phenylalanine substitutions at positions 42, 68, 130, and 156 were constructed using a PCR-based mutagenesis procedure (Stratagene). Sequential mutations were made using complementary primer pairs that altered the Trp codon TGG to Phe TTT (Invitrogen). The mutations were confirmed by sequencing the region of the resulting plasmids (MGH).

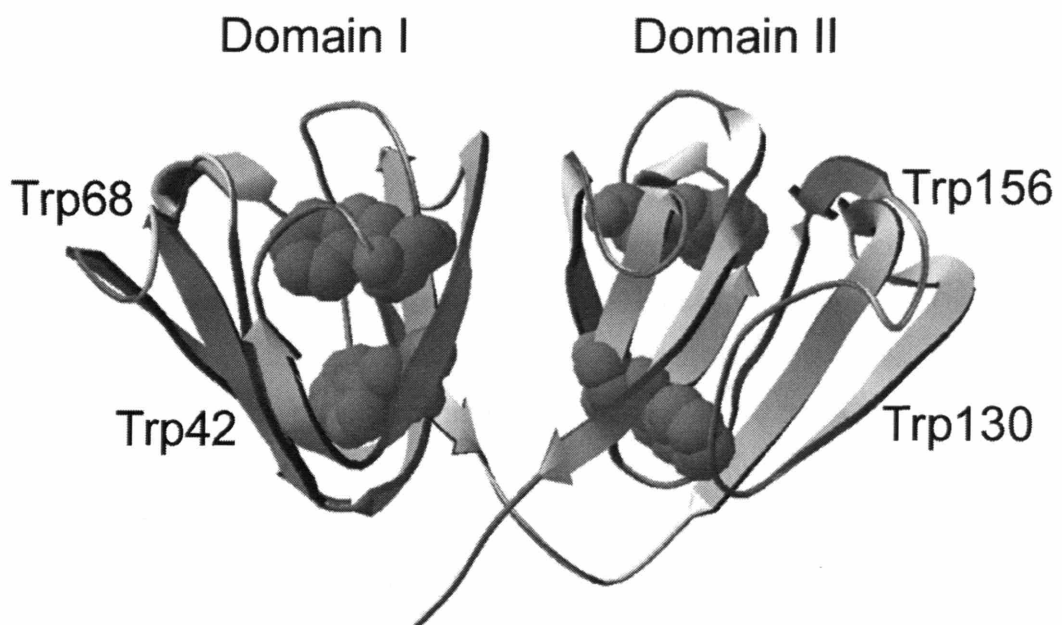


Figure 3-1

Ribbon structure of wild-type human γ D crystallin showing the location of the four native tryptophans at positions 42, 68, 130, and 156 (Basak et al. 2003).

ii. *Expression and purification of triple mutants*

Recombinant H γ D-Crys was prepared by pQE.1 plasmid transformation into *E. Coli* M15 (pREP4) cells. Protein production was induced by addition 1 mM IPTG and allowing 4 hours of incubation at 37°C. Cultures were pelleted by centrifugation for 15 min and cells were resuspended in a 10 mM imidazole, 10 mM Tris, 0.5 M NaCl solution. Cells were lysed using six sequential 20 second bursts of sonication followed by 40 second rest cycles. Lysates were spun at 17,000 RPM for 30 minutes. The resulting supernatant was then applied to a Ni-NTA column and protein was eluted using an increasing concentration of 250 mM imidazole, 10 mM Tris, 0.5 NaCl at room temperature. Fractions containing protein were dialyzed four times against four liters of 10 mM ammonium acetate pH 7.0 for four hours. Maldi mass spectroscopy was performed on all proteins to confirm the presence of the desired amino acid substitutions (MIT-Biopolymers Lab). All proteins were transferred from ammonium acetate to S Buffer (10 mM NaPO₄, 5 mM DTT, 1 mM EDTA, pH 7.0) by dilution.

iii. *Circular dichroism*

CD spectra of wild type and mutant H γ D-Crys proteins were collected on an Aviv Associates (Lakewood, NJ) model 202 circular dichroism spectrometer. All readings were performed on 0.3 mg/mL H γ D-Crys protein samples in S buffer at 37°C. CD was measured every 1 nm between 200 and 260 nm. The signals at all wavelengths were averaged over 5 seconds.

iv. *Ultraviolet light absorbance*

Ultraviolet light spectra of wild-type and mutant H γ D-Crys proteins were collected on a Varian Cary 50 Bio ultraviolet light spectrometer. Readings in the peptide

backbone region (190 nm to 240 nm) were performed on 0.1 mg/mL HyD-Crys in S buffer. Absorbance readings in the aromatic region (240 nm to 340 nm) were taken of 0.3 mg/mL protein samples in S buffer (native) or S buffer containing 5.5 M GdnHCl (denatured). Protein concentration was calculated by measuring denatured protein absorbance at 280 nm and using a protein extinction coefficient of 41.04 mM^{-1} for wild-type and 23.97 mM^{-1} for triple mutant tryptophan His-tagged constructs, respectively.

v. *Fluorescence emission spectra*

Fluorescence emission spectra were read on a Hitachi F-4500 fluorimeter with a continuous flow temperature control system. Proteins were diluted to a concentration of $10 \mu\text{g/mL}$ in S buffer or S buffer containing 5.5 M GdnHCl. Samples were excited at 295 nm and emission was measured from 310 nm to 420 nm. The excitation and emission slit widths were both set to 10 nm. The background fluorescence of S buffer or S buffer and 5.5 M GdnHCl was subtracted from the sample reading. Fluorescence emission maxima were calculated by averaging signals over every 5 nm and selecting the midpoint of the five signals that exhibited the highest average.

C. RESULTS

i. *Purification of his-tagged crystallin*

The cloned HyD-Crys gene (Pande et al. 2000) was excised from a pET3a plasmid and ligated into a pQE.1 plasmid as described in Material and Methods. This plasmid added an N-terminal MKHHHHHHQ peptide to HyD-Crys. The addition of this peptide did not affect expression of HyD-Crys or the ability of the protein to fold into a native-like state during purification. In addition, we were not able to detect differences in the fluorescence or circular dichroism spectra (CD) of the His-tagged protein. The thermodynamic and kinetic unfolding and refolding properties of the His-tagged species were identical to the wild-type protein (results not shown).

Four triple mutant proteins were constructed using site-directed mutagenesis. Each contained three tryptophan to phenylalanine substitutions retaining one native tryptophan (W42-only, W68-only, W130-only, and W156-only). All mutant plasmids were transformed into *E. Coli* M15(pREP4) cells and the proteins expressed during incubation at 37°C. Wild type and all four triple mutant proteins of HyD-Crys accumulated primarily in the soluble fraction of the cell lysates (>60%). These proteins were purified using the protocol developed for wild-type His-tagged HyD-Crys. The four mutant proteins behaved similarly to wild type during the purification procedure.

ii. *Structure assignment*

CD, native gel electrophoresis, and ultraviolet light absorbance were performed to assess the overall conformations of the mutant proteins.

Native gel electrophoresis of wild type and the triple mutant tryptophan constructs confirmed that all proteins retained similar native conformations. His-tagged wild-type HyD-Crys shows three distinct bands when separated on a native polyacrylamide gel due to different charge states and/or degradation products. All mutant constructs exhibited similar native bands running with analogous mobility as wild type (results not shown).

Far-UV CD of the native proteins showed a primarily β -sheet structure for wild-type HyD-Crys and all four triple mutant tryptophan constructs at 37°C, pH 7.0 (Figure 3-2). All five proteins had a characteristic β -sheet minimum at 218 nm although the amount of β -structure appeared to vary between the different constructs. In addition, none of the mutant proteins exhibited the inflection shown by wild-type HyD-Crys at 208 nm (Kosinski-Collins and King 2003) indicating that some of the secondary structure of wild type was diminished in the mutant proteins. The tryptophan to phenylalanine mutations may have slightly disrupted local β -sheet structures, leading to the altered CD signal at 218 nm, while maintaining overall conformations similar to wild type as observed by native gels.

The W68-only construct displayed the characteristic inflection of wild type observed in far-UV CD at 235 nm. W156-only showed a slight arc around this wavelength that was not as pronounced as W68-only. This may have been due to the

overall decreased signal of W156-only. Therefore, it seems that the inflection in the wild-type spectrum at 235 nm reports environments around Trp156 and Trp68.

Ultraviolet light absorbance scans in the peptide backbone region of native wild-type, W42-only, W68-only, W130-only, and W156-only HyD-Crys showed similar spectra between 190 nm and 240 nm at 37°C, pH 7.0 (results not shown). Ultraviolet light spectra in the region of tyrosine and tryptophan absorption were also obtained from native and denatured proteins at 37°C, pH 7.0. All proteins had absorbance spectra typical of polypeptides containing high numbers of tryptophan and tyrosine residues with a maximum at 276 nm. Absorption spectra of denatured states of the triple mutant tryptophan mutants overlaid uniformly indicating that all tryptophans were in similar denatured environments (Figure 3-3A). Ultraviolet light absorbance spectra of native triple mutant tryptophan proteins exhibited the same overall shape and character, but had slightly different overall absorbance intensities (Figure 3-3B). This indicated that Trp42, Trp68, Trp130, and Trp156 were found in slightly different tertiary environments.

iii. *Fluorescence spectra*

The overall character of the fluorescence emission spectra of wild-type, W42-only, W68-only, W130-only, and W156-only HyD-Crys were assessed by exciting the protein at 295 nm and observing fluorescence emission intensities from 310 to 420 nm in either S buffer or S buffer and 5.5 M GdnHCl (Figure 3-4 and Table 3-1). S buffer contained 10 mM NaPO₄, 5 mM DTT and 1 mM EDTA and was prepared at pH 7.0.

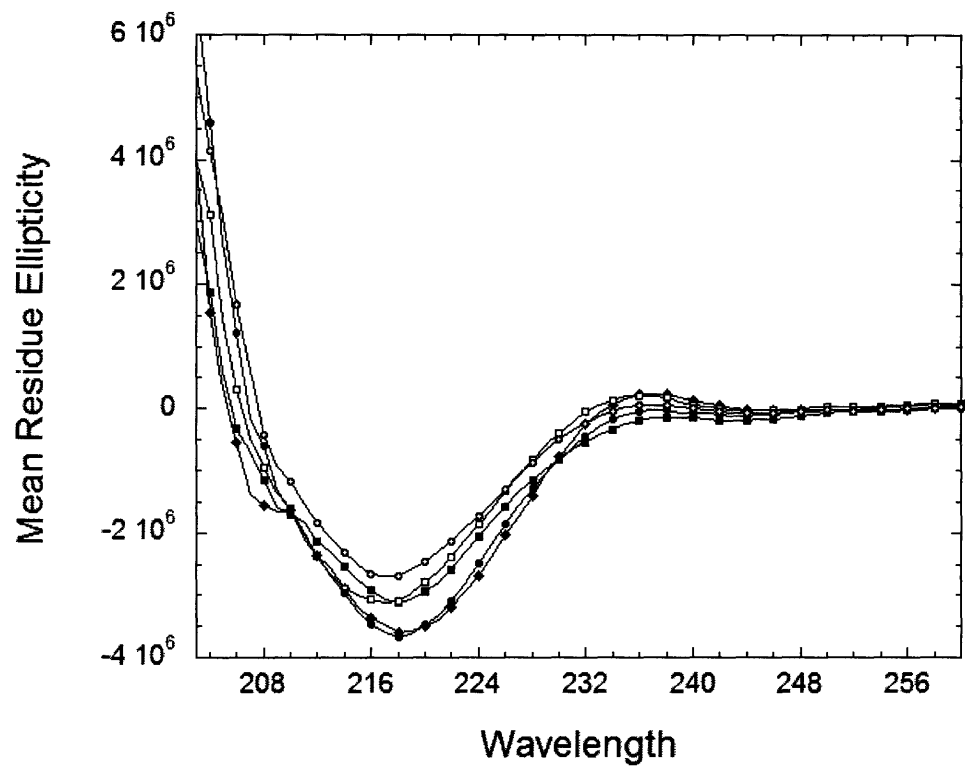


Figure 3-2

Far-UV CD of wild-type His-tagged HyD-Crys (◆), W42-only (■), W68-only (□), W130-only (●), and W156-only (○). Samples were prepared at a 300 $\mu\text{g/mL}$ protein concentration and equilibrated in S buffer at 37°C.

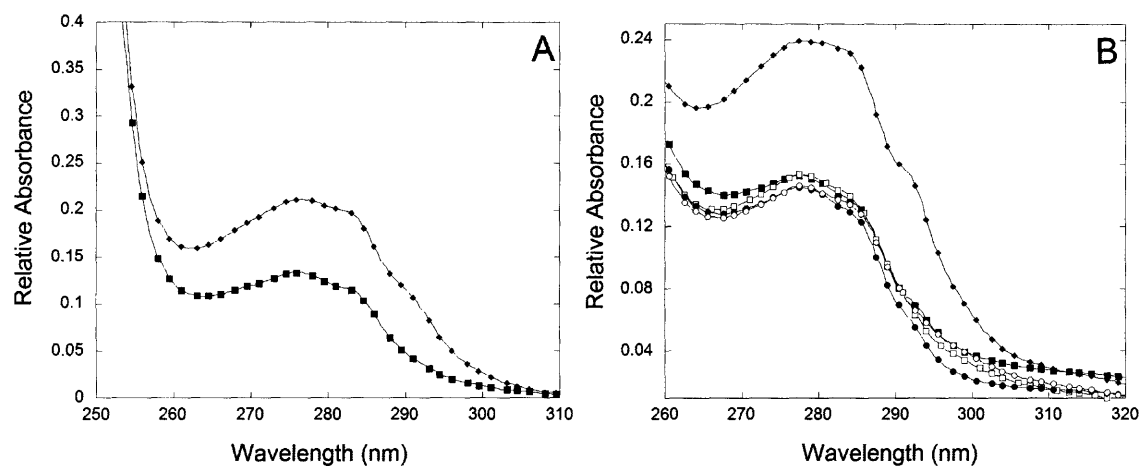


Figure 3-3

Aromatic residue ultra violet light absorbance spectra of His-tagged HyD-Crys at 100 $\mu\text{g/mL}$ protein in S buffer at 37°C. Wild type (\blacklozenge) and a representative triple mutant tryptophan protein, W42-only (\blacksquare) are shown denatured in 5.5 M GdnHCl (A). W68-only, W130-only, and W156-only exhibit denatured spectra indistinguishable from W42-only. Native protein aromatic absorbance is shown of wild-type His-tagged HyD-Crys (\blacklozenge), W42-only (\blacksquare), W68-only (\square), W130-only (\bullet), and W156-only (\circ) (B). Thirty percent of the data points are shown.

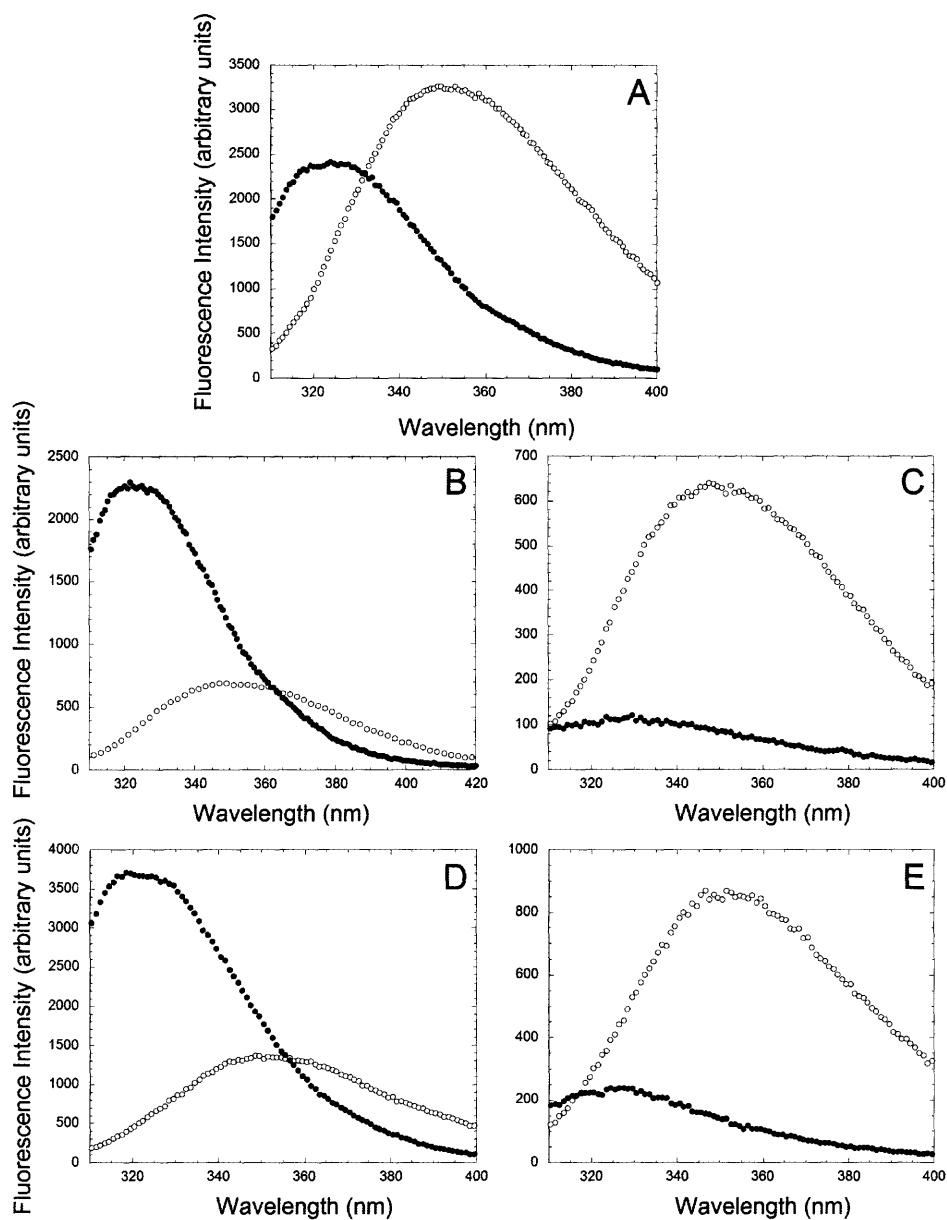


Figure 3-4

Fluorescence emission of native (●) and denatured (○) HyD-Crys. Protein was excited at 295 nm and emission spectra were collected from samples of 10 $\mu\text{g/mL}$ protein in S buffer or S buffer and 5.5 M GdnHCl at 37°C. Fluorescence spectra of wild-type His-tagged HyD-Crys (A), W42-only (B), W68-only (C), W130-only (D), and W156-only (E) are shown.

Wild-type His-tagged HyD-Crys had a native fluorescence emission maximum of 326 nm, and was quenched in the native state. The denatured protein exhibited a fluorescence emission maximum of 350 nm. This is similar to data previously reported for the non-His-tagged construct (Kosinski-Collins and King 2003) (Figure 3-4A). Similarly, all of the triple mutant tryptophan constructs had denatured fluorescence emission maxima of 350 nm. The native fluorescence emission maxima were 327 nm for W42-only, 329 nm for W68-only, 318 nm for W130-only, and 327 nm for W156-only (Table 3-1). W68-only and W156-only retained the native state quenching observed in wild type (Figure 3-4C and 3-4E), while W42-only and W130-only were quenched in the denatured state (Figure 3-4B and 3-4D). W130-only is significantly more fluorescent in its native state than any of the other triple mutant constructs.

Table 3-1

Fluorescence Spectra constants for triple tryptophan mutants of HyD-Crys

	Wild type	W42-only	W68-only	W130-only	W156-only
Native Fluorescence Emission Max	326 nm	327 nm	329 nm	318 nm	327 nm
Denatured Fluorescence Emission Max	350 nm	350 nm	350 nm	350 nm	350 nm
Exhibit Native State Quenching?	Yes	No	Yes	No	Yes

D. DISCUSSION

i. *Expression and structure of crystallin proteins*

H γ D-Crys is a highly stable protein containing four tryptophans buried within two highly hydrophobic protein cores. Four stable, triple mutants of His-tagged H γ D-Crys were produced each containing one native tryptophan and three tryptophan to phenylalanine substitutions. All of the mutant proteins were expressed in *E. coli* in the soluble fraction of the cell lysate. In addition, each had a primarily β -sheet character by circular dichroism measurements and displayed bands of a similar mobility as wild type on a native polyacrylamide gel.

The triple tryptophan mutants display no increased or decreased propensity for aggregation over wild-type H γ D-Crys during either refolding or long-term storage at 4°C. The overall solubility and native-like structure of H γ D-Crys with phenylalanines substituted for tryptophans suggests that the tryptophan residues are not critical for stability at physiological temperature. Likewise, mutation of each tryptophan and substitution with alanine report similar stability (Veronica Zepeda, Melissa Kosinski-Collins, Shannon Flaugh, and Jonathan King, unpublished results). Given the location of these substitutions in the hydrophobic core of H γ D-Crys, the high solubilities and native-like structures of all mutant tryptophan constructs are surprising.

All four tryptophans in H γ D-Crys show high chemical conservation throughout the known members of the γ -crystallin family. Trp42 and Trp130 are 100% conserved throughout the γ -crystallin family as aromatic residues, while Trp68 and Trp156 are 80% and 89% conserved in aromaticity, respectively. The deviations in aromatic composition in Trp68 and Trp156 are found in primarily aquatic animals such as carp and catfish. It is interesting to note that the eyes of these animals are shielded from direct light by their underwater surroundings.

ii. *The quenching of tryptophan emission in the native state*

Based on fluorescence spectra data collected from the triple mutant tryptophan proteins, it appears that Trp68 and Trp156 are responsible for the native state quenching phenomenon in HyD-Crys. Trp68 is in domain I and Trp156 is in domain II. Examination of the 3-dimensional structure of HyD-Crys revealed that both are located in similar positions within the domains.

Because both Trp68 and Trp156 exhibit anomalous quenching, it seemed likely that the residues responsible for this reaction would be similar in the two domains, given the high sequence similarity between domain I and domain II of HyD-Crys. Inspection of the residues within 8 Å of Trp68 and Trp156 showed that the only potential quenchers that were constant between the two domains were two tyrosine residues and histidine. Tyr55, Tyr62, and His65 make a “cage” around Trp68 (Figure 3-5A), while Tyr143, Tyr150, and His122 surround Trp156 (Figure 3-5B). Chen and Barkley have suggested that tyrosine may quench tryptophan fluorescence via a proton-transfer mechanism while histidine may quench via excited state electron transfer (1998). Both interactions have geometry requirements for quenching. As histidine quenches primarily when protonated and as these experiments were performed at pH 7.0, tyrosine is the likely side chain participating in the quenching phenomenon.

Both Trp42 and Trp130 are located within quenching distance of cysteine and histidine side chains. In the original report of the crystal structure of bovine γ B crystallin, Wistow et al. suggested that these interactions may have been the source of the anomalous quenching (1983). Though the results reported here make that mechanism less likely, the cysteine and histidine residues surrounding Trp42 and Trp130 may protect these tryptophans from photo damage *in vivo*.

iii. *Tryptophan quenching and protection from UV radiation*

All lens proteins are subject to irradiation in the visible and ultraviolet light range during the human lifetime. The cornea of the eye absorbs the majority of light at wavelengths less than 295 nm, allowing ultraviolet light at wavelengths longer than 295

nm to pass through to the lens and to the crystallin proteins (Sloney 2002). Although the peak of tryptophan absorbance occurs at 278 nm, like most tryptophan containing proteins, H γ D-Crys shows an absorption tail that extends well beyond 300 nm (3-3A).

It is not obvious why the tryptophan residues are conserved in the γ -crystallin family, but these residues are responsible for the major observed ultraviolet light absorption events of γ -crystallins in vitro. The well-conserved tryptophans in the γ -crystallins may have been maintained during evolution as a part of the mechanism of protecting the retina and other eye structures from UV-B damage. This, however, would then render the γ -crystallin proteins themselves sensitive to ultraviolet light damage. In epidemiological studies, excess UV-B exposure has been shown to be directly correlated with an increase in mature onset cataract in humans (Taylor et al. 1988; Sloney 2002). More specifically, photo oxidation of tryptophan residues is known to be a precursor in the formation of brunescant cataracts and is thought to occur as a result of prolonged exposure to ultraviolet light (Pirie 1971; Kruzel et al. 1973; Zigman et al. 1973; Davies and Truscott 2001; Soderberg et al. 2002). To avoid covalent modification associated with extended exposure of tryptophan residues to ultraviolet light, the crystallins would require a means to dissipate the absorbed excited state energy.

H γ D-Crys was more fluorescent in its native state than in its denatured state (Figure 3-4). Although not emphasized in the literature, γ S and γ B crystallins show this anomalous quenching of their buried tryptophans as well (Rudolf et al. 1990; Wenk et al. 2000). The existence of native-state quenching of tryptophans in a protein exposed to ultraviolet light, but selected for its stability and solubility, may reflect a role in protecting it from the absorption events. Prolonged exposure to ultraviolet light or other oxidative conditions populated in the lens over time by H γ D-Crys may generate ring opening and other covalent damage well characterized in other systems (Balasubramanian et al. 1990; Prinsze et al. 1990; Conti et al. 1988). The generation of a charged species within the buried core of the crystallins would be expected to destabilize the protein and cause full or partial unfolding. These species would be candidates for precursors to the aggregated state of the crystallins found in mature onset cataracts.

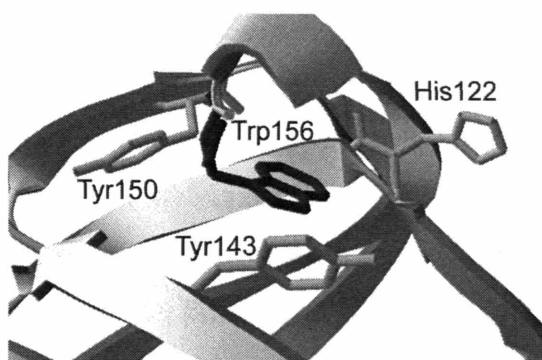
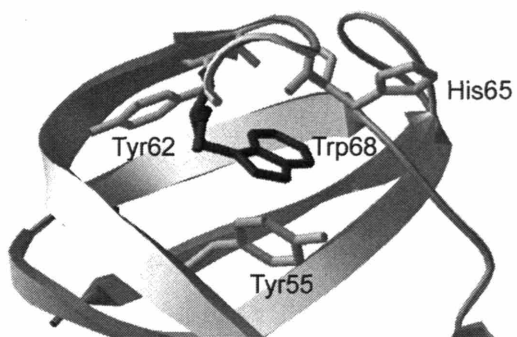


Figure 3-5

X-ray structure of Tyr-His-Tyr aromatic “cage” surrounding Trp68 (A) and Trp 156 (B). Trp68 is enclosed by Tyr55, Tyr62, and His65, while Tyr143, Tyr150, and His122 surround Trp156 (Basak et al. 2003).

CHAPTER IV: IDENTIFICATION OF AN INTERMEDIATE DURING UNFOLDING AND REFOLDING OF HUMAN γ D CRYSTALLIN USING TRIPLE TRYPTOPHAN TO PHENYLALANINE MUTANT PROTEINS

A. INTRODUCTION

Human γ D crystallin could be refolded to its native state at 37 °C after dilution out of denaturant (Kosinski-Collins and King 2003). An *in vitro* aggregation pathway of H γ D-Crys that competed with productive refolding was identified and may be related to the mechanism of its involvement in mature onset cataracts (Kosinski-Collins and King 2003). The structures of the intermediates in the aggregation pathway were studied using atomic force microscopy and consisted of ordered, distinct intermediates. Using fluorescence spectroscopy, a partially-folded hydrophobically collapsed intermediate was also identified in the productive refolding pathway. Within the lens, partially unfolded crystallin intermediates are likely to be recognized by the chaperone, α -crystallin (Fu and Liang 2003; Cobb and Petrash 2002; Bron et al. 2000).

Given the two-domain structure of H γ D-Crys, it seemed possible that the intermediate identified in the aforementioned kinetic experiments had one intact domain and one partially unfolded domain. Previous studies of a bovine homolog of H γ D-Crys, γ B crystallin, showed that the protein could be denatured by urea at pH 2.0, but not at pH 7.0, and could be refolded over all reported concentrations of urea. The protein exhibited a three-stage transition in these equilibrium studies, representing sequential denaturation of the C-terminal and N-terminal domains at pH 2.0 (Rudolph et al. 1990; Mayr et al. 1997; Jaenicke 1999). Conversely, the closely homologous human γ S crystallin, a protein also containing two domains with four Greek keys, unfolded and refolded without evidence of separate domain transitions (Wenk et al. 2000). H γ D-Crys may possess differential domain stability that was not detected in the apparent two-state unfolding transition observed during equilibrium unfolding in GdnHCl. A partially-folded intermediate with only one domain structured might be involved in the aggregation pathway of H γ D-Crys.

We used triple mutant tryptophan constructs each containing only one of the four native tryptophans of HyD-Crys to provide reporters of conformation for different regions of the protein. This has made it possible to assess the thermodynamic stability and kinetic properties of the two domains individually.

B. MATERIALS AND METHODS

i. *Equilibrium refolding and unfolding*

For the unfolding process, purified wild-type or mutant HyD-Crys was diluted to 10 $\mu\text{g/mL}$ in increasing amounts of GdnHCl in S buffer from 0 to 5.5 M. The samples were incubated at 37°C until equilibrium was reached (about 6 hours). For the refolding titration, 100 $\mu\text{g/ml}$ protein was denatured in 5.5 M GdnHCl in S buffer at 37°C for five hours. The protein was subsequently refolded by dilution to 10 $\mu\text{g/ml}$ into decreasing concentrations of GdnHCl from 5.5 to 0.55 M. The fluorescence spectra of the equilibrated samples were determined using a Hitachi F-4500 fluorimeter equipped with a continuous temperature control system with an excitation wavelength at 295 nm and emission monitored from 310 to 420 nm. The excitation and emission slits were both set to 10 nm. The ratios of emission intensities of 360 nm over 320 nm were used for data analysis of wild type, W68-only, and W156-only, W42-only and W130-only. Fraction unfolded values were calculated using the method of Pace et al. (1989) and denaturation midpoints were calculated using the Kaledagraph (Synergy Software) curve fitting function.

ii. *Unfolding fluorescence kinetics*

Tryptophan environment changes with refolding were monitored using a Hitachi 4500 fluorimeter equipped with a continuous temperature control system. Native protein

(100 $\mu\text{g}/\text{mL}$ in S buffer at 37°C) was unfolded by dilution into S buffer to final concentrations of 10 $\mu\text{g}/\text{ml}$ HyD-Crys and 5.5 M GdnHCl using a syringe port injection system exhibiting a dead-time of 1 second. Loss of global structure was monitored with continuous excitation at 295 nm at 37°C for 1 hour. Emission intensities during kinetic unfolding were collected at 350 nm for wild type, W68-only, and W156-only while the emission intensities were collected at 320 nm for W42-only and W130-only. The fluorescence curves were fit to series of consecutive first-order exponentials using the method described by Fersht (1999). The signals were all fit using the KaliedaGraph (Synergy Software) curve fitting algorithm to mechanisms having one, two, and three exponentials, and the best fit was selected by inspection. All proteins were unfolded in at least two separate experiments to ensure the accuracy of the observed fluorescence and curve fitting given the high levels of noise.

iii. Refolding fluorescence kinetics

Changes in tryptophan environment during refolding were monitored using a Hitachi 4500 fluorimeter equipped with a continuous temperature control system. Native protein was denatured at 100 $\mu\text{g}/\text{mL}$ in 5.5 M GdnHCl in S buffer at 37°C for two hours. The unfolded protein was refolded by dilution into S buffer to final concentrations of 10 $\mu\text{g}/\text{ml}$ HyD-Crys and 1.0 M GdnHCl using a syringe port injection system with a dead-time of 1 second. Increase in global structure during refolding was monitored with continuous excitation at 295 nm and emission at 350 nm for wild type, W68-only, and W156-only and emission at 320nm for W42-only and W130-only. The fluorescence curves were fit to series of consecutive first-order exponentials using the method described by Fersht (1999). The signals were all fit using the KaliedaGraph (Synergy Software) curve fitting algorithm to mechanisms having one, two, and three exponentials, and the best fit was selected by inspection. These experiments were repeated for each protein refolding to 1.0 M and 1.5 M GdnHCl.

C. RESULTS

i. *Equilibrium Refolding and Unfolding*

Equilibrium unfolding and refolding experiments were performed on the triple mutant tryptophan constructs of HyD-Crys to determine the stability of the individual domains. Fluorescence spectra were collected as a function of GdnHCl concentration at 37°C as previously described (Kosinski-Collins and King 2003).

All four proteins showed a midpoint for denaturant-induced unfolding in the range of 1-2 M GdnHCl. All four proteins also exhibited reversible refolding above 1.0 M GdnHCl. However, upon dilution to lower concentrations of denaturant, the proteins exhibited a polymerization behavior similar to that described for wild-type HyD-Crys (Kosinski-Collins and King 2003). Solution turbidity measurements of refolded samples confirmed that the apparent increase in fraction unfolded values at low GdnHCl concentrations was caused by signal obstruction due to aggregate formation (results not shown). The equilibrium unfolding and refolding results are summarized in Table 4-1.

The equilibrium unfolding and refolding data were similar for mutant proteins with tryptophans located in the same domain (Figure 4-1). The midpoint of the denaturation transitions was 1.3 M for the two constructs that retain a native tryptophan in domain I (W42-only and W68-only) and 2.0 M for those in domain II (W130-only and W156-only). Only W156-only exhibited the slight hysteresis between unfolding and refolding observed for wild type, having a midpoint of renaturation of 1.7 M. The equilibrium unfolding and refolding transition of W42-only, W68-only, and W130-only did not exhibit hysteresis. Further studies of residues in domain II near Trp156 may elucidate the discrepancies in the unfolding and refolding pathway that are causing the observed hysteresis.

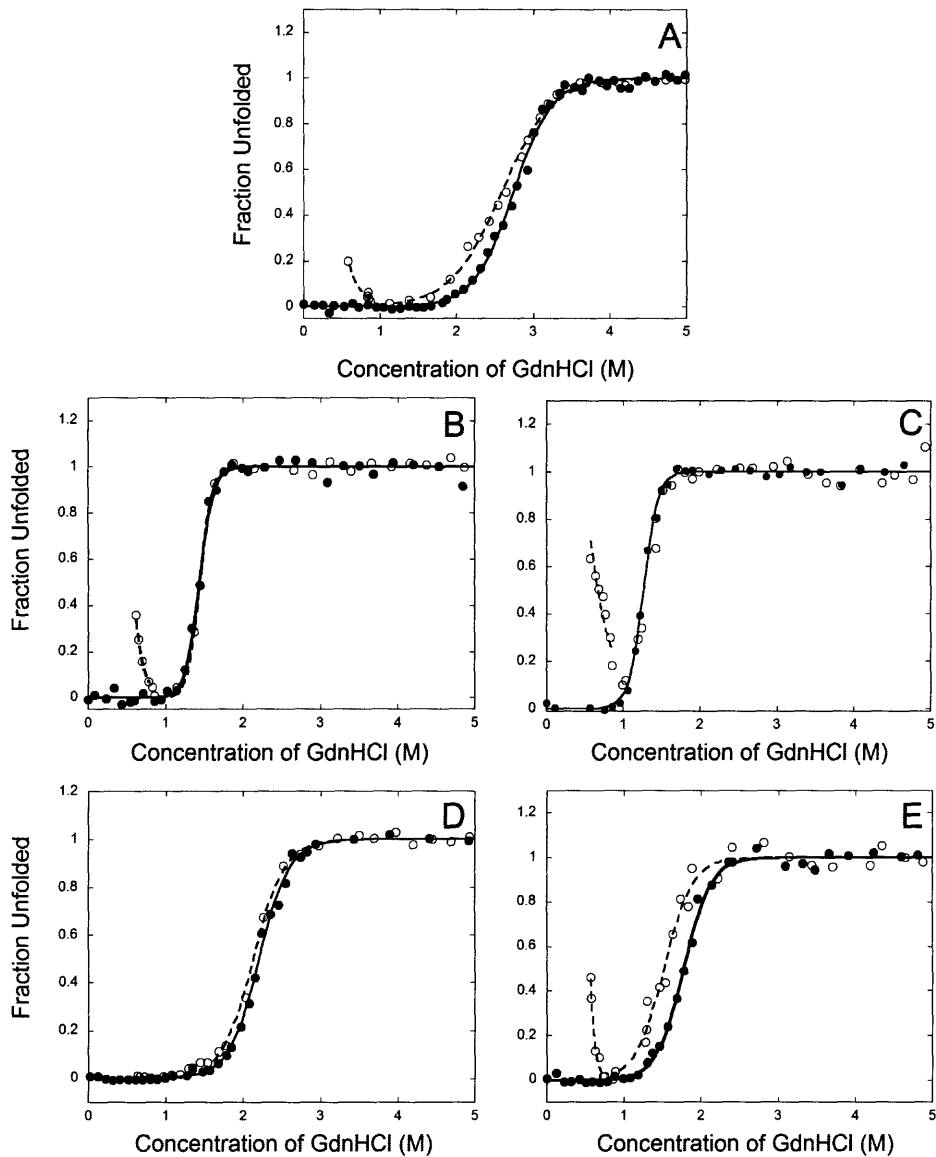


Figure 4-1

Equilibrium unfolding and refolding of His-tagged HyD-Crys in GdnHCl. Tryptophan fluorescence was monitored during unfolding and refolding and all samples were equilibrated in S buffer at a protein concentration of 10 $\mu\text{g/ml}$. A representative set of fraction unfolded data is shown for wild-type (A), W42-only (B), W68-only (C), W130-only (D), and W156-only (E). All protein data were analyzed using the ratio of fluorescence emission intensities at 360 nm over 320 nm. Fraction unfolded values were calculated from raw fluorescence intensity ratio measurements using the method described by Pace et al. (1989). Unfolding (\circ) and refolding (\bullet) transitions are presented for each protein at 37°C.

Based on the resistance of W42-only and W68-only to solvent denaturation, domain I was less stable than domain II. Domain I exhibited an unfolding and refolding transition midpoint at 1.3 M GdnHCl whereas domain II had a midpoint of 2.0 M GdnHCl. This indicated that wild type equilibrium unfolding and refolding probably contained a partially denatured intermediate not readily visible in the fluorescence spectra of wild-type chains. However, it is important to note that if the tryptophan residues are vital to stability of HyD-Crys, triple substitutions would likely affect the overall thermodynamic parameters of the molecule. The altered transition midpoints of the mutant proteins may simply reflect a global destabilization of the molecule due to these changes.

Table 4-1

Equilibrium unfolding and refolding constants for the triple tryptophan mutant proteins of HyD-Crys

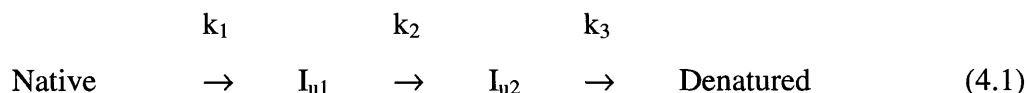
	Wild type	W42-only	W68-only	W130-only	W156-only
Native Fluorescence Emission Max	326 nm	327 nm	329 nm	318 nm	327 nm
Denatured Fluorescence Emission Max	350 nm	350 nm	350 nm	350 nm	350 nm
Exhibit Native State Quenching?	Yes	No	Yes	No	Yes
Exhibit Aggregation at Low [GdnHCl]?	Yes	Yes	Yes	Yes	Yes
[GdnHCl] at ½ Denaturation	2.8 M	1.3 M	1.3 M	2.0 M	2.0 M
[GdnHCl] at ½ Renaturation	2.1 M	1.3 M	1.3 M	2.0 M	1.7 M

ii. Unfolding kinetics

The unfolding kinetics of HyD-Crys were studied by dilution of native protein into 5.5 M GdnHCl and S buffer at 37°C, pH 7.0 and monitoring the changes in the fluorescence emission (Figure 4-2 and Table 4-2). A syringe injection port that exhibited a dead-time of approximately 1 second was used as the mechanism of dilution. Figure 3-6A shows the change in raw fluorescence signal for the four mutant proteins. Because the quenching characteristics vary for each of the mutants, fluorescence emission of W42-only and W130-only decreased upon denaturation whereas emission of W68-only and W156-only increased upon unfolding. In figure 4-2B, the data have been normalized between the native and denatured states for ease of visual comparison. As shown in figure 4-2A, the major changes in fluorescence took place in the first thirty seconds of the reaction. Millisecond time-scale intermediate(s) may have formed within the dead-time of these experiments and are not directly addressed here.

The kinetic data suggest the presence of partially unfolded intermediates in the transition between the native and unfolded states.

Wild-type His-tagged HyD-Crys was best fit with a four-state unfolding pathway (Fersht 1999). The protein started as a native species and then formed two sequential intermediates before becoming completely unfolded. In this model an early intermediate (I_{u1}) was populated within a $t_{1/2}$ of 1.0 s (Figure 4-2). A second partially unfolded intermediate (I_{u2}) followed, forming with a $t_{1/2}$ of 55 s. This intermediate was not as quenched as native indicating the polar-tryptophan interaction had been disrupted. A final unfolding transition ($I_{u2} \rightarrow$ denatured) occurred with a $t_{1/2}$ of 120 s.



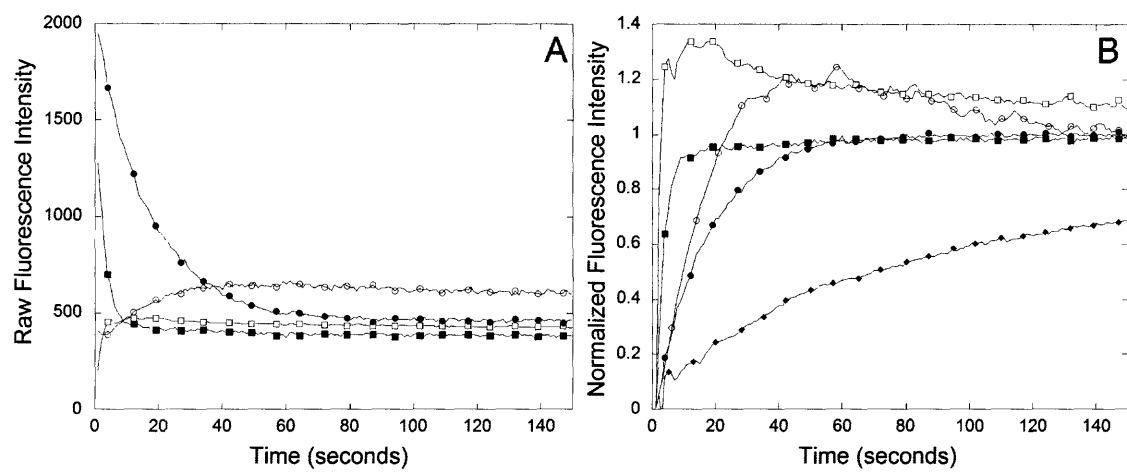


Figure 4-2

Unfolding kinetics of His-tagged HyD-Crys monitoring intrinsic tryptophan fluorescence with excitation at 295 nm. Emission was monitored at 350 nm for wild type, W68-only and W156-only while emission at 320 nm was used for W42-only and W130-only. HyD-Crys was denatured by rapid dilution into 5.5 M GdnHCl at 37°C in S buffer to a final protein concentration of 10 µg/ml. A representative protein unfolding time course is shown for wild-type His-tagged HyD-Crys (◆), W42-only (■), W68-only (□), W130-only (●), and W156-only (○). Data are shown over the first 150 seconds of unfolding, although the data were collected over the entire unfolding process (~1 hour). Raw fluorescence emission values (A) and data normalized between native and denatured emission intensities (B) are shown.

The unfolding process for the domain I tryptophan constructs, W42-only and W68-only, exhibited two transitions corresponding to the early unfolding steps of wild type. These two proteins exhibited changes in fluorescence signal that best fit to a three-state unfolding model. Both proteins unfolded to the intermediate conformation with an initial $t_{1/2}$ of 1.4 s. W42-only had a subsequent transition from intermediate to denatured with a $t_{1/2}$ of 18 s and W68-only had a secondary $t_{1/2}$ of 46 s. These values represent rate constants that are similar to the k_1 and k_2 rate constants observed in wild type (Table 4-2).

In comparison, the changes in the fluorescence signals upon unfolding for the domain II tryptophan constructs, W130-only and W156-only, proceeded more slowly than the changes in emission for domain I. The unfolding curves for tryptophans from domain II were best fit to a three-state unfolding process involving a partially unfolded intermediate. Neither mutant had a k_1 rate constant that was as large as that observed for wild type. Instead, the domain II mutants showed an unfolding process with rate constants that were similar to the k_2 and k_3 values of wild type. W130-only had $t_{1/2}$ values of 12 s and 150 s for its two unfolding transitions, and W156-only had comparable values of 15 s and 36 s.

When comparing the tryptophan residues in homologous positions in the two domains, the unfolding curve trends were similar. The fluorescence signal observed for Trp42 was similar to that of Trp130 in that both had two unfolding transitions and populated an partially unfolded intermediate. The overall trend of the unfolding curve was similar for Trp68 and Trp156 as well and both had a hyper-fluorescent folding intermediate. This intermediate was likely a result of rapid relaxation of the tertiary structure surrounding Trp68 and Trp156 resulting in a partial release of the tryptophan-quencher interaction.

In addition, examination of tryptophans in homologous domain positions showed that domain I unfolded before domain II. Trp42 had a more rapid change of global environment than Trp130 while the fluorescence emission of Trp68 increased its fluorescent signal more rapidly than Trp156. Wild type exhibited both of these transitions and probably underwent sequential unfolding in which domain I unfolded before domain II and an intermediate was populated with domain II folded but domain I denatured.

Table 4-2

Unfolding kinetic rate constants for the triple tryptophan mutants of HyD-Crys

	k_1	$t_{1/2}$	k_2	$t_{1/2}$	k_3	$T_{1/2}$
Wild type	0.2	3.5 s	0.038	18 s	0.0057	120 s
W42-only	0.1	7.9 s	0.039	18 s	NA	NA
W68-only	0.2	3.5 s	0.015	46 s	NA	NA
W130-only	NA	NA	0.060	12 s	0.0045	150 s
W156-only	NA	NA	0.046	15 s	0.019	36 s

iii. Refolding kinetics

In order to study the kinetic rates and intermediates formed during refolding of HyD-Crys, fluorescence emission was monitored during refolding of denatured HyD-Crys at 37°C to a final denaturant concentration of 1.0 M GdnHCl in S buffer (Figure 4-3, Table 4-3). Solution turbidity scans affirmed that no aggregate was formed under these conditions for any of the proteins described.

Figure 4-3A shows the changes in the raw fluorescence signal during refolding for the four triple tryptophan mutants. The major transitions occurred within the first 500 seconds of refolding. For ease of visual comparison, Figure 4-3B shows the refolding fluorescence signals normalized between the denatured and native states. Intermediates formed within the first second of refolding were not detected in these experiments.

Wild-type HyD-Crys was best fit to a three-state model suggesting the presence of one partially folded intermediate along the productive refolding pathway (Fersht 1999). The refolding reaction exhibited an early transition from denatured to a partially refolded intermediate with a $t_{1/2}$ of 15 s and a second transition of intermediate to native with a $t_{1/2}$ of 190 s. The intermediate was more fluorescent than the denatured protein and the native state was more fluorescent than the intermediate.

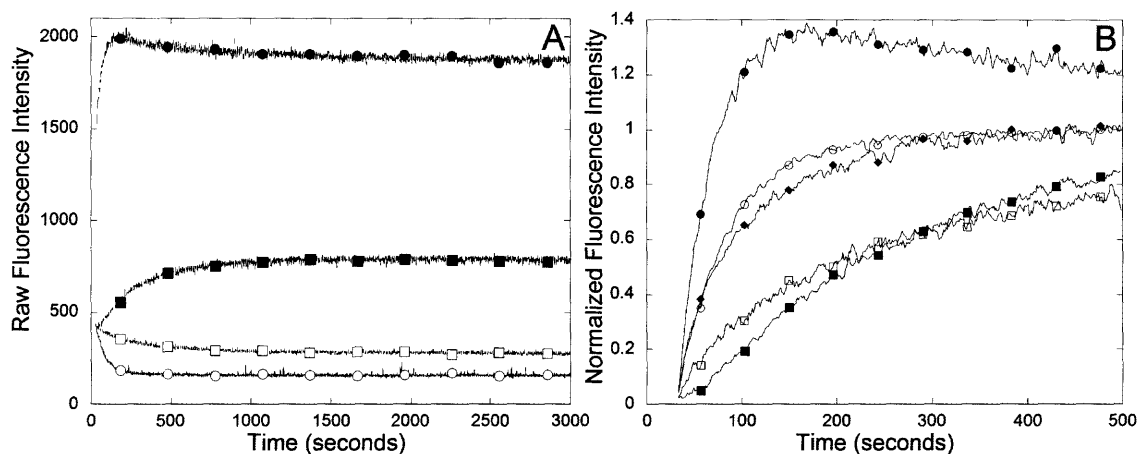


Figure 4-3

Refolding kinetics of His-tagged HyD-Crys monitoring intrinsic tryptophan fluorescence with excitation at 295 nm. Emissions were measured at 350 nm for wild type, W68-only and W156-only while emission wavelengths of 320 nm were used for W42-only and W130-only. HyD-Crys was denatured in 5.5 M GdnHCl at 37°C in S buffer for 3 hours. HyD-Crys was refolded by rapid dilution with S buffer to a final GdnHCl concentration of 1.0 M and a final protein concentration of 10 $\mu\text{g/ml}$ for wild-type His-tagged HyD-Crys (\blacklozenge), W42-only (\blacksquare), W68-only (\square), W130-only (\bullet), and W156-only (\circ). A representative scan of raw fluorescence signals (A) and signals normalized between native and denatured fluorescence values (B) are shown. Data are shown over the first 5000 seconds for raw fluorescence or 500 seconds for normalized values for refolding, although the data was collected over 2 hours.



The fluorescence signals of the domain I tryptophans reached a native-like state more slowly than the signals from the domain II tryptophans. The fluorescence of unfolding curves observed for W42-only and W68-only could be best fit to one exponential suggesting a two-state refolding pathway. These chains refolded with $t_{1/2}$ values of 190 s and 210 s respectively and neither populated an observable intermediate. The refolding transitions of these domain I tryptophans correspond closely with the second transition observed in wild type. The domain I constructs do not account for the burst fluorescence kinetics observed in wild type.

The domain II tryptophan constructs, W130-only and W156-only underwent a refolding process which was initially faster than domain I. These two proteins had fluorescence refolding signals that were best fit to a three-state model suggesting the existence of an intermediate. Both had initial $t_{1/2}$ values that were similar to the transitions observed for wild type. W130-only had an initial $t_{1/2}$ of 27 s and a secondary $t_{1/2}$ of 300 s while the transitions of W156-only had values of 31 s and 150 s, respectively.

W130-only had a hyper-fluorescent intermediate and was the only construct displaying a significant transformation within the dead-time of these experiments as shown by a very high initial fluorescence signal upon the onset of refolding (Figure 4-3A). It is possible that a millisecond refolding intermediate was populated by W130-only that was not readily visible using the syringe injection port system. Further studies using stopped-flow devices should be performed on this protein to assess the significance of this putative early transformation in the refolding pathway.

From these data, it appeared that domain II refolded first and was then followed by refolding of domain I (Figure 4-3B). A global hydrophobic collapse likely occurred in domain II that was then slowly followed by tight packing of the domain II tryptophans into their proper orientation. This resulted in the population of an intermediate with a primarily intact domain II, but a denatured domain I. While Trp130 and Trp156 were being tightly packed, a nearly simultaneous hydrophobic collapse of domain I occurred as evidenced by the change in fluorescence of Trp42 and Trp68. This two-step process

resulted in the formation of a protein with native-like tertiary structure as seen by a stabilization in tryptophan fluorescence emission.

Table 4-3

Refolding kinetic rate constants for the triple tryptophan mutants of HyD-Crys

	k_1	$t_{1/2}$	k_2	$t_{1/2}$
Wild type	0.047	15 s	0.0036	190 s
W42-only	NA	NA	0.0036	190 s
W68-only	NA	NA	0.0033	210 s
W130-only	0.022	30 s	0.0023	300 s
W156-only	0.022	30 s	0.0047	150 s

D. DISCUSSION

i. *Stability of crystallin triple tryptophan mutants*

Equilibrium unfolding and refolding experiments in GdnHCl demonstrated that the triple mutant tryptophan constructs were destabilized compared to wild type (Figure 4-1). Native to denatured transition midpoints of 1.3 M GdnHCl were measured for W42-only and W68-only and 2.0 M GdnHCl for W130-only and W156-only. Wild type had an equilibrium unfolding midpoint of 2.8 M GdnHCl. The decreased transition midpoints of the triple tryptophan mutants are likely due to destabilization of the molecule cause by the triple mutations. Substitution of three tryptophans for phenylalanines probably strains the hydrophobic cores of the molecules allowing GdnHCl to enter and denature these areas of the protein at lower concentrations than wild type. However, HyD-Crys with single tryptophans substituted by alanine folded into native-like soluble monomers with stabilities equal to that of wild-type HyD-Crys (Veronica Zepeda, Melissa Kosinski-Collins, Shannon Flaugh, and Jonathan King, unpublished results).

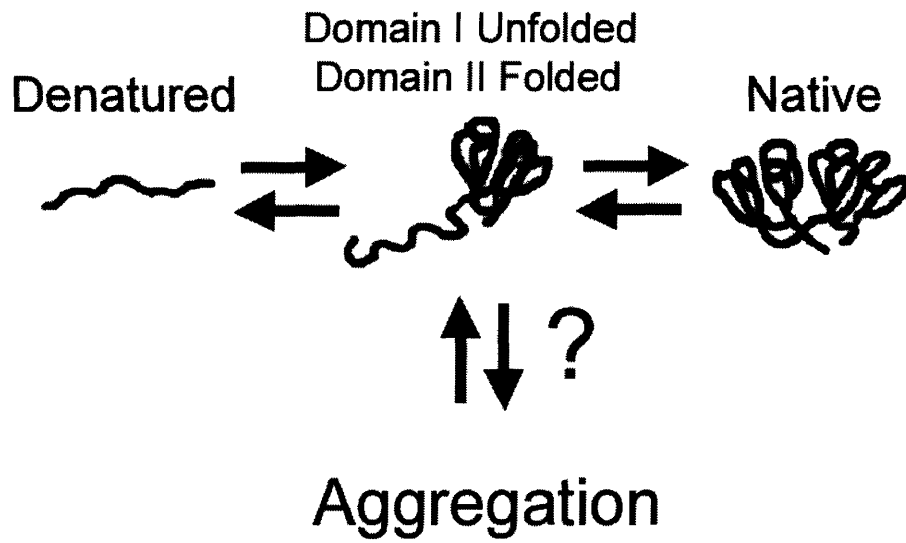


Figure 4-4

Model of HyD-Crys folding and aggregation. Upon rapid dilution into refolding buffer, denatured HyD-Crys had a putative intermediate that had domain II folded and domain I unfolded. During unfolding in GdnHCl, domain I unfolded earlier than domain II likely populating a similar intermediate.

ii. *Unfolding and refolding intermediates*

Kinetic analysis of H γ D-Crys revealed that unfolding and productive refolding involve a similar intermediate state in which tertiary structure is absent from domain I, but present in domain II (Figure 4-4). The population of this single domain conformer may represent an intermediate important in the *in vitro* aggregation pathway previously described (Kosinski-Collins and King 2003). Structurally distinct folding intermediates have been shown to be important in many aggregation pathways and disease systems (Haase-Pettingell and King 1988; Wetzel 1994; Speed et al. 1995). The domain-swapping model would provide a mechanism to explain polymerization of such partially-folded two-domain species into an ordered fibrillar state (Rousseau et al. 2003; Liu and Eisenberg 2002). Additionally, such partially folded species may be related to the crystallin conformers recognized by α -crystallin (Das et al. 1999; Cobb and Pettrash 2002). Specifically, α -crystallin has been shown to bind and recognize molten globule, partially folded states of the $\beta\gamma$ -crystallins as well as many other non-lens proteins such as alkaline phosphatase, alpha-lactalbumin, and apolipoprotein C-II (Tanksale et al. 2002, Bettelheim 2002, Hatters et al. 2001)

These data provide an explanation for tertiary structure of the domain cores. However, we do not yet know what contributions interface formation makes in the folding of H γ D-Crys. Future studies will investigate the kinetic and thermodynamic significance of domain interface residues in folding of the molecule.

CHAPTER V: CHARACTERIZING THE UNFOLDING AND REFOLDING PATHWAY OF HUMAN γ S CRYSTALLIN AND N143D; A DEAMIDATED MUTANT IDENTIFIED IN AGE-ONSET CATARACTS

A. INTRODUCTION

Human mature onset cataracts affect nearly 15% of the US population over 40 years of age and are the leading cause of blindness worldwide (NEI 2002). Pathological studies of cataractous lenses have revealed that cataracts are composed of protein aggregates that precipitate or polymerize in lens cells of the eye (Oyster 1999).

The proteins removed from cataractous human lenses are highly covalently modified and often have cysteine or methionine oxidations, premature truncations, and glutamine and asparagine deamidations, and/or tryptophan ring cleavages. It is still unclear whether these modifications are the causative agents of the cataracts or occur after the protein aggregates. It is possible, however, that changes such as these locally or globally destabilize the native state making it more susceptible to aggregation. Aggregation of a few protein molecules may seed the subsequent destabilization of other polypeptides as seen with other protein folding disorders such as Alzheimer's disease (AD), thus facilitating cataractogenesis.

The human α -crystallins are thought to perform chaperone-like functions in the lens, while the $\beta\gamma$ -crystallins are primarily structural proteins. The β -crystallins are multimeric in solution while the γ -crystallins are monomeric. Human γ C, γ D, and γ S crystallin are three γ -crystallins that are appreciably expressed in the lens.

Human γ S crystallin (H γ S-Crys) is the major protein component of the adult human lens and is primarily expressed in the lens fiber cells (Harding and Crabbe 1984). Covalently modified H γ S-Crys has been recovered in protein aggregates removed from aged, cloudy lenses. The majority of high molecular weight H γ S-Crys removed from cataractous lenses is disulfide bonded (Lapko et al. 2002, Takemoto and Boyle 2000, Takemoto 2001).

H γ S-Crys has 178 amino acids and shows high sequence similarity to other γ -crystallins. The protein is presumed to be a two domain protein with both domains

showing high levels of sequence conservation appearing to be the result of gene duplication during evolution (Jaenicke 2000). Like most soluble γ -crystallins, H γ S-Crys is monomeric in solution, but unlike other γ -crystallins, the interdomain interactions are weak (Wenk et al. 2000).

The C-terminal domain of H γ S-Crys has been crystallized and shows high structural similarity to other members of the crystallin family of proteins (Purkiss et al. 2002). Based on the crystal packing observed in x-ray diffraction of the C-terminal domain, a probable model of H γ S-Crys structure has been constructed using two interacting C-termini to reflect the position and structure of entire molecule. In this model, each domain of H γ S-Crys is composed of anti-parallel β -sheets arranged in two Greek-key motifs (Figure 5-1). There are four tryptophans in H γ S-Crys (two in each domain) that may be used to monitor fluorescence changes during unfolding and refolding.

The stability and folding of intact H γ S-Crys and its isolated N- and C-terminal domains have been studied at pH 7.0 at 20°C by Wenk et al. (2000). Equilibrium unfolding and refolding as a function of guanidine hydrochloride (GdnHCl) concentration demonstrated that H γ S-Crys displayed two-state reversible folding under these conditions, although Chevron plot analysis suggested the presence of a kinetic intermediate that was not directly identified. We have further investigated this intermediate in fluorescence studies in an attempt to elucidate its involvement in cataractogenesis *in vivo*.

Deamidated variants of H γ S-Crys are present in many positions in cataractous lenses including glutamine 92, glutamine 96, and asparagine 143 (Lapko et al. 2002; Takemoto 2001; Hanson et al. 1998). A study of the high molecular weight fraction of the insoluble portion of cataractous lenses revealed that, of these three positions, asparagine 143 was the preferentially deamidated polar amino acid of H γ S-Crys (Takemoto 2001). No deamidation of this residue, however, was detected in the normal lens (Takemoto and Boyle 2000). Harding and Crabbe have suggested that the conversion of a neutral amino acid to a charged, unprotonated amino acid occurring during deamidation may destabilize the protein perhaps favoring protein unfolding and

aggregation (1984). Studies of β B1-crystallins have shown that proteins deamidated in the connecting peptide have altered domain association characteristics (Harms et al.). We have investigated the *in vitro* unfolding, refolding, and aggregation pathways of N143D H γ S-Crys and compared these to wild type to determine if an alteration in folding makes the mutant protein more susceptible to polymerization in the lens.

B. MATERIALS AND METHODS

i. Expression and purification

Wild-type and deamidated forms of H γ S-Crys were expressed from a pQE plasmid (Qiagen) in *E. Coli* M15 cells. Transformed cells were grown in 100 μ g/mL Ampicillin and 25 μ g/mL Kanamycin at 37°C until log phase was reached. Protein expression was induced with addition of 1.5 mM IPTG and grown for an additional four to six hours. The cells were pelleted by centrifugation at 5,000xg for 10 minutes. The cells were then lysed with 50 mM Tris, 1 mM EDTA, 300 mM NaCl, pH 8.0 and subject to eight freeze-thaw cycles. The lysed pellets were dialyzed at 4°C three times against 0.1 M Na₂SO₄, 0.06 M NaH₂PO₄, pH 7.0. The protein was resolved on a SW3000 gel filtration column and the γ S peak was dialyzed against distilled water with three changes at 4°C. Protein concentration was determined using a Bradford assay and then the samples were lyophilized (Bradford 1976). The protein samples were stored at -20°C and were weighed and diluted directly into the buffers as needed for the described experiments.

His-tagged H γ D-Crys was purified as described from a pQE.1 plasmid (Qiagen) (Kosinski-Collins et al. 2003). The protein contained an N-terminal 6 histidine tag to assist with purification.

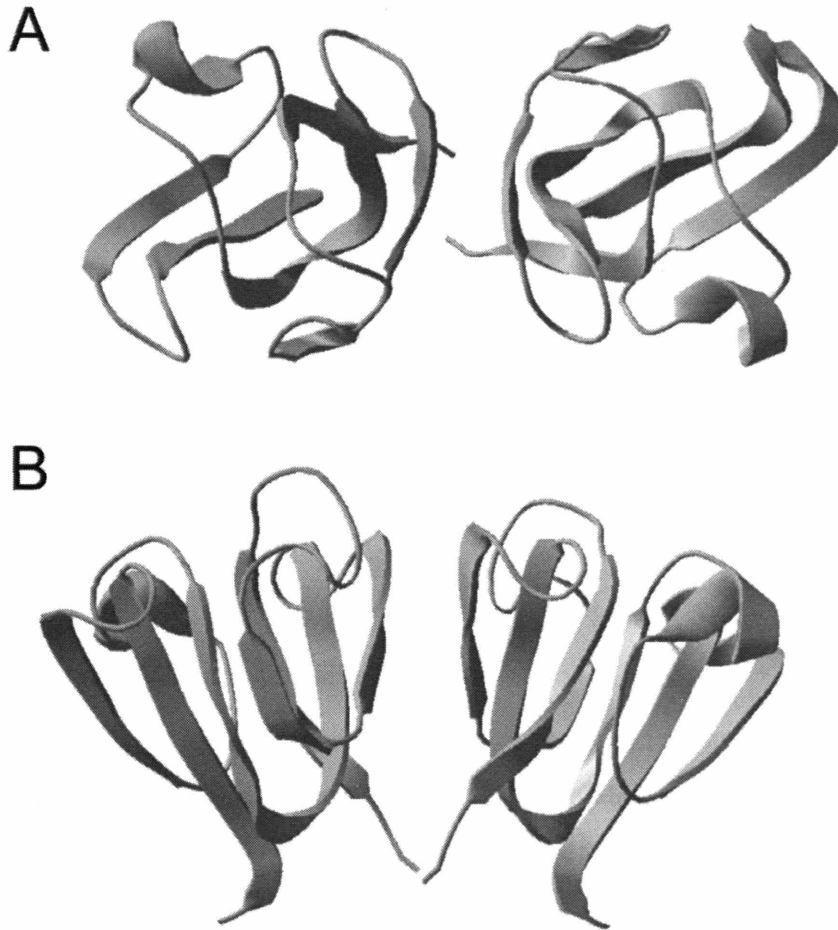


Figure 5-1

Ribbon structure of the crystal packing of two C-termini of wild-type human γ S crystallin showing the location of residue 143 (Purkiss et al. 2002). Both the top view (A) and side view (B) are shown.

ii. Circular dichroism

CD spectra of wild-type and mutant HyS-Crys proteins were collected on an Aviv Associates (Lakewood, NJ) model 202 circular dichroism spectrometer. All readings were performed on 0.3 mg/mL HyD-Crys protein samples in 10 mM NaPO₄, 5 mM DTT, 1 mM EDTA, pH 7.0. For wavelength spectra, CD was measured every 1 nm from 260 nm to 200 nm and averaged over 5 seconds. For thermal stability, CD was measured at 220 nm every 0.5°C between 20°C and 90°C. The protein was equilibrated at each temperature for one minute and the signals were averaged over 5 seconds.

iii. Ultraviolet light absorbance

Ultraviolet light spectra of proteins were collected on a Varian Cary 50 Bio ultraviolet light spectrometer. Concentration was calculated by measuring protein absorbance at 280 nm and using an extinction coefficient of 41.04 mM⁻¹ for both wild-type and N143D HyS-Crys and 41.04 mM⁻¹ for his-tagged HyD-Crys in 6.0 M GdnHCl. Wild-type and N143D HyS-Crys had molecular weights of 21,006 Da. Mass spectroscopy was performed on the purified protein to confirm that no DNA mutations had been accumulated during the PCR plasmid preparation procedure. Isoelectric focusing was used to confirm the presence of the deamidation mutation. His-tagged HyD-Crys had a molecular weight of 21,844 Da.

iv. Fluorescence emission spectra

Fluorescence emission spectra were read on a Hitachi F-4500 fluorimeter with a continuous flow temperature control system. Proteins were diluted to a concentration of 200 mM ammonium acetate, pH 7.0 or 200 mM ammonium acetate containing 6 M GdnHCl. Samples were excited at 295 nm and emission was measured from 310 nm to 420 nm. The excitation and emission slit widths were both set to 10 nm. The background fluorescence of buffer or buffer and 6 M GdnHCl was subtracted out from the sample reading. Fluorescence emission maxima were calculated by averaging signals over every 5 nm and selecting the midpoint of the five signals that exhibited the highest average.

v. Equilibrium refolding and unfolding

For the unfolding process, 100 µg/mL purified wild-type or N143D HyS-Crys was diluted to 10 µg/mL in increasing amounts of GdnHCl in 10 mM NaPO₄, 5 mM DTT, 1 mM EDTA, pH 7.0, from 0 to 5.5 M. The samples were incubated at 37°C until equilibrium was reached (about 6 hours). For the refolding titration, 100 µg/ml protein was denatured in 5.5 M GdnHCl in at 37°C for five hours. The protein was subsequently refolded by dilution to 10 µg/ml into decreasing concentrations of GdnHCl from 5.5 to 0.55 M. The fluorescence spectra of the equilibrated samples were determined using a Hitachi 4500 fluorimeter equipped with a continuous temperature control system with an excitation wavelength at 295 nm and emission monitored from 310 to 420 nm. The excitation and emission slits were both set to 10 nm. The ratios of emission intensities of 360 nm over 320 nm were used for data analysis. Fraction unfolded values were calculated using the method of Pace et al. (1989) and denaturation midpoints were calculated using the Kaliedagraph (Synergy Software) curve fitting function.

vi. Unfolding fluorescence kinetics

Tryptophan environment changes with refolding were monitored using a Hitachi F-4500 fluorimeter equipped with a continuous temperature control system. Native protein at 100 $\mu\text{g}/\text{mL}$ in 10 mM NaPO_4 5 mM DTT, 1 mM EDTA, pH 7.0 at 37°C was unfolded by dilution to final concentrations of 10 $\mu\text{g}/\text{ml}$ HyD-Crys and 5.5 M GdnHCl using a syringe port injection system exhibiting a dead-time of 1 second. Loss of global structure was monitored with continuous excitation at 295 nm at 37°C for 2 hours. Emission intensities during kinetic unfolding were collected at 350 nm for wild-type and N143D HyS-Crys. The fluorescence curves were fit to series of consecutive first-order exponentials using the method described by Fersht (1999). The signals were all fit using the Kaliedagraph (Synergy Software) curve fitting algorithm to mechanisms having one, two, and three exponentials, and the best fit was selected by inspection. All proteins were unfolded in at least two separate experiments to ensure the accuracy of the observed fluorescence and curve fitting.

vii. Refolding fluorescence kinetics

Changes in tryptophan environment during refolding were monitored using a Hitachi F-4500 fluorimeter equipped with a continuous temperature control system. Native protein was denatured at 100 $\mu\text{g}/\text{mL}$ in 5.5 M GdnHCl in 10 mM NaPO_4 , 5 mM DTT, 1 mM EDTA, pH 7.0 at 37°C for two hours. The unfolded protein was refolded by dilution into 10 mM NaPO_4 , 5 mM DTT, 1 mM EDTA, pH 7.0 to final concentrations of 10 $\mu\text{g}/\text{ml}$ HyD-Crys and 0.55 M GdnHCl using a syringe port injection system with a dead-time of 1 second. Increase in global structure during refolding was monitored with continuous excitation at 295 nm and emission at 350 nm for both wild-type and N143D

H γ S-Crys. The fluorescence curves were fit to series of consecutive first-order exponentials using the method described by Fersht (1999). The signals were all fit using the Kaledagraph (Synergy Software) curve fitting algorithm to mechanisms having one, two, and three exponentials, and the best fit was selected by inspection. These experiments were repeated for each protein.

viii. Refolding native gel electrophoresis

Wild-type and N143D H γ S-Crys was unfolded in 10 mM NaPO₄, 5 mM DTT, 1 mM EDTA, 5.5 M GdnHCl overnight at 37°C. The protein was then refolded by dilution in 10 mM NaPO₄, 5 mM DTT, 1 mM EDTA at 37°C to a final protein concentration of 10 μ g/mL for various times. Native gel loading buffer containing DTT was added to each sample and the protein was placed on ice. Samples were separated by electrophoresis through a 9% polyacrylamide gel. The gel was run for three hours at 4°C. Visualization was performed by silver staining as described (Rabilloud et al. 1988).

C. RESULTS

i. Purification and structure assignment of H γ S-Crys

Both wild-type and N143D H γ S-Crys expressed in the soluble portion of the cell lysate and folded into a native-like conformations.

Far-UV CD of the native proteins showed a primarily β -sheet structure for wild-type H γ S-Crys and the N143D mutant protein 37°C, pH 7.0 (data not shown). Both proteins had a characteristic β -sheet minimum at 218 nm indicating that the native states were similar for both proteins.

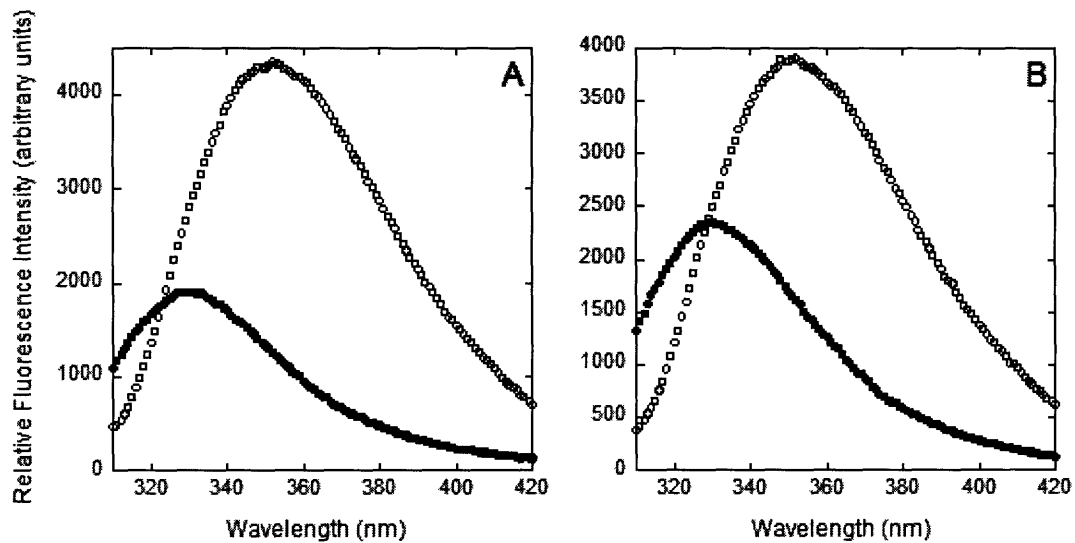


Figure 5-2

Fluorescence emission of native (●) and denatured (○) H γ S-Crys. Protein was excited at 295 nm and emission spectra were collected from samples of 10 μ g/mL protein in 200 mM ammonium acetate, pH 7.0 or buffer and 6 M GdnHCl at 37°C. Fluorescence spectra of wild-type (A) and N143D (B) H γ S-Crys are shown.

The overall character of the fluorescence emission spectra of wild-type and the N143D mutant of H γ S-Crys were assessed by exciting the protein at 295 nm and observing fluorescence emission intensities from 310 to 420 nm in either 200 mM ammonium acetate, pH 7.0 buffer or buffer and made up to 6.0 M GdnHCl (Figure 5-2 and Table 5-1). Wild-type H γ S-Crys had a native fluorescence emission maximum of 329 nm, and was quenched in the native state.

The denatured protein exhibited a fluorescence emission maximum of 350 nm. This is similar to data previously reported (Wenk et. al. 2000). N143D displayed a fluorescence emission maximum of 330 nm in the native state and 350 nm in the denatured state. The N143D mutant displayed similar fluorescence and quenching characteristics as wild-type although the overall fluorescence of the native state appeared to be slightly higher.

ii. Stability of human γ S crystallin

To assess the stability of the proteins, equilibrium unfolding and refolding experiments were performed on both mutant and wild-type H γ S-Crys as a function of GdnHCl concentration (Figure 5-3). Fluorescence spectra were collected at 37°C, pH 7.0 as previously described (Kosinski-Collins and King 2003).

Briefly, for unfolding, protein was added to increasing concentrations of GdnHCl and allowed to incubate until equilibrium was reached. For refolding, the protein was denatured for 5 hours and then refolded by addition to samples of decreasing GdnHCl concentration and incubated until equilibrium was reached.

Both proteins exhibited a single, smooth transition between native and denatured with no evidence of any stabilized intermediates. The refolding and unfolding equilibrium curves were indistinguishable between the two proteins each having the same transition midpoints at 37°C.

These reversible, two-state protein unfolding/refolding transitions were analyzed using the method described by Pace et al. (1997). Both proteins showed a midpoint for denaturant-induced unfolding and refolding at 2.3 M GdnHCl with m-values of 5.5. Wild

type had a ΔG_{H_2O} of 13 kcal/mol while N143D exhibited a ΔG_{H_2O} of 12 kcal/mol. The equilibrium unfolding and refolding results are summarized in Table 5-1.

Table 5-1

Equilibrium unfolding and refolding constants of HyS-Crys

	Native Fluorescence Emission Max	Denatured Fluorescence Emission Max	ΔG_{H_2O}	m-value	Concentration of GdnHCl at $\frac{1}{2}$ denatured	Melting Temp
Wild type	329 nm	350 nm	13 kcal/mol	5.5	2.3 M	76°C
N143D	330 nm	350 nm	12 kcal/mol	5.5	2.3 M	76°C

Wild-type HyS-Crys showed the same characteristics at 25°C as at 37°C. There was no evidence of hysteresis in that the equilibrium unfolding and refolding equilibrium transitions were the same (Figure 5-3C).

To further assess the stability of the proteins, CD temperature melts were performed on wild-type and N143D HyD-Crys (Figure 5-4). CD measured at 220 nm in 10 mM NaPO₄, 5 mM DTT, and 1 mM EDTA, pH 7.0 as a function of temperature. Both proteins displayed a single unfolding transition with a midpoint 76°C. Solution turbidity measurements indicated that both proteins aggregated at this temperature and neither could be refolded by decreasing the temperature of the solution.

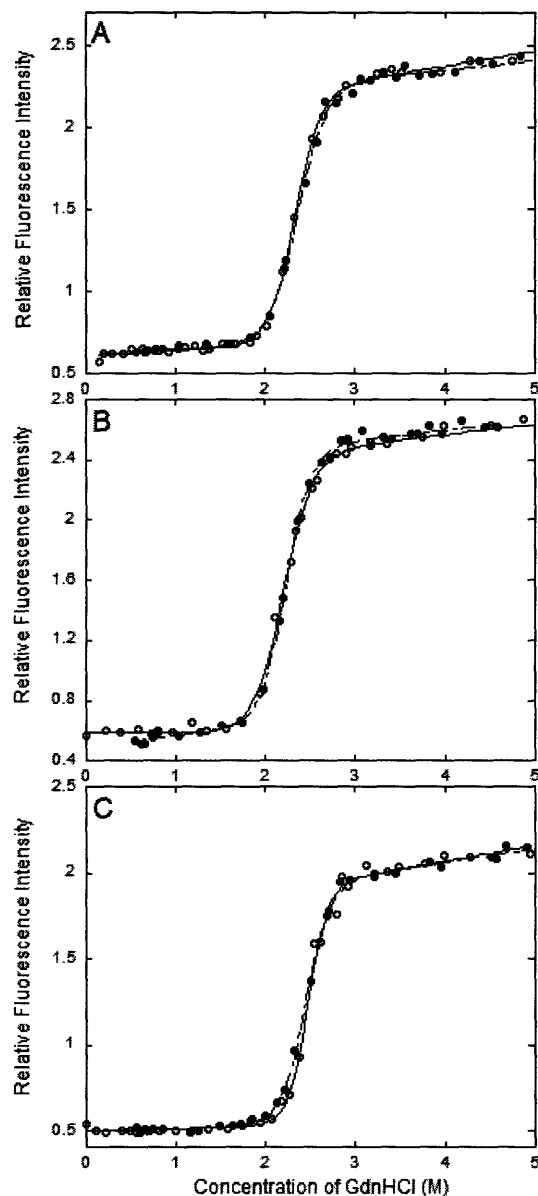


Figure 5-3

Equilibrium unfolding and refolding of H γ S-Crys in GdnHCl. Tryptophan fluorescence was monitored during unfolding and refolding and all samples were equilibrated in 10 mM NaPO₄, 5 mM DTT, 1 mM EDTA, pH 7.0 at a protein concentration of 10 μ g/ml. A representative set of fluorescence data is shown for wild-type H γ S-Crys at 37°C (A), N143D 37°C (B), and wild-type at 25°C (C). All protein data was analyzed using the ratio of fluorescence emission intensities at 360 nm over 320 nm. Unfolding (○) and refolding (●) transitions are presented for each protein.

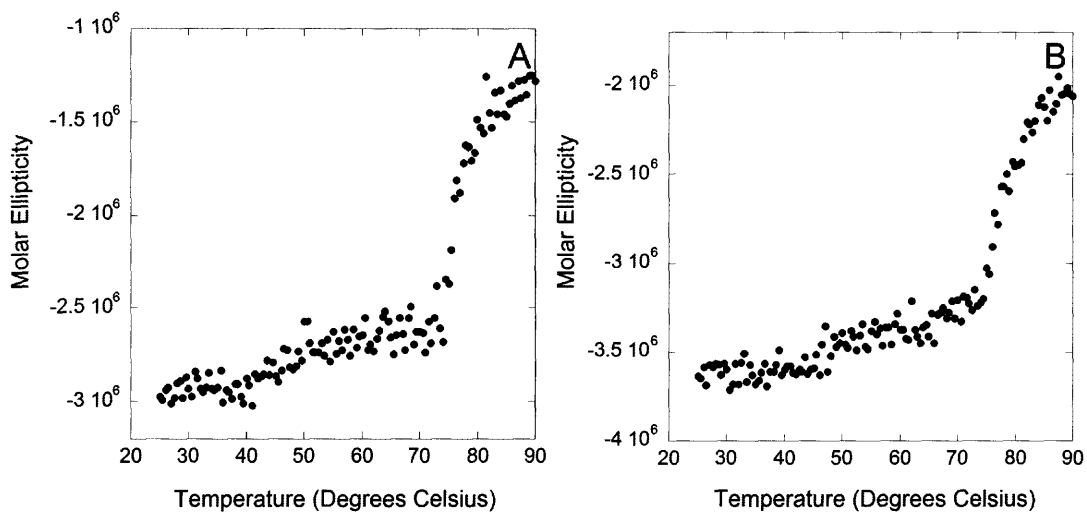


Figure 5-4

Temperature denaturation of HyS-Crys monitored with circular dichroism. Protein was prepared at 300 $\mu\text{g}/\text{mL}$ in 10 mM NaPO_4 , 5 mM DTT, 1 mM EDTA, pH 7.0. The far-UV CD was monitored at 220 nm as the temperature was increased in the cuvette by 0.5°C for both wild-type HyS-Crys (A) and N143D (B).

iii. *Unfolding kinetics*

The unfolding kinetics of H γ S-Crys was studied by dilution of native protein into 5.5 M GdnHCl at 37°C and monitoring the changes in the fluorescence emission (Figure 5-5, Table 5-2). A syringe injection port that exhibited a dead-time of approximately 1 second was used as the mechanism of dilution. Because of the quenching characteristics of H γ S-Crys, fluorescence emission intensity increased upon unfolding for the proteins. Millisecond time-scale intermediate(s) may have formed within the dead-time of these experiments and are not directly addressed here.

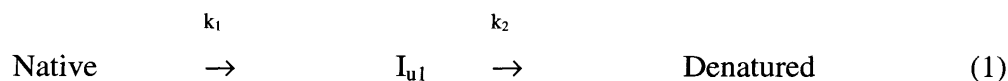
Table 5-2

Unfolding kinetic rate constants of H γ S-Crys

	k_1	$t_{1/2}$	k_2	$t_{1/2}$
Wild type	0.13	5.3 s	0.0037	190 s
N143D	0.16	4.3 s	0.0077	90 s

The kinetic data suggest the presence of partially unfolded intermediates in the transition between the native and unfolded states for both wild type and N143D. The intermediates appear to be hyper-fluorescent in both species.

The wild type unfolding signals were best fit with a three-state unfolding pathway (Fersht 1999). The protein started as a native species and then formed an intermediate before becoming completely unfolded. In this model, an early intermediate (I_u) was populated within a $t_{1/2}$ of 5.3 s. A final unfolding transition ($I_u \rightarrow$ denatured) occurred with a $t_{1/2}$ of 190 s.



The N143D mutant displayed an unfolding fluorescence signal that fit with similar unfolding kinetic rates as wild-type. The unfolding intermediate formed with a $t_{1/2}$ of 4.3 s and the intermediate unfolded with a $t_{1/2}$ of 90 s.

These experiments were repeated in the absence of DTT and similar results were obtained for both the wild-type and mutant HyS-Crys.

iv. Refolding kinetics

In order to study the kinetic rates and intermediates formed during refolding of HyS-Crys, fluorescence emission was monitored during refolding of denatured HyS-Crys at 37°C to a final denaturant concentration of 0.5 M GdnHCl in 10 mM NaPO₄, 5 mM DTT, 1 mM EDTA, pH 7.0 (Figure 5-6, Table 5-3). Solution turbidity scans affirmed that no aggregate was formed under these conditions for either wild-type or mutant HyS-Crys.

Table 5-3

Refolding kinetic rate constants of HyS-Crys

	k_1	$t_{1/2}$	k_2	$t_{1/2}$
Wild type	0.084	8.3 s	0.0034	200 s
N143D	0.12	5.8 s	0.0014	490 s

Figure 6 shows the changes in the raw fluorescence signal during refolding for both wild-type and N143D HyS-Crys. The major transitions occurred within the first 20 minutes of refolding. Intermediates formed within the first second of refolding were not detected in these experiments.

Wild-type HyS-Crys was best fit to a three-state model suggesting the presence of one partially folded intermediate along the productive refolding pathway (Fersht 1999). The refolding reaction exhibited an early transition from denatured to a partially refolded intermediate with a $t_{1/2}$ of 8.3 s and a second transition of intermediate to native with a $t_{1/2}$ of 200 s.

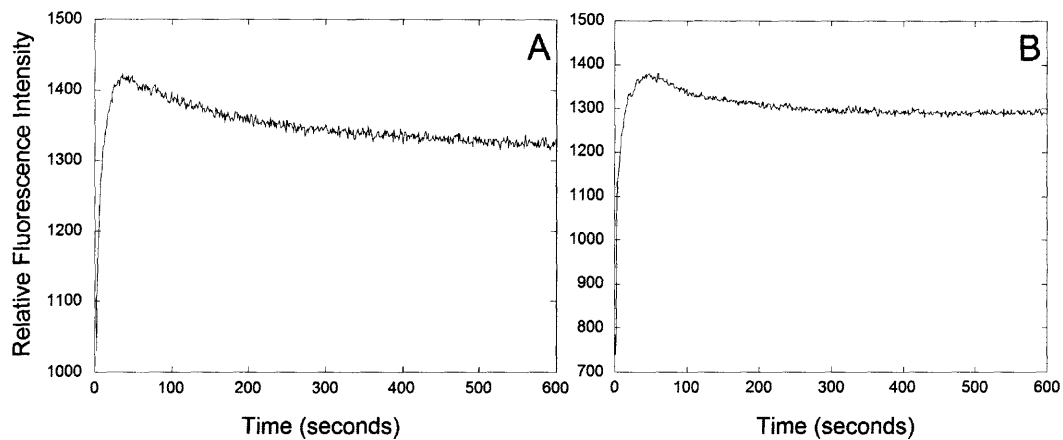


Figure 5-5

Unfolding kinetics of H γ S-Crys monitoring intrinsic tryptophan fluorescence with excitation at 295 nm. Emissions were monitored at 350 nm for wild type and N143D. H γ S-Crys was denatured by rapid dilution into 5.5 M GdnHCl at 37°C in 10 mM NaPO₄, 5 mM DTT, 1 mM EDTA, pH 7.0 to a final protein concentration of 10 μ g/ml. Raw fluorescence signal from a representative protein unfolding time course is shown for wild-type (A) and N143D (B). Data are shown over the first 600 seconds of unfolding, although the data were collected over the entire unfolding process (~1 hour).

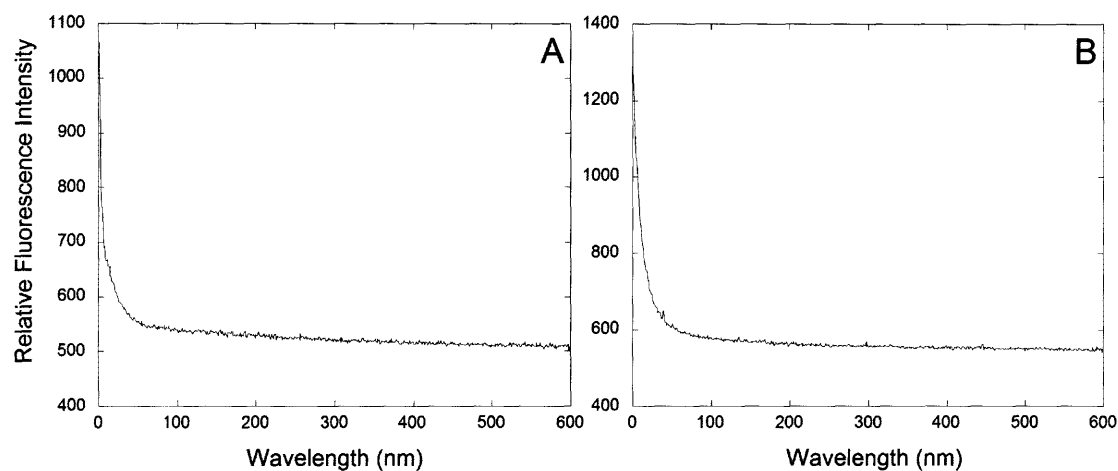


Figure 5-6

Refolding kinetics of H γ S-Crys monitoring intrinsic tryptophan fluorescence with excitation at 295 nm and emission at 350 nm. H γ D-Crys was denatured in 5.5 M GdnHCl at 37°C in 10 mM NaPO₄, 5 mM DTT, 1 mM EDTA, pH 7.0 for 3 hours. H γ S-Crys was refolded by rapid dilution with S buffer to a final GdnHCl concentration of 0.55 M and a final protein concentration of 10 μ g/ml for wild-type (A) and N143D (B) H γ S-Crys. A representative scan of raw fluorescence signals are shown over the first 600 seconds of refolding, although the data were collected over 2 hours.



The fluorescence signal observed during the refolding of N143D could be fit to a similar three state folding pathway as that observed for wild-type. The intermediate formed with a $t_{1/2}$ of 5.8 s and then intermediate to native transition occurred with a $t_{1/2}$ of 490 s.

We repeated these refolding fluorescence experiments in the absence of DTT and obtained similar results for both proteins with no evidence of aggregation.

Wild-type and N143D H γ S-Crys was refolded for various times and the resulting species were resolved on a 9% native polyacrylamide gel.

For wild-type H γ S-Crys, after five minutes of refolding one band was visible in the lanes that corresponded to productively refolded native protein (Figure 5-7). An intermediate formed in during the first 4 minutes of refolding. The intermediate did not resolve into a discreet band, but was visualized as a wide, smeared mass of protein that migrated significantly farther than the native protein. The native gel observation of the formation of an intermediate within the first few seconds of refolding that disappears within 4 minutes is consistent within the intermediate predicted in fluorescence experiments.

For N143D H γ S-Crys, a band corresponding to the migration distance of native N143D was observed after 4 minutes of refolding (Figure 5-8). An intermediate that migrated a farther distance than native formed within the first few seconds of refolding similar to wild type. Traces of this intermediate, however, were visible in lanes corresponding to up to 15 minutes of refolding. How and why this intermediate is stabilized in N143D may provide insight into the process and effect of deamidation in the lens.

IV. DISCUSSION

i. Aggregation and refolding in human γ S crystallin

H γ S-Crys and H γ D-Crys are found in the aggregates extracted from old cataractous lenses (Lapko et al. 2002, Takemoto and Boyle 2000, Takemoto 2001, Hanson et al. 2000). It is not known whether these proteins are co-aggregates or whether their aggregation pathways are independent. An *in vitro* aggregation pathway was identified for H γ D-Crys when the protein was unfolded in GdnHCl and then refolded by dilution to concentrations of GdnHCl lower than 1.0 M (Kosinski-Collins and King 2003). When wild-type and N143D H γ S-Crys were refolded under the same conditions, there was no evidence of aggregation in light scattering or solution turbidity measurements (Figure 5-2).

ii. Deamidation at position 143 in human γ S crystallin

Many posttranslationally modified proteins have been identified from cataractous lenses. Studies have revealed that deamidation at asparagine 143 in H γ S-Crys is a primary covalent modification observed in the insoluble portion of cataractous human lenses (Takemoto 2001). It is thought that deamidation may destabilize the native state of H γ S-Crys by charge addition, or alternately, the oxidative damage may follow initial unfolding or aggregation.

The stability of wild-type and N143D H γ D-Crys were assessed through equilibrium unfolding and refolding experiments. The ΔG_{H_2O} of wild type was 13 kcal/mol and was 12 kcal/mol for N143D. Both proteins exhibited a two-state denaturation transition at 2.3 M GdnHCl and neither protein exhibited any evidence of hysteresis as observed with H γ D-Crys (Kosinski-Collins and King 2003). In addition, temperature melt analysis in the absence of GdnHCl revealed that both proteins had the same denaturation temperature of 76°C. Together, this data demonstrates that N143D was not significantly destabilized as compared to wild-type H γ S-Crys.

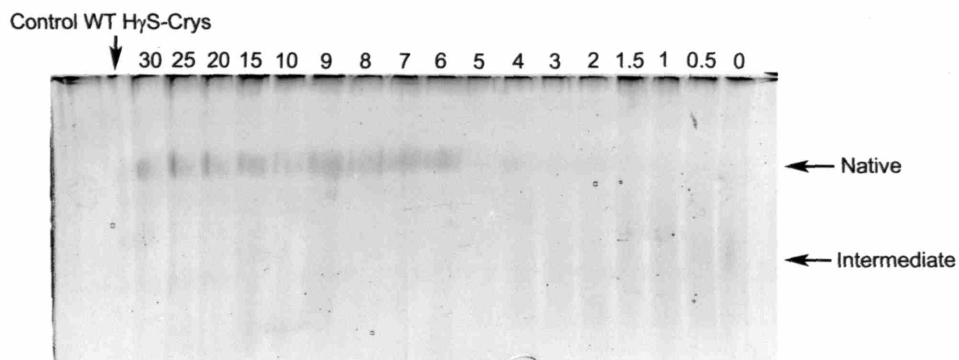


Figure 5-7

Wild-type HyS-Crys was unfolded in 5.5 M GdnHCl overnight and then refolded by dilution into 10 mM NaPO₄, 5 mM DTT, 1 mM EDTA, pH 7.0. From right to left, the gel shows folded wild-type HyS-Crys and samples of 30 min., 25 min., 20 min., 15 min., 10 min., 9 min., 8 min., 7 min., 6 min., 5 min., 4 min., 3 min., 2 min., 1.5 min., 1 min., 0.5 min., and 0 min. of refolding. The location of the native protein and intermediate bands are designated. The gel was run for three hours at 4°C.

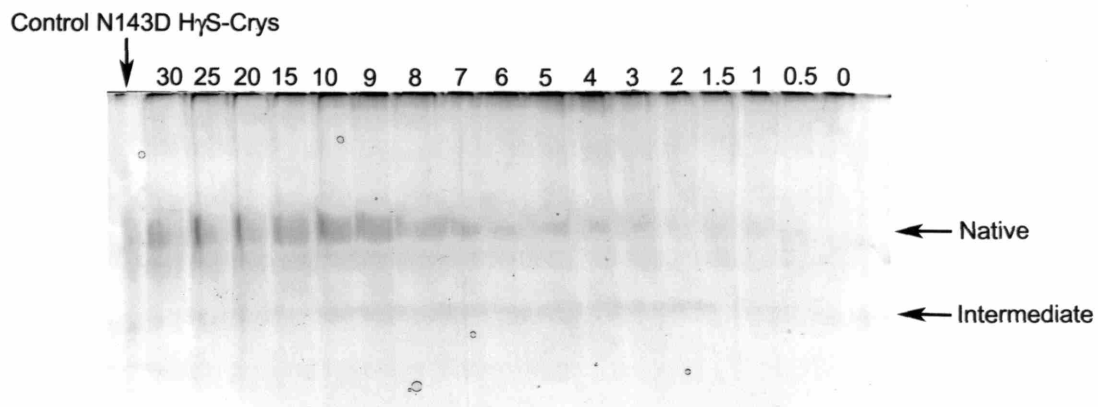
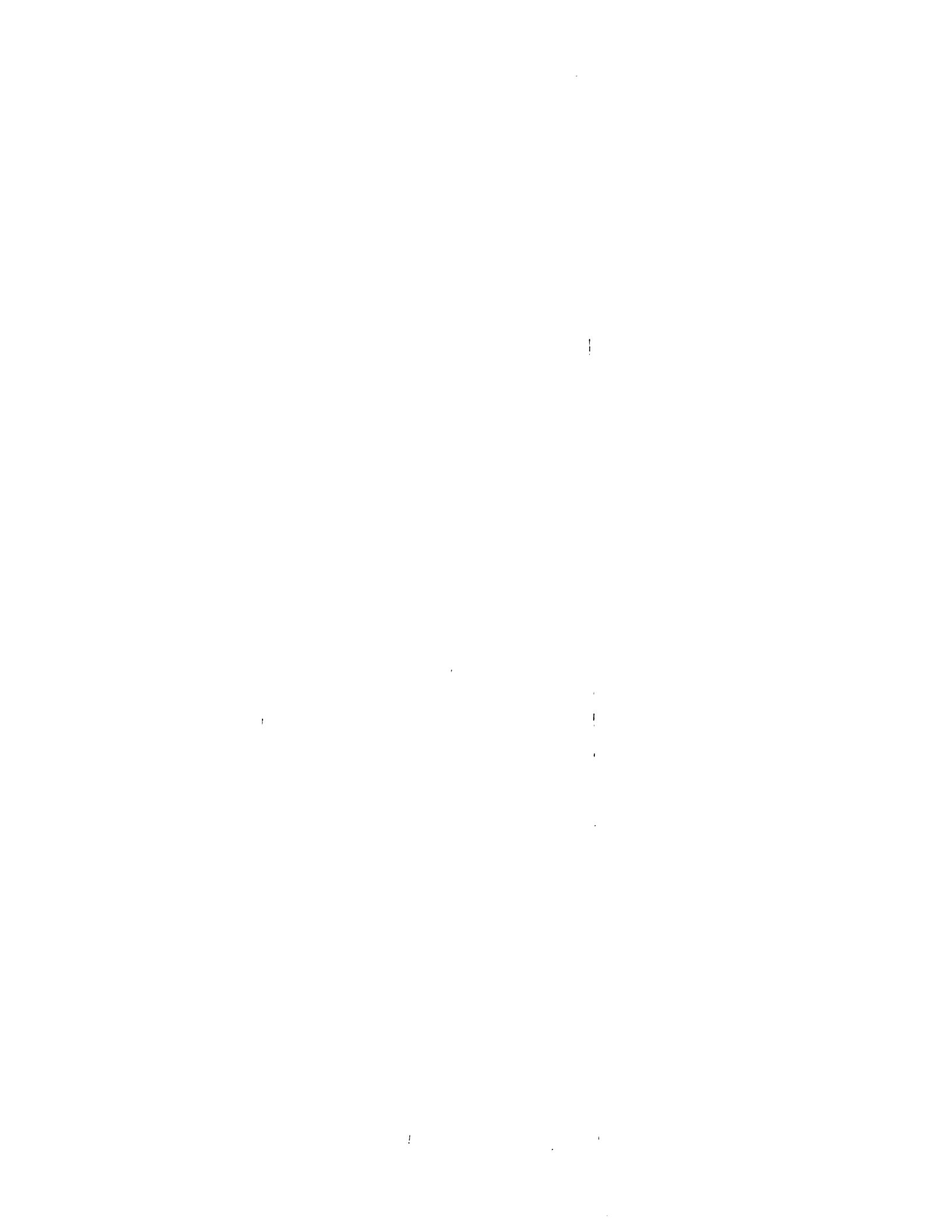


Figure 5-8

N143D HyS-Crys was unfolded in 5.5 M GdnHCl overnight and then refolded by dilution into 10 mM NaPO₄, 5 mM DTT, 1 mM EDTA, pH 7.0. From right to left, the gel shows folded N143D HyS-Crys and samples of 30 min., 25 min., 20 min., 15 min., 10 min., 9 min., 8 min., 7 min., 6 min., 5 min., 4 min., 3 min., 2 min., 1.5 min., 1 min., 0.5 min., and 0 min. of refolding. The location of the native protein and intermediate bands are designated. The gel was run for three hours at 4°C.



We further studied the *in vitro* unfolding and refolding pathway of wild-type and N143D H γ D-Crys. During unfolding, fluorescence changes revealed that both proteins went through a partially denatured, hyper-fluorescent intermediate (Table 5-2). Based on the time of formation of the intermediate for wild-type and N143D H γ S-Crys, the partially unfolded species more than likely had the same overall structure for both proteins. The fluorescence signals observed during the refolding of wild-type and N143D H γ D-Crys are similar as well. Both proteins refolded through a partially unfolded conformation and had similar kinetic rate constants (Table 5-3). This is more than likely the intermediate proposed to exist in wild type during isolated domain studies performed by Wenk et al. (2000). It appeared as though this intermediate was longer lived in N143D than in wild-type H γ S-Crys in native gels, although this may simply be an artifact of the ice incubation of the protein before resolution on the gel. We cannot precisely determine whether or not the unfolding and refolding intermediates were the same for either protein. Overall, these data suggest that an aspartic acid at position 143 did not affect the overall characteristics of protein folding *in vitro*.

Our results reveal that the unfolding and refolding pathways of wild-type and N143D H γ S-Crys were very similar and N143D was not destabilized as compared to wild type. This indicated that the final deamidated form of aspartic acid at position 143 was not important for H γ S-Crys folding and stability. This, however, does not reveal the role of deamidation to H γ S-Crys destabilization and unfolding in general. Previous studies have suggested that deamidation in proteins may occur via a succinimide intermediate (Fujii et al. 1999; Aswad et al. 2000). Specifically, studies by Takemoto et al. have shown that this is likely in the N143D mutation of H γ S-Crys (2001). Insoluble H γ S-Crys removed from cataractous lenses contained both L-beta-aspartate and L-alpha-aspartate indicating that it formed through an L-Succinimidyl intermediate. It is possible that it is the succinimidyl intermediate of N143D that alters H γ S-Crys folding and stability *in vivo* and future studies should examine the affect of the cyclic intermediate on the protein.

iii. Role of chaperonin recognition

The discussion above assumes that the impact of deamidation is to perturb the initial conformation of H γ S-Crys creating a disrupted structure that is more prone to proceed down an aggregation pathway. As noted earlier, the recovery of deamidated proteins from the cataract leaves open the question of whether the events are primary or secondary.

An alternate possibility is that deamidation may interfere with recognition of H γ S-Crys by the α -crystallins or a deamidated residue in H γ S-Crys may serve as the α -crystallin recognition site directly. In fact, deamidated amino acids may be among the signals that α -crystallin uses to recognize non-native crystallin species. In such models some initial stress would result in partial unfolding, sensitizing the protein to oxidative damage.

Unfortunately it is not known whether the oxidized $\beta\gamma$ -crystallins found in cataracts are complexes with α -crystallin or with each other. Inclusion in the cataract could reflect either efficient recognition, or failure of recognition of α -crystallin. To sort this out, we should determine the partners of oxidized crystallins within the cataract, and study the interaction of H γ S-Crys and α -crystallin *in vitro*.

CHAPTER VI: RECRUITMENT OF HUMAN γ S CRYSTALLIN TO THE HUMAN γ D CRYSTALLIN AGGREGATION PATHWAY

A. INTRODUCTION

H γ S-Crys has 69% sequence similarity and 50% sequence identity with H γ D-Crys. H γ S-Crys has a three amino acid N-terminal extension of Ser-Lys-Thr and a two amino acid domain linker extension of His-Leu not present in H γ D-Crys. H γ S-Crys was historically thought to be present only in the lens fiber cells; however, recent evidence from Beebe and colleagues suggests that it may be in the lens epithelium as well.

H γ D-Crys is expressed primarily early in life and predominately in the lens nucleus. H γ D-Crys could be refolded to its native state at 37°C after dilution out of denaturant (Kosinski-Collins and King 2003). An *in vitro* aggregation pathway of H γ D-Crys that competed with productive refolding was identified and may be related to the mechanism of its involvement in mature onset cataracts (Kosinski-Collins and King 2003). It is important to note, however, that native, soluble, H γ D-Crys is thought to be more nuclear in the lens and H γ S-Crys is thought to be more cortical. It is unlikely, therefore, that disrupted H γ D-Crys localized to the lens nucleus would be able to recruit native, unsecreted H γ S-Crys. But, because both H γ D-Crys and H γ S-Crys are found in cataract, we have investigated whether or not aggregation of H γ D-Crys can recruit or seed H γ S-Crys aggregation *in vitro*.

B. MATERIALS AND METHODS

H γ D-Crys and H γ S-Crys were unfolded in 5.5 M GdnHCl at 100 ug/mL overnight at 37°C. Equal amount of denatured or native H γ S-Crys and denatured H γ D-Crys were mixed and then refolded by a ten-fold dilution into 10 mM NaPO₄, 5 mM DTT, 1 mM EDTA, pH 7.0 to a final denaturant concentration of 0.55 M. For the time-dependent aggregation recruitment experiments, samples of denatured H γ S-Crys were begun

refolding and then equivalent amounts of denatured HyD-Crys were added to the samples after 0 minutes, 2 minutes, 10 minutes, 30 minutes, 1 hour, and 2 hours of HyS-Crys refolding. The protein samples were allowed to refold together at 37°C for 3 hours and then the solutions were centrifuged at 40,000 RPM for 1 hour. The supernatant was decanted and the pellet was resolubilized in sodium dodecyl sulfate (SDS) containing β -mercaptoethanol. The SDS containing protein samples were then boiled at 95°C for 30 minutes and resolved on a 14% SDS polyacrylamide gel. The protein bands were then visualized using silver nitrate staining as described (Rabilloud et al. 1988).

In the case of the time-dependent aggregation recruitment experiment, the overall amount of protein found in each lane was quantitated by measuring the intensity of the bands. The calculated values were then corrected for the background lane intensity and then normalized to the most intense band which was the HyD-Crys band observed from HyS-Crys and HyD-Crys mixing at 0 minutes for each separate experiment. Final values were calculated by averaging data collected from two separate experiments.

C. RESULTS

Though HyS-Crys did not aggregate on its own under the aforementioned refolding conditions, we were interested in its behavior in the presence of aggregating HyD-Crys.

Wild-type His-tagged HyD-Crys was denatured in 5.5 M GdnHCl and then refolded to 0.55 M GdnHCl. Previous studies have shown that HyD-Crys aggregates under these conditions (Kosinski-Collins and King, 2003). The resulting solution was centrifuged at high speed and the aggregate was resolubilized in SDS and boiled. Lane 2 of figure 6-1A and 6-1B show the resolution of this species on a polyacrylamide gel. There was only one band in each of these lanes corresponding to approximately 20,000 Da. There was no evidence of any high molecular weight species that were resistant to SDS denaturation.

Wild-type His-tagged HyD-Crys and wild-type HyS-Crys were denatured separately in 5.5 M GdnHCl at 37°C. Equivalent amounts of both protein solutions were added together and refolded by rapid dilution into phosphate buffer at pH 7.0. When the

resulting aggregate was resolubilized in SDS and resolved on a 14% SDS polyacrylamide gel, it contained two major bands that migrated to slightly different distances (Figure 6-1A, Lane 1). These bands corresponded to the migration distance of H γ S-Crys monomers and H γ D-Crys monomers of approximately 20,000 Da. SDS denatured His-tagged H γ D-Crys migrates further on a polyacrylamide gel than H γ S-Crys despite having a larger molecular weight. In mass spectroscopy analysis of His-tagged H γ D-Crys, we have found that the His-tag degrades. It is possible that during SDS boiling and denaturation, the His-tag of H γ D-Crys is more susceptible to degradation thus decreasing the predicted molecular weight of the protein and altering the gel migration distance. In addition to the two major crystallin monomer bands, there were a large number of higher molecular weight species resistant to SDS denaturation. The most visible bands corresponded to species of a molecular weight between approximately 40,000 Da and 45,000 Da. This more than likely represents dimer species of various combinations of H γ S-Crys and H γ D-Crys. There were several bands of lower molecular weight than 20,000 Da that probably represent degradation products of H γ S-Crys and H γ D-Crys. The N143D mutant of H γ S-Crys co-aggregated with H γ D-Crys to approximately the same degree as wild-type H γ S-Crys and showed similar high molecular weight species (data not shown). This data seems to indicate that the dimer is more resistant to SDS denaturation than the aggregate.

When native H γ S-Crys was added to refolding H γ D-Crys, however, only one band was visible on the SDS polyacrylamide gel corresponding to the migration distance of H γ D-Crys (Figure 6-1B, Lane 1). This indicated that native H γ S-Crys could not be recruited to the H γ D-Crys *in vitro* aggregate under these conditions.

The recruitment of wild-type and N143D H γ S-Crys to the H γ D-Crys aggregate was further studied in a time dependent manner. Denatured samples of H γ S-Crys were begun refolding and an equivalent amount of denatured H γ D-Crys was added after 0, 2, 10, 30, 60, and 120 minutes. The samples were allowed to completely refold together overnight and were then centrifuged at high speed. The resulting pellet was resolubilized in SDS, boiled, and separated on a 14% acrylamide gel (Figure 6-2).

Both H γ S-Crys and H γ D-Crys were observed in all pellets although the amount of protein incorporated into the aggregate decreased as H γ S-Crys was given additional time

to refold. Even though the overall concentrations of H γ D-Crys and H γ S-Crys were the same, the amount of H γ D-Crys found in the pellet was always more than H γ S-Crys. The intensity of the wild-type H γ S-Crys resuspended pellet band was approximately 40-70% of that of H γ D-Crys. The intensity of the N143D H γ S-Crys band was about 20-45% of the H γ D-Crys band. This data suggests that wild-type H γ S-Crys is more prone to be recruited to the H γ D-Crys aggregate than N143D.

The intensity of the bands of H γ S-Crys and H γ D-Crys were quantitated and the amount of protein relative to that observed when H γ S-Crys and H γ D-Crys were refolded together from 0 minutes was determined (Figure 6-3). The intensity of the monomer bands observed for wild-type and N143D H γ S-Crys aggregation with H γ D-Crys decreased proportionally as a function of time. In the case of both wild-type and N143D H γ S-Crys, at least some of the protein refolded for two hours could still be pulled into the H γ D-Crys aggregate as demonstrated by the presence of a H γ S-Crys monomer band in these lanes. This indicated that perhaps some non-native protein was present in these samples and was susceptible to aggregation.

In all samples of the time-dependent aggregation experiment, SDS-resistant, high molecular weight aggregate formed. Additionally, there was one aggregate species that was found in wild-type H γ S-Crys samples that was not present in N143D samples. This band is indicated on Figure 6-2.

D. DISCUSSION

H γ S-Crys did not aggregate on its own when refolded by dilution out of denaturant *in vitro*. Solution turbidity measurements of refolded H γ S-Crys in both kinetic and equilibrium experiments showed no evidence of polymerized protein species. Another member of the γ -crystallin protein family, H γ D-Crys, aggregates when refolded under these conditions (Kosinski-Collins and King, 2003).

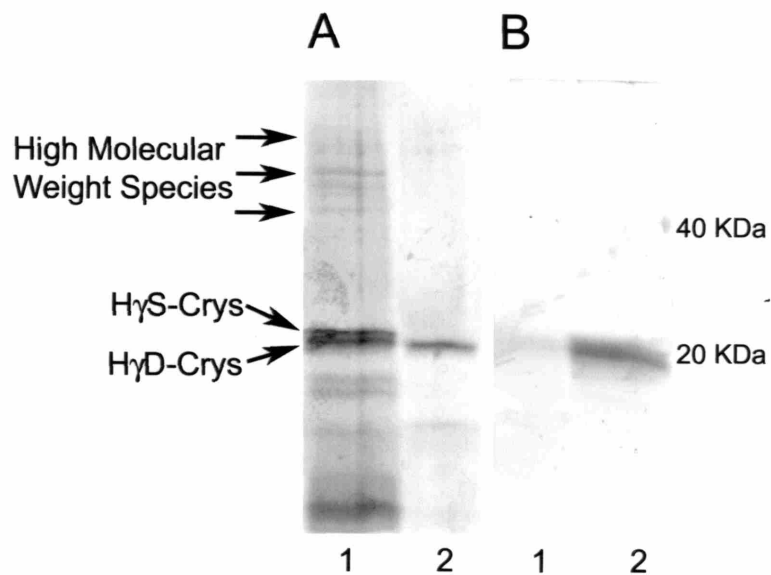
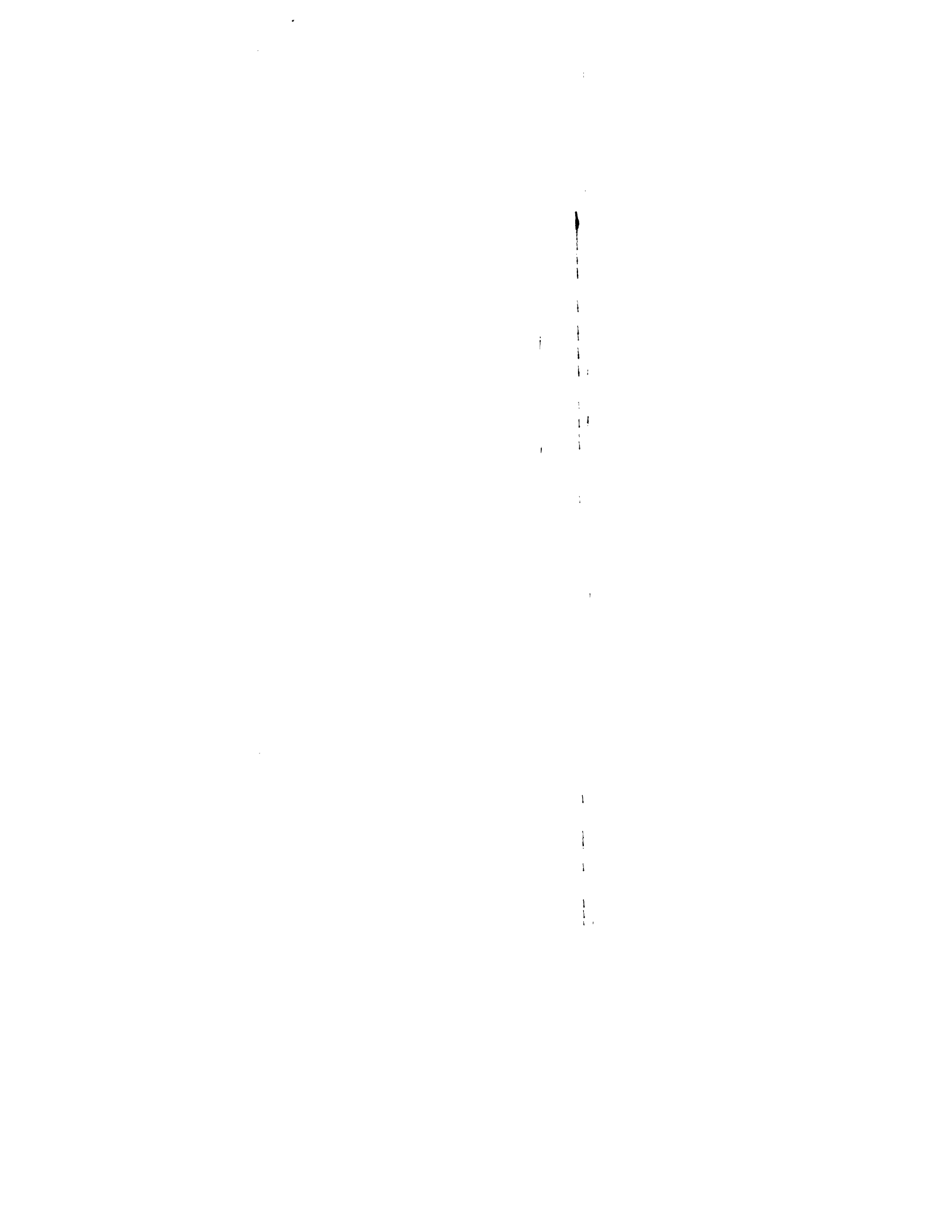


Figure 6-1

Recruitment of H γ S-Crys to the H γ D-Crys aggregate. Denatured (A) or native (B) H γ S-Crys protein was added to aggregating H γ D-Crys. The solution was allowed to refold for 5 hours and then the resulting aggregate was resolubilized by boiling in SDS. The resolubilized pellet was resolved on a 14% polyacrylamide gel. Lane 1 is SDS denatured H γ D-Crys aggregate alone and lane 2 is H γ D-Crys aggregating with H γ S-Crys.



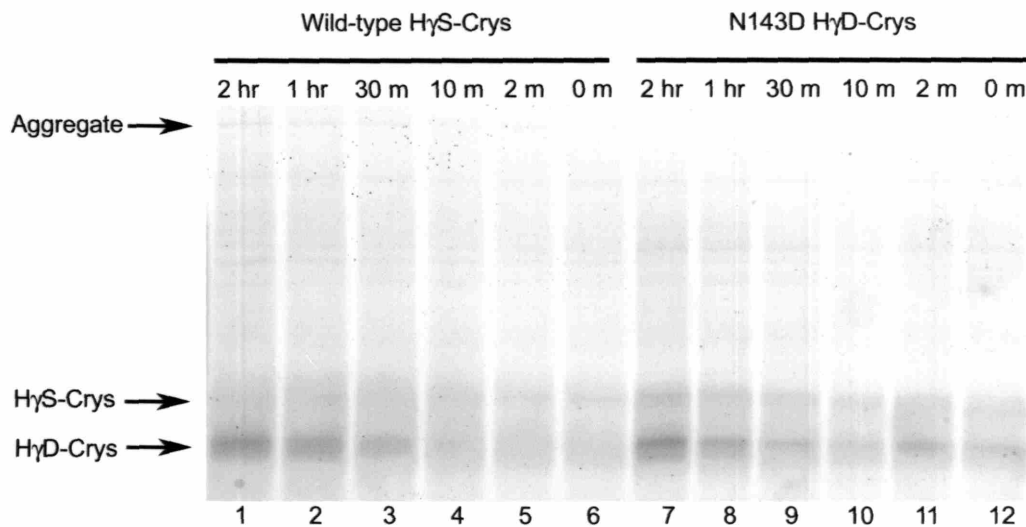


Figure 6-2

Time dependent recruitment of wild-type and N143D H γ S-Crys to the H γ D-Crys aggregate. Refolding of denatured H γ S-Crys was initiated and an equal concentration of denatured H γ D-Crys was added after 0, 2, 10, 30, 60, or 120 minutes. The protein samples were allowed to refold together overnight and then the solutions were centrifuged at 40,000 RPM. The resolubilized pellet was resolved on a 14% polyacrylamide gel. Lanes 1 through 6 show the time dependent refolding and aggregation of wild-type H γ S-Crys and H γ D-Crys, while lanes 7 to 12 show N143D H γ S-Crys and H γ D-Crys. The time of addition of denatured H γ D-Crys after H γ S-Crys refolding was begun is indicated at the top of the lanes.

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

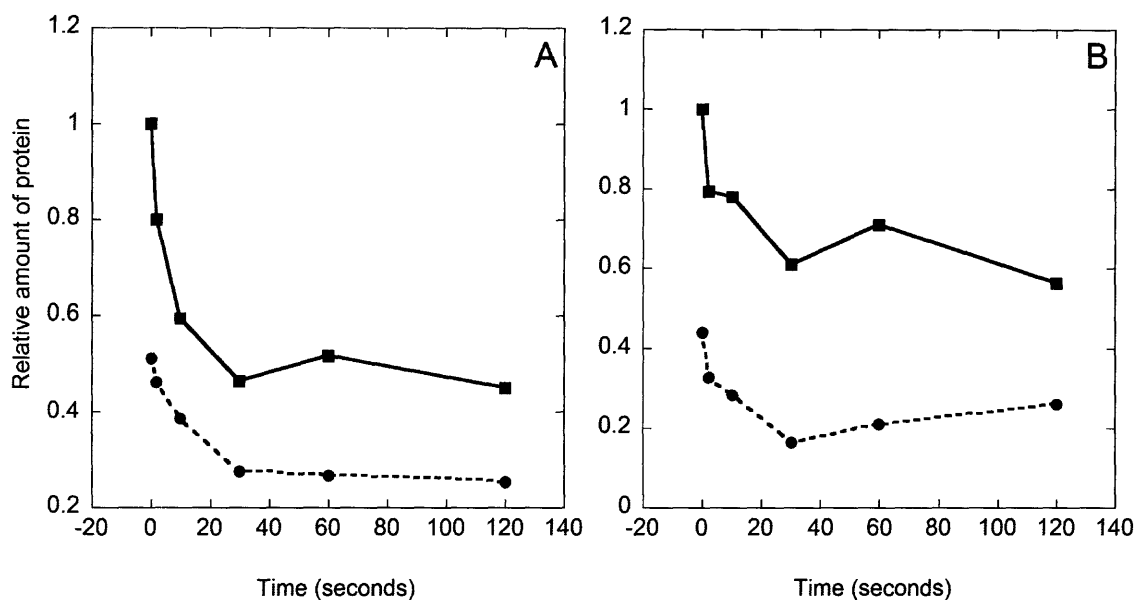


Figure 6-3

Quantitation of loss of band intensities for time dependent recruitment of H γ S-Crys to the H γ D-Crys aggregate. The overall intensity of the monomer bands of both H γ D-Crys and H γ S-Crys were measured and averaged from two separate refolding/aggregation experiments. The percentages of protein found in the bands were calculated relative to the most intense band observed in the experiment which was always the H γ D-Crys monomer band visualized at 0 minutes. Results from wild-type (A) and N143D (B) H γ S-Crys aggregation recruitment are shown. The relative amounts of the H γ S-Crys (●, dotted line) and H γ D-Crys (■, solid line) monomers are depicted.

Although denatured H γ S-Crys did not aggregate alone *in vitro*, it could be recruited to aggregating H γ D-Crys as shown by the double band present in SDS gel of the solubilized pellet. Native H γ S-Crys, however, could not be recruited to the H γ D-Crys aggregate. The coaggregation phenomenon may have implications for *in vivo* polymerization and cataract formation.

At least some H γ S-Crys could be recruited to the H γ D-Crys aggregate after it had been refolded for up to two hours (Figure 6-2). This indicated that at least some non-native species was present in these samples. There may have been a minutely populated slow folding species of H γ S-Crys. Kinetic refolding experiments showed that all tryptophans had reached a native-like environment by thirty minutes. This suggested that the non-native species of H γ S-Crys that could be recruited to the H γ D-Crys aggregate at two hours had proper hydrophobic core packing but had incorrect domain interaction or β -strand alignment.

When H γ S-Crys and H γ D-Crys were refolded together, high molecular weight species of approximately 40,000 to 45,000 Da were formed that could not be denatured with SDS. These species were most likely dimers of various combinations of H γ S-Crys and H γ D-Crys, because refolding H γ D-Crys did not form any SDS resistant multimers. These dimers may represent a highly stable species capable of nucleating further aggregation *in vitro*. Further studies should address the mechanism of coaggregation of H γ S-Crys and H γ D-Crys and determine the overall composition of the polymerized state.

When aggregated proteins are removed from old, cataractous lenses, they contain many different crystallin proteins including H γ S-Crys and H γ D-Crys. When H γ D-Crys becomes destabilized in the lens and begins to aggregate, perhaps other destabilized proteins that cannot necessarily polymerize on their own such as H γ S-Crys can be recruited into the oligomerizing H γ D-Crys species or vice versa. As the age of the protein increases, this results in the accumulation of covalent modifications such as deamidation and the formation of a succinimide intermediate of H γ S-Crys. These covalent changes may provide just enough tertiary structure destabilization of H γ S-Crys to allow it to be pulled into the H γ D-Crys aggregate. High molecular weight species removed from the lens should be studied to see if they can seed *in vitro* crystallin aggregation.

CHAPTER VII: CONCLUDING DISCUSSION

A. AROMATIC-AROMATIC RING INTERACTIONS IN HUMAN γ D CRYSTALLIN

The nuclear crystallins present in the human eye lens are expressed early in life and the proteins are present at up to 700 mg/mL. There is no protein turnover in mature lens cells and, therefore, the crystallins are exposed to electromagnetic radiation and other oxidizing stresses throughout the human lifetime. For all of these reasons, the crystallin proteins must be extremely stable to stay in solution for proper function. Below we consider the possibility of a network of aromatic interactions contributing to crystallin stability.

Human γ D crystallin (HyD-Crys) is a protein containing an extensive network of aromatic ring interactions including four tryptophans, fourteen tyrosines and six phenylalanines (Figure 7-1). Raman spectroscopy has revealed that the tyrosine residues from all bovine γ -crystallins are marginally hydrogen bonded (Pande et al. 1991). The interactions between the aromatic residues of HyD-Crys may function as a stabilization “network.” Each individual ring-ring contact alone probably does not contribute very much to the stability of the molecule given the tolerance of the molecule to tryptophan substitution. However, the sum contribution of all of the aromatic stacks may be important. This is different from the stabilization observed in coiled-coils where isolated amino acid contacts contribute significantly to the stability of the complex (Lupas 1996).

Our studies have shown that the four tryptophans at the center of the hydrophobic core of this protein are not vital for stability and can be triply substituted by phenylalanine or singly replaced with alanine with no significant destabilization in terms of protein solubility and/or Gibb’s free energy; however, the quenching of Trp 68 and Trp 156 implies particularly intimate noncovalent native-state interactions. Perhaps it is not any individual residue that controls or dictates the stability of the molecule, but perhaps it is the overall impact of the pi bond interactions caused by the entire aromatic-aromatic ring stacking network that allow HyD-Crys to remain stable.

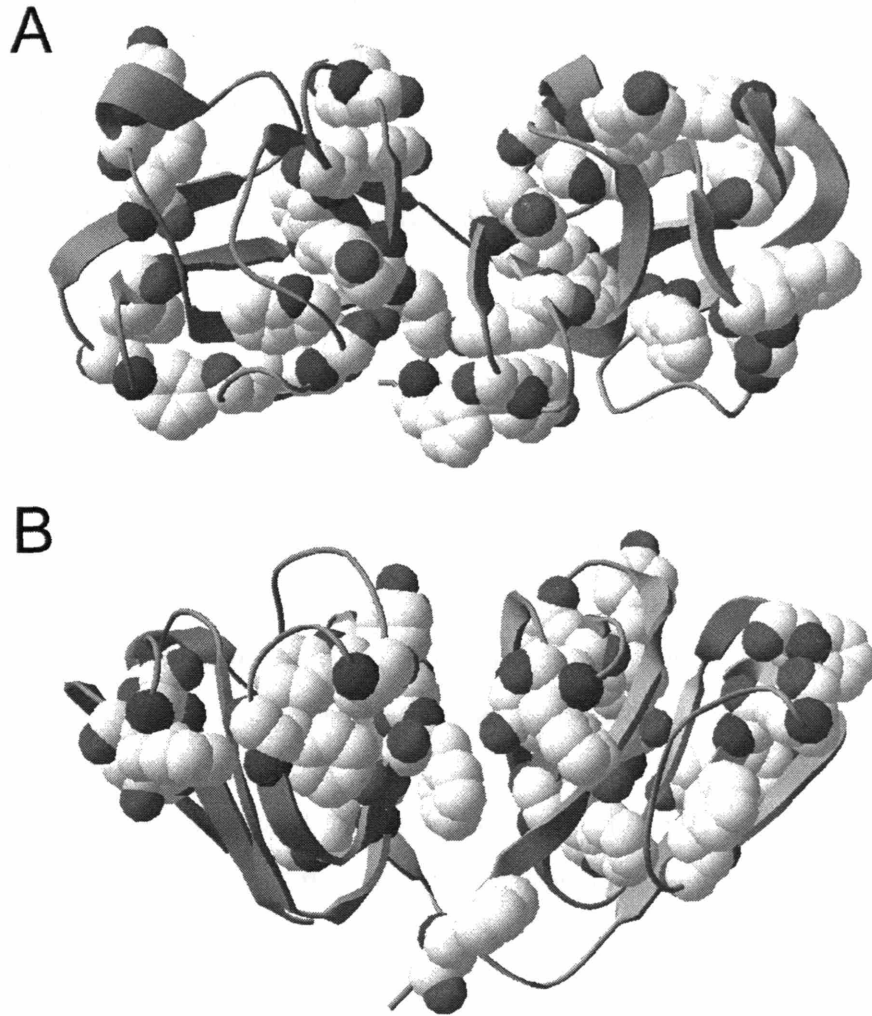


Figure 7-1

Ribbon structure of human γ D crystallin showing the location of aromatic network (Basak et al. 2003). The structure is shown from the top (A) and the side (B).

Many proteins are stabilized by aromatic-aromatic ring interactions including small globular proteins like barnase and larger fibrous proteins such as collagen. One example is the P22 tailspike protein which is a parallel β -helix protein. This protein has a highly aromatic stack running up the center of the helix in which the rings sit one on top of the other and the pi bond orbitals are aligned slightly divergent from parallel (Betts et al. 2004).

In 1985, a study was published by Burley and Petsko that analyzed the frequency and orientation of aromatic ring interactions in proteins. They found that aromatic rings are most often found at an interaction distance of 4.5 to 7.0 Å with 90° angles separating the pi bond orbitals. Further analysis of paired aromatic residues argued that aromatic rings preferred to align themselves in an “off-centered parallel” arrangement whereby the ring of one residue is perfectly stacked flat upon the ring of the adjacent but not centered completely (McGaughey et al. 1998). In addition, studies using double mutant cycle analysis have shown that the contribution of an aromatic-aromatic pair to protein stability is -0.6 to -1.3 kcal/mol (Serrano et al. 1991).

There is an extensive global aromatic network in HyD-Crys, and studies using triple tryptophan mutants showed that the fluorescence quenching observed in the molecule’s native state is most likely due to two smaller local aromatic networks. One network surrounds Trp68 and involves Tyr55, Tyr62, and His65, while the other surrounds Trp 156 and involves Tyr143, Tyr150, and His122 (Basak et al. 2003). It seems likely that HyD-Crys has multiple levels of aromatic ring stabilization involving local and global protein sequence and structure. Although we have not determined the structure of aggregated HyD-Crys yet, it is likely that there is a significant reorganization of the aromatic network in the self-associated species. Raman spectroscopy of the aggregated protein may provide insight into the orientation of the aromatic rings in this species.

Inspection of HyD-Crys reveals that the majority of the aromatic residues are located near the “top” of the molecule clustering around Trp68 and Trp156. With exception of the tryptophans, many of the aromatic amino acids are found near the outer edges of domain I and domain II and are at least partially solvent exposed.

The tyrosine and phenylalanine residues of HyD-Crys seem to create a “ring” around the outer edge of both domains. Because these residues are solvent exposed, it is possible that the aromatic rings from one protein may interact with the tyrosine rings from an adjacent protein. Particularly striking is the fact that the solvent exposed rings tend to be found in pairs.

Close to the domain interface of the protein, Tyr45 interacts with Tyr50 at a 5.0 Å distance, and Tyr134 interacts with Tyr139 at a 4.9 Å distance. Both of these pairs have a side to face ring arrangement and may be important during interface formation or domain reorganization in the folding pathway.

Of further interest in HyD-Crys are the interactions between Tyr16 and Tyr28 located 4.5 Å apart and Tyr93 and Tyr98 found 3.8 Å apart. In both of these cases, the two interacting tyrosine rings are virtually parallel to each other. The pairs are located at opposite ends of the protein. Similarly, Tyr6 and Phe11 interact at a distance of 4 Å, and Phe116 and Phe118 interact at a 3.5 Å distance. These two pairs show a more typical face to end ring configuration, but are also located at opposite ends of domain I and domain II. Together, these four pairs create a symmetrical rectangle around the outside edge of HyD-Crys and may provide an alignment platform for protein self association *in vivo*. Aromatic ring interactions between multiple crystallin proteins may be important for maintaining crystallin solubility and stability at the high protein concentrations found in the eye lens.

Future studies should analyze that angle at which the rings interact in the x-ray structure of HyD-Crys and assess the contribution of tyrosine and phenylalanine to protein stability. It is possible that insight into ring stabilization and stacking in crystallin may be applicable to other highly soluble proteins as well.

B. CRYSTALLIN FOLDING, AGGREGATION, AND HYSTERESIS

Cataract is caused by polymerization of the crystallin proteins. Given the high concentration and stability of native crystallin in the lens, the aggregation-prone state of the protein is most likely a partially unfolded or disrupted conformation. We have investigated the thermodynamics and kinetics of HyD-Crys and human γ S crystallin

(H γ S-Crys) unfolding, refolding, and aggregation *in vitro* to identify the partially folded conformation and characterize the resulting aggregation process.

H γ D-Crys exhibited an unfolding and refolding hysteresis at 37°C that became more pronounced at 25°C (Figure 2-3) (Kosinski-Collins and King 2003). H γ S-Crys did not have an equilibrium unfolding and refolding hysteresis at 37°C or 25°C (Figure 5-3). Recent studies by Flaugh and King have shown that the unfolding and refolding transitions of H γ D-Crys overlay exactly after a 24 hour equilibration eliminating the hysteresis (personal communication). In all of these studies, it was the position of the unfolding transition that changed as a function of temperature and time. This indicated that the unfolding pathway of H γ D-Crys had a kinetically controlled barrier while the unfolding pathway of human γ S crystallin (H γ S-Crys) did not. In addition, aggregation of H γ D-Crys was observed when the protein was denatured in GdnHCl and then refolded into buffer, while H γ S-Crys showed no aggregation propensity. It is possible that the observed hysteresis and aggregation are related to direct differences in sequence between H γ D-Crys and H γ S-Crys. Differences between the H γ S-Crys and the H γ D-Crys primary sequences must control the *in vitro* aggregation process. H γ S-Crys has 69% sequence similarity and 50% sequence identity with H γ D-Crys. H γ S-Crys has a three amino acid N-terminal extension of Ser-Lys-Thr and a two amino acid domain linker extension of His-Leu not present in H γ D-Crys.

Studies performed in our lab have shown that the interface residues of H γ D-Crys are important in unfolding and refolding of H γ D-Crys (Flaugh et al. 2004, manuscript submitted and S. L. Flaugh and J. A. King, manuscript in preparation). When the domain interface residues of H γ D-Crys were mutated, a partially unfolded folding intermediate was stabilized. Jaenicke has suggested that the domain interactions are weaker in H γ S-Crys than in H γ D-Crys, and it seems likely that the interface may be important in the different aggregation character between the two proteins (1999).

The domain interface of H γ D-Crys contains two interacting polar pairs of residues around the two edges of the interface surrounding three pairs of hydrophobic interactions near the center. The polar pairs of H γ D-Crys are Glu54/Gln143 and Arg79/Met147 while the hydrophobic pairs are Met43/Val132, Phe56/Leu145, and Iso81/Val170. Based on alignments of H γ S-Crys with H γ D-Crys, the hydrophilic pairs of H γ S-Crys are

Met57/Gln148 and Arg82/Asp152 while the hydrophobic pairs are Ala46/Iso137, Iso59/Leu150, and Val84/Iso175. The only notable differences between the two interfaces are the loss of the aromatic (Phe56) in HyS-Crys and the placement of the salt bridge at opposite sides of the domain interface. The changes in interface residues between HyD-Crys and HyS-Cry appear minor and most likely do not account for the differences in aggregation propensity or hysteresis formation observed in the two molecules.

The unfolding and refolding equilibrium hysteresis may be a clue to the mechanism of aggregation of HyD-Crys. Studies using triple tryptophan to phenylalanine mutants of HyD-Crys showed that the origin of the hysteresis in HyD-Crys came from residues surrounding Trp156 (Figure 4-1). The only triple tryptophan mutant displaying an equilibrium unfolding and refolding hysteresis was Trp156-only. Given that a kinetic barrier exists to productive unfolding, there is probably a rigid region in this area of the molecule that is resistant to denaturation (Figure 2-3). Perhaps this area of the molecule forms a structured core from which the remainder of the molecule folds, and must remain stable even in extreme conditions to retain the native structure of the overall protein. Alternatively, perhaps as this region of the molecule unfolds, a non-native core of interaction is stabilized that is not necessarily a productive on-pathway conformation. As observed in interleukin-1 β , this region may represent an aggregation-prone “micro-domain” that is resistant to unfolding and that forms rapidly when exposed to denaturant-free buffer upon refolding. We cannot say decisively whether the “micro-domain” is made up of native or non-native inter-residue contacts. Formation of isolated secondary structural loops in this region may provide not only a nucleus for productive folding, but may also be an area where potential non-native association may occur. It is of interest to note that a truncation at Trp156 in HyD-Crys has been implicated in juvenile onset cataract (Knoch et al. 2000). Perhaps elimination of the “micro-domain” in this region decreases stability of the molecule and opens a pocket for self-association. Clearly this region must be important for maintaining global solubility of HyD-Crys.



Figure 7-2

Ribbon structure of human γ S crystallin showing the location of the conformationally important prolines, Pro154 and Pro162 (Purkiss et al. 2002).

We were interested in determining the residues responsible for the unfolding and refolding hysteresis in this “micro-domain” so we examined all residues within 8 Å of Trp156 in the crystal structure of native H γ D-Crys. We then aligned these residues between H γ S-Crys and H γ D-Crys to identify any differences in the two sequences within this region (Table 8-1).

Table 7-1

The residues within 8 Å of Trp156 in H γ D-Crys that have amino acid composition differences between H γ S-Crys and H γ D-Crys.

Residue in H γ D-Crys	Residue in H γ S-Crys
Glutamic Acid 96	Glycine
Leucine 124	Cysteine
Asparagine 125	Lysine
Leucine 133	Phenylalanine
Tyrosine 154	Proline
Glutamine 155	Isoleucine
Threonine 160	Alanine
Alanine 162	Proline
Arginine 163	Alanine

In general, H γ S-Crys seemed to have more hydrophobic residues surrounding Trp156 than H γ D-Crys. H γ D-Crys had many acidic and basic residues in this region of the molecule. It is possible that the presence of charges actually stabilizes the domain II core to GdnHCl denaturation. It may require a higher concentration or a longer exposure to GdnHCl for the denaturant to out-compete the existing polar interactions present in this area in H γ D-Crys. Hydrogen-deuterium exchange experiments coupled with NMR will determine if residual structure is present in the core surrounding Trp156 in denaturing conditions.

There were two specific differences that were particularly striking in the sequence alignment between H γ D-Crys and H γ S-Crys. Position 154 is a tyrosine and position 162

is an alanine in H γ D-Crys and both align to two different prolines in H γ S-Crys (Table 8-1). These proline residues are located at the top of the molecule in a loop forming region of H γ S-Crys and are closely associated with the β -strand which forms the domain interface of domain II (Figure 8-2).

Proline is the only amino acid that exhibits constrained movement around the α -carbon bond in the peptide backbone. Perhaps by constraining the folding states that the polypeptide backbone can sample in H γ S-Crys, the prolines limit the number of folding intermediates thus eliminating the kinetic barrier to productive unfolding or refolding observed in H γ D-Crys.

If formation or destruction of the “micro-domain” is an important step for the productive unfolding of the molecule, it seems likely that any structural rigidity either increasing the kinetic rate of “micro-domain” formation or limiting the number of conformations sampled by the molecule would expedite H γ S-Crys productive folding and decrease or eliminate the aggregation pathway as well. Experiments performed by Clark and colleagues at Notre Dame have shown that proline functions to prevent aggregation of the β -helical pertactin protein of *Bordetella Pertussis* (M. Junker, C. Schuster, K. Whiteman, and P. L. Clark, personal communication). Alanine scanning of the “micro-domain” residues of H γ D-Crys may elucidate which if any of these residues are responsible for the observed hysteresis and aggregation pathway.

C. DOMAIN SWAPPING AS A MECHANISM FOR CRYSTALLIN AGGREGATION

Given the high structural similarity between the two domains of HyD-Crys (49% similarity and 36% identity), domain swapping is a likely mechanism of aggregation for this protein. In addition, studies utilizing triple tryptophan to phenylalanine mutant proteins showed that a structure formed during unfolding and refolding that had domain I denatured and domain II in a native-like conformation (Kosinski-Collins et al. 2004). Disruption of interface contacts have further been shown to stabilize this intermediate (Flaugh et al. 2004). If a partially unfolded conformation of HyD-Crys exists in solution, it is likely that slow renaturation of the N-terminal domain may involve the formation of both native intramolecular domain formation and non-native intermolecular domain contacts resulting in domain swapping events.

Proline has been shown to be an important residue in domain swapping in p13suc1 (Rousseau et al. 2001). In this protein, two prolines are located in the conformational hinge region. One proline gives the protein the flexibility to rotate into a domain swapped structure while the other proline provides a conformational constraint that prevents high order oligomerization and aggregation (Silow et al. 1999). We have hypothesized that the prolines in HyS-Crys are important to the stability and refolding of HyS-Crys in the region of the “micro-domain”. It is possible that the presence of the two prolines in HyS-Crys function as “traffic guards” limiting the ability of the protein to domain swap as well. HyD-Crys does not, however, have these two prolines in its “micro-domain.” Instead, HyD-Crys has an alanine and a tyrosine residue in place of the prolines. These residues (alanine especially) do not force the conformational restrictions on HyD-Crys inherent to proline cis/trans isomerization. Perhaps structural flexibility related to domain interface formation in HyD-Crys allows domain swapping and aggregation to occur in this protein.

More exhaustive study of the aforementioned prolines may elucidate the mechanism of hysteresis and aggregation in HyD-Crys.

D. CONCLUDING REMARKS

Studies of HyD-Crys have identified an intermediate common to both the unfolding and refolding pathway that retains native structure of the C-terminal domain, but a disordered conformation in the N-terminal domain at pH 7.0, 37°C (Kosinski-Collins et al. 2004). This refolding intermediate has recently been shown to be common to the productive refolding and aggregation pathways (I. A. Mills, S. L. Flaugh, and J. A. King, unpublished results). Studies by Rudolph et al. with bovine γ B crystallin identified a refolding intermediate that had a structure, N-terminal domain and an unfolded C-terminal domain, when refolded in urea at pH 2.0 (1990). Folding studies in urea at 20°C with rat β B2 crystallin, showed that an intermediate formed during unfolding that had an intact C-terminal domain and a denatured N-terminal domain (Wieligmann et al. 1999). A destabilized intermediate form of β B2 has further been shown to be preferentially bound by α -crystallin (Sathish et al. 2004). The presence of a partially unfolded conformer of the $\beta\gamma$ -crystallins may be important to the mechanism of aggregation and cataract formation.

If one domain is unfolded and one domain of a $\beta\gamma$ -crystallin is native-like, we can imagine several possible scenarios for protein polymerization. The unfolded domain may remain unfolded and simply aggregate as a disordered mat surrounded by the folded other domain. It is possible that the unstructured domain may refold into a native-like conformation, but use an adjacent molecule for formation of the domain interface forming a domain-swapped structure. Alternatively, the unstructured domain may form isolated β -strands that may insert into other productively refolded β -sheets in a loop-sheet insertion mechanism. Perhaps though, the unfolded domain refolds into a completely different conformation like a parallel β -helix that has exposed interaction surfaces simply more prone to polymerization.

These possibilities merely assess the contribution of the $\beta\gamma$ -crystallins, but we know cataracts contain the α -crystallins as well. Assessing what type of disrupted

conformations of the $\beta\gamma$ -crystallins are present in the aged lens, and the propensity with which these structures bind α -crystallin, may elucidate how and why protein polymerization occurs in the lens.

The exact mechanism of cataract formation in the aged lens is unknown to date. Numerous studies have been focused on identifying the causative agent of cataract, but we have yet to identify a clear polymerization pathway for the crystallin proteins. It is important to note that we rarely think of cataract as a specifically organized polymer because of the amorphous nature of the high molecular weight species removed from old lenses, but that does not mean that an organized core structure was never present. The fibrous protein deposits identified in the brains of patients diagnosed with AD contain many associated proteins in aggregated mats, even though the A β peptide forms amyloid fibrils as the aggregation initiating species. In addition, studies of many different aggregating proteins have demonstrated that a seed of polymerized protein may recruit other protein species to the multimerizing chain. The organized nature of hereditary, juvenile onset cataracts and the ability of the $\beta\gamma$ -crystallins to form *in vitro* amyloid fibers suggests that mature onset cataracts may involve an ordered aggregate intermediate or nucleus for formation.

We need to approach cataract as a possible ordered aggregation mechanism to understand the underlying basis for the disease. We may have been misled by the “unstructured” nature of the aggregated species removed from cataractous lenses and we need to remember that organized aggregation is still a possibility for mature onset cataract even though it is not immediately obvious from the lack of structure of material removed. We need to also remember that there may be multiple paths that lead to crystallin polymerization and aggregation and that no one mechanism may solely be responsible for cataract. Ultimately, the process of cataractogenesis is extremely complex and may require years of study and experimentation to understand.

CHAPTER VIII: REFERENCES

- Aquilina, J.A., Benesch, J.L., Bateman, O.A., Slingsby, C., and Robinson, C.V. 2003. Polydispersity of a mammalian chaperone: mass spectrometry reveals the population of oligomers in alphaB-crystallin. *Proc. Natl. Acad. Sci.* **100**:10611-10616.
- Aswad, D.W., Paranandi, M.V., and Schurter, B.T. 2000. Isoaspartate in peptides and proteins: formation, significance, and analysis. *J. Pharm. Biomed. Anal.* **21**: 1129-1136.
- Basak, A., Bateman, O., Slingsby, C., Pande, A., Asherie, N., Ogun, O., Benedek, G.B., and Pande, J. 2003. High resolution x-ray crystal structures of human γ D crystallin (1.25 Å) and the R58H mutant (1.15 Å) associated with aculeiform cataract. *J. Mol. Biol.* **328**: 1137-1147.
- Balasubramanian, D., Du, X., and Zigler, J.S.J. 1990. The reaction of singlet oxygen with proteins, with special reference to crystallins. *Photochem. Photobiol.* **52**: 761-768.
- Baldwin, M.A., Pan, K.M., Nguyen, J., Huang, Z., Groth, D., Serban, A., Gasset, M., Mehlhorn, I., Fletterick, R.J., and Cohen, F.E. 1994. Spectroscopic characterization of conformational differences between PrPC and PrPSc: an alpha-helix to beta-sheet transition. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **343**: 435-441.
- Banaszak, L., Winter, N., Xu, Z., Bernlohr, D.A., Cowan, S., and Jones, T.A. 1994. Lipid binding proteins: a family of fatty acid and retinoid transport proteins. *Adv. Prot. Chem.* **45**: 89-151.
- Banzon, J.A. and Kelly, J.W. 1992. β -sheet rearrangements, serpins, and beyond. *Protein Eng.* **5**: 113-115.
- Baskakov, I.V., Legname, G., Prusiner, S.B., and Cohen, F. E. 2001. Folding of prion protein to its native α -helical conformation is under kinetic control. *J. Biol. Chem.* **276**: 19687-19690.
- Bateman, O.A., Sarra, R., van Genesen, S.T., Kappe, G., Lubsen, N.H., and Slingsby, C. 2003. The stability of human acidic β -crystallin oligomers and hetero-oligomers. *Exp. Eye Res.* **77**: 409-422.
- Bax, B., Lapatto, R., Nalini, V., Driessen, H., Lindley, P. F., Mahadevan, D., Blundell, T. L., and Slingsby, C. 1990. X-ray analysis of beta B2-crystallin and evolution of oligomeric lens proteins. *Nature* **347**: 776
- Bihoreau MT, Baudin V, Marden M, Lacaze N, Bohn B, Kister J, Schaad O, Dumoulin A, Edelstein SJ, Poyart C, et al. 1992. Steric and hydrophobic determinants of the solubilities of recombinant sickle cell hemoglobins. *Protein Sci.* **1**: 145-150.

- Blundell, T.L., Lindley, P., Miller, L., Moss, D., Slingsby, C., Tckley, I., Turnell, B., and Wistow, G. 1981. The molecular structure and stability of the eye lens: x-ray analysis of gamma-crystallin II. *Nature* **289**: 771-777.
- Bradford, M.M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**: 248.
- Brodsky, B. and Ramshaw, J.A. 1997. The collagen triple-helix structure. *Matrix Biol.* **15**: 545-554.
- Benedek, G. 1997. Cataract as a protein condensation disease: the Proctor Lecture. *Invest. Ophthalmol. Vis. Sci.* **38**: 1911-1921.
- Bennett, M.J., Choe, S., and Eisenberg, D.S. 1994. Domain swapping: Entangling alliances between proteins. *Proc. Natl. Acad. Sci.* **91**: 3127-3131.
- Bettelheim, F.A. 2002. Kinetics of chaperoning of dithioreitol-denatured alpha-lactalbumin by alpha crystallin. *Int. J. Biol. Macromol.* **18**: 161-169.
- Betts, S.D., Speed, M., and King, J. 1999. Detection of early aggregation intermediates by native gel electrophoresis and native western blotting. *Methods Ezymol.* **309**: 333-350.
- Bradley, P., Cowen, L., Menke, M., King, J., and Berger, B. 2001. BETAWRAP: Successful prediction of parallel beta-helices from primary sequence reveals an association with many microbial pathogens. *Proc. Natl. Acad. Sci.* **98**: 14819-14824.
- Broide, M.L., Berland, C.R., Pande, J., Ogun, O.O., and Benedek, G.B. 1991. Binary-liquid phase separation of lens protein solutions. *Proc. Natl. Acad. Sci.* **88**: 5660-5664.
- Bron, A.J., Vrensen, G.F., Koretz, J., Maraini, G., and Harding J.J. 2000. The ageing lens. *Ophthalmologica* **214**: 86-104.
- Bruce, D., Perry, D.J., Borg, J-Y., Carrell, R.W., and Wardell, M.R. 1994. Thromboembolic disease due to thermolabile conformational changes of antithrombin Rouen-VI. *J. Clin. Invest.* **94**: 2265-2274.
- Burley, S.K. and Petsko, G.A. 1985. Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science* **229**: 23-28.
- Capaldi, A.P. and Radford, S. E: 1998. Kinetic studies of β -sheet protein folding. *Curr. Opin. Struct. Biol.* **8**: 86-9.
- Carrell, R.W. and Gooptu, B. 1998. Conformational changes and disease- serpins prions and Alzheimer's. *Curr. Opin. Struct. Biol.* **8**: 799-809.

- Carrell, R.W., Whisstock, J., and Lomas, D.A. 1994. Conformational changes in the serpins and the mechanism of α -1 antichymotrypsin deficiency. *Amer. J. Respir. Crit. Care Med.* **150**: 171-175.
- Caughey, B. and Lansbury, P.T. 2003. Protofibrils, pores, fibrils, and neurodegeneration: separating the responsible protein aggregates from the innocent bystanders. *Annu. Rev. Neurosci.* **26**: 267-298.
- Chen, B. and King, J. 1991. Thermal unfolding pathway for the thermostable P22 tailspike endorhamnosidase. *Biochemistry* **30**: 6260-6269.
- Chen, Y. and Barkley, M.D. 1998. Towards understanding tryptophan fluorescence in proteins. *Biochemistry* **37**: 9976-9982.
- Chung, J., Yang, H., de Beus, M.D., Ryu, C.Y., Cho, K., and Colon, W. 2003. Cu/Zn superoxide dismutase can form pore-like structures. *Biochem. Biophys. Res. Commun.* **312**: 873-876.
- Clark, J.I. 1994. *Principle and Practice of Ophthalmology*. Saunders College Publishing, Philadelphia, PA.
- Clark, J.I. and Muchowski, P.J. 2000. Small heat-shock proteins and their potential role in human disease. *Curr. Opin. Struct. Biol.* **10**: 52-59.
- Clovis, C. and Garland, D. 2002. Posttranslational modification of human alphaA-crystallin: correlation with electrophoretic migration. *Arch. Biochem. Biophys.* **397**: 319-323.
- Cohen, C. and Parry, D.A.D. 1990. α -helical coiled coils and bundles: how to design and α -helical protein. *Proteins* **7**: 1-15.
- Colon, W. and Kelly, J.W. 1992. Partial denaturation of transthyretin is sufficient for amyloid fibril formation in vitro. *Biochemistry* **31**: 8654-8660.
- Colon, W. and Kelly, J.W. 1991. Transthyretin acid induced denaturation is sufficient for amyloid fibril formation in vitro. In *Applications of Enzyme Biotechnology*. (eds. Kelly, J.W. and Baldwin, T.O.), pp 99-108. Plenum, New York, NY.
- Conti, F., Cantu, A.M., and Duclouhier, H. 1988. Orientation of the tryptophans responsible for the photoinactivation of nerve sodium channels. *Eur. Biophys. J.* **16**: 73-81.
- Conway, K.A., Lee, S.J., and Lansbury, P.T.Jr. 1998. Accelerated *in vitro* fibril formation by a mutant alpha-synuclein linked to early-onset Parkinson's disease. *Nat. Med.* **4**: 1318-1320.

- Dalessio, P. M. & Ropson, I. J. 1998. pH dependence of the folding pathway of intestinal fatty acid binding protein. *Arch. Biochem. Biophys.* **359**: 199-208.
- Das, K.P., Choo-Smith, L-P., Petrash, J.M., Surewicz, W.K. 1999. Insight into the secondary structure of non-native proteins bound to a molecular chaperone α -crystallin. *J. Biol. Chem.* **274**: 33209-33212.
- Davies, M.J. and Truscott, R.J.W. 2001. Photo-oxidation of proteins and its role in cataractogenesis. *J. Photochem. & Photobiol.* **63**: 114-125.
- Delage, M.T. and Tardieu, A. 1983. Short-range order of crystallin proteins accounts for eye lens transparency. *Nature* **302**: 415-417.
- Devlin, G.L., Carver, J.A., Bottomley, S.P. 2003. The selective inhibition of serpin aggregation by the molecular chaperone, alpha crystallin, indicates a nucleation-dependent specificity. *J. Biol. Chem.* **278**: 48644-48650.
- Dill, K.A. and Chan, H.S. 1997. From Levinthal pathways to funnels. *Nat. Struct. Biol.* **4**: 10-19.
- Dunker, A.K. and Jones, T.C. 1978. Proposed knobs-into-holes packing for several membrane proteins. *Membr. Biochem.* **2**: 1-16.
- Dunstone, M.A., Dai, W., Whisstock, J.C., Rossjohn, J., Pike, R.N., Feil, S.C., LeBonniec, B.F., Parker, M.W., and Bottomley, S.P. 2000. Cleaved antitrypsin polymers at atomic resolution. *Protein Sci.* **9**: 417-420.
- Eaton, W.A., Muñoz, V., Thompson, P.A., Chan, C-K., and Hofrichter, J. 1997. Submillisecond kinetics of protein folding. *Curr. Opin. Struct. Biol.* **7**: 10-14.
- Emsley, P., Charles, I.G., Fairweather, N.F., and Isaacs, N.W. 1996. Structure of Bordetella Pertussis virulence factor P.69 pertactin. *Nature* **381**: 90-92.
- Eyles, S.J. and Gierasch, L.M. 2000. Multiple roles of prolyl residues in structure and folding. *J. Mol. Biol.* **301**: 737-747.
- Fersht, A. 1999. Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding. W.H. Freeman and Company. New York, NY.
- Finke, J.M., Roy, M., Zimm, B.H., and Jennings, P.A. 2000. Aggregation events occur prior to stable intermediate formation during refolding of interleukin 1- β . *Biochemistry* **39**: 575-583.

- Fischer, B., Sumner, I., and Goodenough, P. 1993. Renaturation of lysozyme-- temperature dependence of renaturation rate, renaturation yield, and aggregation: identification of hydrophobic folding intermediates. *Arch. Biochem. Biophys.* **306**: 183-187.
- Frederikse, P.H. 2000. Amyloid-like protein structure in mammalian ocular lens. *Curr. Eye Res.* **20**: 462-468.
- Flaugh, S.L., Kosinski-Collins, M.S., and King, J.A. The role of the hydrophilic interface residues in folding and aggregation of human γ D crystallin. *Exp. Eye Res.* Manuscript submitted for review.
- Fu, L. and Liang, J.J-N. 2002. Unfolding of human lens recombinant β B2- and γ C-crystallin. *J. Struct. Biol.* **139**: 191-198.
- Fu, L. and Liang, J. 2003. Alteration of protein-protein interactions of congenital cataract crystallin mutants. *Invest. Ophthalmol. Vis. Sci.* **44**: 1155-1159.
- Fujii, N., Awakura, M., Takemoto, L., Inomata, M., Takata, T., Fujii, N., and Saito, T. 2003. Characterization of alpha-crystallin from high molecular weight aggregates in the normal human lens. *Mol. Vis.* **9**: 315-322.
- Fujii, N., Harada, K., Momose, Y., Ishii, N., and Akaboshi, M. 1999. D-amino acid formation induced by a chiral field within a human lens protein during aging. *Biochem. Biophys. Res. Comm.* **263**: 322-326.
- Ganesh, C., Zaidi, F.N., and Udgaonkar, J.B., and Varadarajan, R. 2001. Reversible formation of on-pathway macroscopic aggregates during the folding of maltose binding protein. *Protein Sci.* **10**: 1635-1644.
- Ghelis, C. and Yon., J. 1982. Overview of the reversibility of the unfolding-folding process. In *Protein Folding*. (eds. B. Horecker, N.O. Kaplan, J. Marmur, & H. Scheraga), pp 285-295. Academic Press, New York, NY.
- Goldberg, M.E. 1969. Tertiary structure of *Escherichia Coli* β -D-galactosidase. *J. Mol. Biol.* **46**: 441-446.
- Guex, N.P., and Peitsch, M.C. 1997. SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* **18**: 2714-2723.
- Gunasekaran, K., Eyles, S.J., Hagler, A.T., and Gierasch, L.M. 2001. Keeping it in the family: folding studies of related proteins. *Curr. Opin. Struct. Biol.* **11**: 83-93.
- Haase-Pettingell, C.A. and King J. 1988. Formation of aggregates from a thermolabile folding intermediate in P22 tailspike maturation. A model for inclusion body formation. *J. Biol. Chem.* **263**: 4977-4983.

Haley, D.A., Bova, M.P., Huang, Q.L., Mchaourab, H.S., Stewart, P.L. 2000. Small heat-shock protein structures reveal a continuum from symmetric to variable assemblies. *J. Mol. Biol.* **298**: 261-72.

Hanson, S.R., Smith, D.L., and Smith, J. 1998. Deamidation and disulfide bonding in human lens gamma-crystallins. *Exp. Eye Res.* **67**: 301-312.

Hanson, S.R., Hasan, A., Smith, D.L., and Smith, J. 2000. The major in vivo modifications of the human water-insoluble lens crystallins are disulfide bonds, deamidation, methionine oxidation and backbone cleavage. *Exp. Eye Res.* **71**: 195-207.

Harding, J.J. and Crabbe, M.J.C. 1984. The lens: development, proteins, metabolism, and cataract. In, *The Eye*, vol.1B. (ed. Dawson, H.), pp. 207-492. Academic Press, London.

Harper, J.D., Lieber, C.M., and Lansbury, P.T. Jr. 1997. Atomic force microscopic imaging of seeded fibril formation and fibril branching by the Alzheimer's disease amyloid-beta protein. *Chem. Biol.* **4**: 951-959.

Harper, J.D., Wong, S.S., Lieber, C.M., and Lansbury, P.T. Jr. 1997. Observation of metastable Abeta amyloid protofibrils by atomic force microscopy. *Chem. Biol.* **4**: 119-125.

Harper, J.D., Wong, S.S., Lieber, C.M., and Lansbury, P.T. Jr. 1999. Assembly of A β amyloid protofibrils: an in vitro model for a possible early event in Alzheimer's disease. *Biochemistry* **38**: 17851-17864.

Hatters, D.M., Lindner, R.A., Carver, J.A., Howlett, G.J. 2001. The molecular chaperone, alpha crystallin, inhibits amyloid formation by apolipoprotein C-II. *J. Biol. Chem.* **276**: 33755-33761.

Hay, R.E., Andley, U.P., and Petrash, J.M. 1993. Expression of recombinant bovine γ B-, γ C-, and γ D-crystallins and correlation with native protein. *Exp. Eye Res.* **58**: 573-584.

He, Q., Mason, A.B., Lyons, B.A., Tam, B.M., Nguyen, V., MacGillivray, R.T.A., and Woodworth, R.C. 2001. Spectral and metal-binding properties of three single-point tryptophan mutants of the human transferrin N-lobe. *Biochem. J.* **354**: 423-429.

Heon, E., Priston, M., Schorderet, D.F., Billingsley, G.D., Girard, P.O., Lubsen, N., and Munier, F.L. 1999. The γ -crystallins and human cataracts: the puzzle made clearer. *Am. J. Hum. Genet.* **65**: 1261-1267.

Hollien, J. and Marqusee, S. 2002. Comparison of the folding processes of *T. thermophilus* and *E. Coli* ribonucleases H. *J. Mol. Biol.* **316**: 327-340.

- Horovitz, A. and Fersht, A. 1990. Strategy for analyzing the cooperativity of intramolecular interactions in peptides and proteins. *J. Mol. Biol.* **214**: 613-617.
- Horwitz, J. 2000. The function of alpha-crystallin in vision. *Sem. Cell & Dev. Biol.* **11**: 53-60.
- Horwitz, J., Bova, M.P., Ding, L.L., Haley, D.A., Stewart, P.L. 1999. Lens alpha crystallin: structure and function. *Eye* **13**: 403-408.
- Hotze, E.M., Heuck, A.P., Czajkowsky, D.M., Shao, Z., Johnson, A.E., and Tweten, R.K. 2002. Monomer-monomer interactions drive the prepore to pore conversion of a beta-barrel-forming cholesterol-dependent cytolyisin. *J. Biol. Chem.* **277**: 11597-11605.
- Hu, J.C., O'Shea, E.K., Kim, P.S., and Sauer, R.T. 1990. Sequence requirements for coiled-coils: analysis with lambda repressor-GCN4 leucine zipper fusions. *Science* **250**: 1400-1403.
- Huntington, J.A., Pannu, N.S., Hazes, B., Read, R.J., Lomas, D.A., and Carrell, R.W. 1999. A 2.6 Å structure of a serpin polymer and implications for conformational disease. *J. Mol. Biol.* **293**: 449-455.
- Ingram, V.M. 1956. A specific chemical difference between globins of normal human and sickle cell anaemia haemoglobin. *Nature* **178**: 792-794.
- Inoue, S., Kuroiwa, M., Ohashi, K., Hara, M., and Kisilevsky, R. 1997. Ultrastructural organization of hemodialysis-associated beta 2-microglobulin amyloid fibrils. *Kidney Int.* **52**: 1543-1549.
- Jaenicke, R. 1999. Stability and folding of domain proteins. *Prog. Biophys. & Mol. Biol.* **71**: 155-241.
- Janowski, R., Kozak, M., Jankowska, E., Grzonka, Z., Grubb, A., Abrahamson, M., and Jaskolski, M. 2001. Human cystatin C, an amyloidogenic protein dimerizes through three-dimensional domain swapping. *Nat. Struct. Biol.* **8**: 316-320.
- Jenkins, J. and Pickersgill, R. 2001. The architecture of parallel beta-helices and related folds. *Prog Biophys Mol Biol.* **77**: 111-175.
- Jensen, P.H., Nielsen, M.S., Jakes, R., Dotti, C.G., and Goedert, M. 1998. Binding of α -synuclein to brain vesicles is abolished by familial Parkinson's disease mutation. *J. Biol. Chem.* **273**: 292-294.
- Kamtekar, S. and Hecht, M.H. 1995. Protein Motifs 7: The four-helix bundle: what determines a fold? *FASEB J.* **9**: 1013-1022.

Kawahara, M. and Kuroda, Y. 2000. Molecular mechanism of neurodegeneration induced by Alzheimer's β -amyloid protein: Channel formation and disruption of calcium homeostasis. *Brain Res. Bul.* **53**: 389-397.

Kayed, R., Head, E., Thompson, J.L., McIntire, T.M., Milton, S.C., Cotman, C.W., and Glabe, C.G. 2003. Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science* **300**: 486-489.

Keating AE, Malashkevich VN, Tidor B, and Kim PS. 2001. Side-chain repacking calculations for predicting structures and stabilities of heterodimeric coiled coils. *Proc. Natl. Acad. Sci.* **98**: 14825-14830.

Kelly, J.W. 1996. Alternative conformations of amyloidogenic proteins govern their behavior. *Curr. Opin. Struct. Biol.* **6**: 11-17.

Kim, S.-J., Woo, J.-R., Seo, E.J., Yu, M.-H., Ryu, S.-E. 2001. A 2.1 Å resolution structure of an uncleaved alpha1-antitrypsin shows variability of the reactive center and other loops. *J. Mol. Biol.* **306**: 109-119.

King, J.A., Haase-Pettingell, Robinson, A.S., Speed, M., and Mitraki, A. 1996. Thermolabile folding intermediates: inclusion body precursors and chaperonin substrates. *FASEB J.* **10**: 57-66.

Kivela, T., Tarkkanen, A., Frangione, B., Ghiso, J., and Haltia, M. 1994. Ocular amyloid deposition in familial amyloidosis, an analysis of native and variant gelsolin in Meretoja's syndrome. *Invest. Ophthalmol. Vis. Sci.* **10**: 3759-3769.

Kmoch, S., Brynda, J., Asfaw, B., Bezouska, K., Novak, P., Rezacova, P., Ondrova, L., Filipic, M., Sedlacek, J., and Elleder, M. 2000. Link between a novel human γ D-crystallin allele and a unique cataract phenotype explained by protein crystallography. *Hum. Mol. Genet.* **9**: 1779-1786.

Knaus, K.J., Morillas, M., Swietnicki, W., Malone, M., Surewicz, W.K., and Yee, Y.C. 2001. Crystal structure of the human prion protein reveals a mechanism for oligomerization. *Nat. Struct. Biol.* **8**: 770-774.

Kosinski-Collins, M.S., Flaugh, S.L., and King, J. 2004. Probing Folding and Fluorescence Quenching in Human γ D Crystallin Greek Key Domains Using Triple Tryptophan Mutant Proteins. *Protein Sci.* **13**: 2223-2235.

Kosinski-Collins, M.S. and King, J. 2003. In vitro unfolding, refolding, and polymerization of human γ D crystallin, a protein involved in cataract formation. *Protein Sci.* **12**:480-490.

Kurzel, R.B., Wolbarsht, M., Yamanashi, B.S., Staton, G.W., and Borkman, R.F. 1973. Tryptophan excited states and cataracts in the human lens. *Nature* **241**: 132-133.

- Kuwajima, K. 1992. Protein folding in vitro. *Curr. Opin. Biotechnol.* **3**: 462-467.
- Lai, Z., Colon, W., and Kelly, J.W. 1996. The acid-mediated denaturation of transthyretin proceeds through an intermediate that partitions into amyloid. *Biochemistry* **35**: 6470-6482.
- Lai, Z., McCulloch, J., Lashuel, H.A., Kelly, J.W. 1997. Guanidine hydrochloride-induced denaturation and refolding of transthyretin exhibits a marked hysteresis: equilibria with high kinetic barriers. *Biochemistry* **36**: 10230-10239.
- Lakowicz, J.R. 1999. Principles of Fluorescence Spectroscopy. Kluwer Academic Plenum Publishers, New York, New York.
- Lampi, K.J., Oxford, J.T., Bachinger, H.P., Shearer, T.R., David L.L., and Kapfer, D.M. 2001. Deamidation of human beta B1 alters the elongated structure of the dimer. *Exp. Eye Res.* **72**: 279-288.
- Lander E.S. et al. International Human Genome Sequencing Consortium. 2001. Initial sequencing and analysis of the human genome. *Nature* **409**: 860-921.
- Lansbury, P.T. 1995. The chemistry of the scrapie infection- Implications of the ICE 9 metaphor. *Chem. Biol.* **2**: 1-5.
- Lapko, V.N., Purkiss, A.G., Smith, D.L., and Smith, J.B. 2002. Deamidation in human γ S-crystallin from cataractous lenses is influenced by surface exposure. *Biochemistry* **41**: 8638-8648.
- Lapko, V.N., Smith, D.L., and Smith, J.B. 2002. S-methylated cysteines in human lens gamma S-crystallins. *Biochemistry* **41**: 14645-14651.
- Lashuel, H., Lai, Z., and Kelly, J.W. 1998. Characterization of the transthyretin acid denaturation pathways by analytical ultracentrifugation: implications for wild-type, V30M, and L55P amyloid fibril formation. *Biochemistry* **37**: 17851-17864.
- Lashuel, H.A., Hartley, D.M., Petre, B.M., Wall, J.S., Simon, M.N., Walz, T., and Lansbury, P.T.Jr. 2003. Mixtures of wild-type and a pathogenic (E22G) form of A β 40 *in vitro* accumulate protofibrils, including amyloid pores. *J. Mol. Biol.* **332**: 795-808.
- Li, J., Uversky, V.N., and Fink, A.L. 2001. Effect of familial Parkinson's disease point mutations A30P and A53T on the structural properties, aggregation, and fibrillation of human alpha-synuclein. *Biochemistry* **40**: 11604-11613.
- Lin P.P., Barry R.C., Smith D.L., and Smith J.B. 1998. In vivo acetylation identified at lysine 70 of human lens alphaA-crystallin. *Protein Sci.* **7**: 1451-1457.

- Liu, Y. and Eisenberg, D. 2002. 3D domain swapping: As domains continue to swap. *Protein Sci.* **11**: 1285-1299.
- Liu, Y., Gotte, G., Libonati, M., and Eisenberg, D. 2001. A domain swapped RNase dimer with implications for amyloid formation. *Nat. Struct. Biol.* **8**: 211-214.
- Liu, Y., Gotte, G., Libonati, M., and Eisenberg, D. 2002. Structures of the two 3D domain swapped RNase trimers. *Protein Sci.* **11**: 371-380.
- London, J., Skrzynia, C., and Goldberg, M.E. 1974. Renaturation of Escherichia Coli tryptophanase after exposure to 8 M urea. Evidence for the existence of nucleation center. *Eur. J. Biochem.* **47**: 409-415.
- Lotharius, J. and Brundin, P. 2002. Impaired dopamine storage resulting from α -synuclein mutations may contribute to the pathogenesis of Parkinson's disease. *Hum. Mol. Genet.* **11**: 2395-23407.
- Lubsen, N.H., Aarts, H.J.M., and Schoenmakers, J.G.G. 1988. The evolution of lenticular proteins: β - and γ -cystallin super gene family. *Prog. Biophys. Mol. Biol.* **51**: 47-67.
- Luhrs, T., Riek, R., Guntert, P., and Wuthrich, K. 2003. NMR structure of the human doppel protein. *J. Mol. Biol.* **326**: 1549.
- Lupas, A. 1996. Coiled coils: new structures and new functions. *Trends Biochem. Sci.* **21**: 375-382.
- Ma, Z., Hanson, S.R., Lampi, K.J., David, L.L., Smith, D.L., and Smith, J.B. 1998. Age-related changes in human lens crystallins identified by HPLC and mass spectrometry. *Exp. Eye Res.* **67**: 21-30.
- MacRae, T. 2000. Structure and function of small heat shock/ α -crystallin proteins: established concepts and emerging ideas. *Cell Mol. Life Sci.* **57**: 899-913.
- Marini, D.M., Hwang, W., Laffenburger, D.A., Zhang, S., and Kamm, R.D. 2002. Left-handed helical ribbon intermediates in the self-assembly of a β -sheet peptide. *Nano Lett.* **2**: 295-299.
- Mayr, E.M., Jaenicke, R., and Glockshuber, R. 1997. The domains in γ B crystallin: Identical fold-different stabilities. *J. Mol. Biol.* **269**: 260-269.
- McCutchen, S.L., Lai, Z., Miroy, G., Kelly, J.W., and Colon, W. 1995. Comparison of lethal and nonlethal transthyretin variants and their relationship to amyloid disease. *Biochemistry* **34**: 13527-13536.

- McCarty, C.A., and Taylor, H.R. 2002. A review of the epidemiological evidence linking ultraviolet radiation and cataracts. In, *Progress in Lens and Cataract Research*. (eds. Hockwin et al.) *Dev. Ophthalmol.*, **35**: 21-31.
- McGaughey, G.B., Gagné, M., and Rappé, A.K. 1998. π -Stacking interactions: alive and well in proteins. *J. Biol. Chem.* **273**: 15458-15463.
- Meehan, S., Berry, Y., Luisi, B., Dobson, C.M., Carver, J.A., and MacPhee, C.E. 2004. Amyloid fiber formation by lens crystallin proteins and its implications for cataract formation. *J. Biol. Chem.* **279**: 3413-3419.
- Mitraki, A. and King, J. 1989. Protein folding intermediates and inclusion body formation. *Biotechnology* **7**: 690-697.
- Munoz, V. and Serrano, L. 1995. Helix design, prediction and stability. *Curr. Opin. Biotechnol.* **6**: 382-386.
- Najmudin, S. 1993. Structure of the bovine eye lens protein gammaB(gammaII)-crystallin at 1.47 Å. *Acta Crystallogr.* **49**: 223-233.
- Nandrot E., Slingsby C., Basak A., Cherif-Chefchaoui, M., Benazzouz B., Hajaji Y., Boutayeb S., Gribouval O., Arbogast L., Berraho A., Abitbol M., and Hilal L. 2003. Gamma-D crystallin gene (CRYGD) mutation causes autosomal dominant congenital cerulean cataracts. *J. Med. Genet.* **40**: 262-267.
- National Eye Institute (US) 2002. "Vision problems in the US." Prevent Blindness in America.
- Nelson, G.A., Edward, D.P., and Wilensky, J.T. 1999. Ocular amyloidosis and secondary glaucoma. *Ophthalmology* **106**: 1363-1366.
- Nguyen, J., Balwin, M.A., Cohen, F.E., and Prusiner, S.B. 1995. Prion protein peptides induce α -helix to beta sheet conformational transitions. *Biochemistry* **34**: 4186-4192.
- Nolting, B., Golbik, R., and Fersht, A.R. 1995. Submillisecond events in protein folding. *Proc. Natl. Acad. Sci.* **92**: 10668-10672.
- Norledge, B.V., Mayr, E.M., Glockshuber, R., Bateman, O.A., Slingsby, C., Jaenicke, R., and Driessen, H.P.C. 1996. *Nat. Struct. Biol.* **3**: 267-274.
- Norledge, B.V., Har, R.E., Bateman, O.A., Slingsby, C., and Driessen, H.P.C. 1997. Towards a molecular understanding of phase separation in the lens: A comparison of the x-ray structure of two high T_c γ -crystallins, γE and γF , with two low T_c γ -crystallins, γB and γD . *Exp. Eye Res.* **65**: 609-630.

- Onuchic, J.N., Luthey-Schulten, Z., and Wolynes, P.G. 1997. Theory of protein folding: the energy landscape perspective. *Ann. Rev. Phys. Chem.* **48**: 545-600.
- O'Shea, E.K., Klem, C. Kim, P., and Alba, T. 1991. X-ray structure of the Gcn3 Leucine Zipper, A two stranded parallel coiled-coil. *Science* **254**: 539-544.
- Oyster, C.W. 1999. *The Human Eye Structure and Function*. Sinauer Associates, Inc., Sunderland, MA.
- Pace, C.N., Shirley, B.A., and Thomson, J.A. 1989. Measuring the conformational stability of a protein. In *Protein Structure: A Practical Approach*. (Creighton, T.E., ed.), pp 311-330. Oxford University Press, UK.
- Pan, K.M., Baldwin, M., Nguyen, J., Gasset, M., Serban, A., Groth, D., Mehlhorn, I., Huang, Z., Fletterick, R.J., and Cohen, F.E. 1993. Conversion of alpha-helices into beta-sheets features in the formation of the scrapie prion proteins. *Proc. Natl. Acad. Sci.* **90**: 10962-10966.
- Pande, J., McDermott, M.J., Callender, R.H., and Spector, A. 1991. The calf γ -crystallins- A raman spectroscopic study. *Exp. Eye Res.* **52**: 192-197.
- Pande, A., Pande, J., Asherie, N., Lomakin, A., Ogun, O., King, J.A., Lubsen, N.H., Walton, D., and Benedek, G.B. 2000. Molecular Basis of a progressive juvenile-onset hereditary cataract. *Proc. Natl. Acad. Sci.* **97**: 1993-1998.
- Pande, A., Pande, J., Asherie, N., Lomakin, A., Ogun, O., King, J.A., and Benedek, G.B. 2001. Crystal cataracts: Human genetic cataract caused by protein crystallization. *Proc. Natl. Acad. Sci.* **98**: 6116-6120.
- Peitsch, M.C. 1995. Protein Modeling by E-mail. *Biotechnology* **13**: 658-660.
- Peitsch, M.C. 1996. ProMod and Swiss-Model: Internet-based tools for automated comparative protein modeling. *Biochem. Soc. Trans.* **24**: 274-279.
- Pepys, M.B., Hawkins, P.N., Booth, D.R., Vigushin, D.M., Tennet, G.A., Soutar, A.K., Totty, N., Nguyen, O., Blake, C.C.F., Terry C.J. et al. 1993. Human lysozyme gene mutations cause hereditary systemic amyloidosis. *Nature* **362**: 553-557.
- Prinsze C., Dubbleman, T.M., and Van Steveninck, J. 1990. Protein damage induced by small amounts of photodynamically generated singlet oxygen or hydroxyl radicals. *Biochim. Biophys. Acta.* **1038**: 152-157.
- Prusiner, S. 1997. Prion diseases and the BSE crisis. *Science* **278**: 245-251

- Ptitsyn, O.B. 1995. Structures of folding intermediates. *Curr. Opin. Struct. Biol.* **5**: 74-78.
- Ptitsyn, O.B. 1998. Protein folding: nucleation and compact intermediates. *Biochemistry* **63**: 367-73.
- Purkiss, A.G., Bateman, O.A., Goodfellow, J.M., Lubsen, N.H., and Slingsby, C. 2002. The x-ray crystal structure of human γ S crystallin C-terminal domain. *J. Biol. Chem.* **277**: 4199-4205.
- Rabilloud, T., Carpentier, G., and Tarroux, P. 1988. Improvement and simplification of low background silver staining of proteins by using sodium dithionite. *Electrophoresis* **9**: 288-291.
- Rahkit, R., Cunningham, P., Furtos-Matei, A., Dahan, S., Qi, X.F., Crow, J.P., Cashman, N.R., Kondejewski, L.H., and Chakrabarty A. 2002. Oxidation induced misfolding and aggregation of superoxide dismutase and its implication for amyotrophic lateral sclerosis. *J. Biol. Chem.* **277**: 551-556.
- Reader, J.S., Van Nuland, N.A., Thompson, G.S., Ferguson, S.J., Dobson, C.M., and Radford, S.E. 2001. A partially folded intermediate species of the β -sheet protein apopseudoazurin is trapped during proline limited folding. *Protein Sci.* **10**: 1216-1224.
- Reddy, G.B., Narayanan, S., Reddy, P.Y., and Surolia I. 2002. Suppression of DTT-induced aggregation of abrin by alphaA and alphaB crystallins: A model aggregation assay for alpha-crystallin chaperone activity in vitro. *FEBS Lett.* **522**: 59-64.
- Reixach, N., Deechongkit, S., Jiang, X., Kelly, J.W., Buxbaum, J.N. 2004. Tissue damage in the amyloidoses: Transthyretin monomers and nonnative oligomers are the major cytotoxic species in tissue culture. *Proc. Natl. Acad. Sci.* **101**: 2817-2822.
- Ren, Z., Li, A., Shastry, B.S., Padma, T., Ayyagari, R., Scott, M.H., Kaiser-Kupfer, M.I., Hejtmann, J.F. 2000. A 5-base pair insertion in the gammaC crystallin gene is associated with autosomal dominant variable zonular pulverulent cataract. *Hum Genet.* **106**: 531-537.
- Rousseau, F., Schymkowitz, J.W., and Itzhaki, L.S. 2003. The unfolding story of three dimensional domain swapping. *Structure* **11**: 243-251.
- Rousseau, F., Schymkowitz, J.W., Wilkinson, H.R. and Itzhaki, L.S. 2001. Three-dimensional domain swapping in p13suc1 occurs in the unfolded state and is controlled by proline residues. *Structure* **11**: 243-251.

- Rudolph, R., Sienbendritt, R., Nessler, G., Sharma, A.K., and Jaenicke, R. 1990. Folding of an all-beta protein: Independent domain folding in γ II-crystallin from calf eye lens. *Proc. Natl. Acad. Sci.* **87**: 4625-4629.
- Ryu, S.E., Choi, H.J., Kwon, K.S., Lee, K.N., and Yu, M.H. 1996. The native strains in the hydrophobic core and flexible reactive loop of a serine protease inhibitor: crystal structure of an uncleaved alpha1-antitrypsin at 2.7 Å. *Structure* **4**: 1181-1192.
- Safar, J., Roller, P.P., Gajdusek, D.C., and Gibbs, C.J. Jr. 1993. Thermal stability and conformational transitions of scrapie amyloid (prion) protein correlate with infectivity. *Protein Sci.* **2**: 2206-2216.
- SaintJean, A.P., Phillips, K.R., Creighton, D.J., and Stone, M.J. 1998. Active monomeric and dimeric forms of *Pseudomonas putida* glyoxalase I: evidence for 3D domain swapping. *Biochemistry* **37**: 10345-10353.
- Sandilands, A., Hutcheson, A.M., Long, H.A., Prescott, A.R., Vrensen, G., Loster, J., Klopp, N., Lutz, R.B., Graw, J., Masaki, S., Dobson, C.M., MacPhee, C.E., Quinlan, R.A. 2002. Altered aggregation properties of mutant gamma crystallins cause inherited cataract. *EMBO J.* **21**: 6005-6014.
- Santhiya, S.T., Shyam Manohar, M., Rawley, D., Vijayalakshmi, P., Namperumalsamy, P., Gopinath, P.M., Loster, J., and Graw, J. 2002. Novel mutations in the gamma crystallin genes cause autosomal dominant congenital cataracts. *J. Med. Genet.* **39**: 352-358.
- Saraiva, M.J.M. 2001. Transthyretin mutations in hyperthyroxinemia and amyloid diseases. *Hum. Mutat.* **17**: 493-503.
- Sasaki, H., Jonasson, F., Shui, Y.B., Kojima, M., Ono, M., Katoh, N., Cheng, H-M., Takahashi, N., and Sasaki, K. 2002. High prevalence of nuclear cataract in the population of tropical and subtropical areas. In, Progress in Lens and Cataract Research. (eds. Hockwin et al.) *Dev. Ophthalmol.* **35**: 60-69.
- Sathish, H.A., Koteiche, H.A., and McHaourab, H.S. 2004. Binding of destabilized betaB2-crystallin mutants to alpha-crystallin: the role of a folding intermediate. *J. Biol. Chem.* **279**: 16425-16432.
- Schormann, N., Murrell, J.R., and Benson, M.D. 1998. Tertiary structures of amyloidogenic and non-amyloidogenic variants of transthyretin variants : new model for amyloid fibril formation. *Amyloid* **5**: 175-187.
- Serrano, L., Bycroft, M., and Fersht, A.R. 1990. Aromatic-aromatic interactions and protein stability. *J. Mol. Biol.* **218**: 465-475.

- Shentu, X., Yao, K., Xu, W., Zheng, S., Hu, S., and Gong, X. 2004. Special fasciculiform cataract caused by a mutation in the gammaD crystallin gene. *Mol. Vis.* **29**: 233-239.
- Sifers, R.N. 1995. Defective protein folding as a cause of disease. *Nature Struct. Biol.* **2**: 355-357.
- Silow, M., Tan, Y.J., Fersht, A.R., and Oliveberg, M. 1999. Formation of short-lived protein aggregates directly from the coil in two-state folding. *Biochemistry* **38**: 13006-13012.
- Sinclair, J.F., Ziegler M.M., and Baldwin T.O. 1994. Kinetic partitioning during refolding yields multiple native states. *Nat. Struct. Biol.* **1**: 320-326.
- Sipe, J.D. 1994. Amyloidosis. *Crit. Rev. Clin. Lab. Sci.* **31**: 325-354.
- Sliney, D. H. 2002. How light reaches the eye and its components. *Inter. J. Toxicol.* **21**: 501-509.
- Slingsby, C., Norledge, B., Simpson, A., Bateman, O.A., Wright, G., Driessen, H.P.C., Lindley, P.F., Moss, D.S., and Bax, B. 1997. X-ray diffraction and structure of crystallins. *Prog. Retin. & Eye Res.* **16**: 3-29.
- Soderberg, P.G., Lofgren, S., Ayala, M., Dong, X., Kakar, M., and Mody, V. 2002. Toxicity of ultraviolet radiation exposure to the lens expressed by maximum tolerable dose. In, Progress in Lens and Cataract Research. (eds. Hockwin et al.) *Dev. Ophthalmol.* **35**: 70-75.
- Speed, M.A., Wang, D.I.C., and King, J. 1995. Multimeric intermediates in the pathway to the aggregated inclusion body state for P22 tailspike polypeptide chains. *Protein Sci.* **4**: 900-908.
- Srivastava, O.P., Kirk, M.C., and Srivastava, K. 2004. Characterization of covalent multimers of crystallins in aging human lenses. *J. Biol. Chem.* **279**: 10901-10909.
- Srivastava, O.P. and Srivastava, K. 2003. Existence of deamidated alphaB-crystallin fragments in normal and cataractous human lenses. *Mol. Vis.* **9**:110-118.
- Stephan, D.A., Gillanders, E., Vanderveen, D., Freas-Lutz, D., Wistow, G., Baxevanis, A.D., Robbins, C.M., VanAuken, A., Quesenberry, M.I., Bailey-Wilson, J., Juo, S.H., Trent, J.M., Smith, L., Brownstein, M.J. 1999. Progressive juvenile-onset punctate cataracts caused by mutation of the gammaD-crystallin gene. *Proc. Natl. Acad. Sci.* **96**: 1008-1012.

- Stieber, A., Gonatas, J.O., Moore, J.S., Bantly, A., Yim, H.S., Yim, M.B., and Gonatas, N.K. 2004. Disruption of the structure of the Golgi apparatus and the function of the secretory pathway by mutants G93A and G85R of Cu, Zn superoxide dismutase (SOD1) of familial amyotrophic lateral sclerosis. *J. Neurol. Sci.* **219**: 45-53.
- Sturtevant, J.M., Yu., M.-H., Haase-Pettingell, C., and King, J. 1989. Thermostability of temperature sensitive folding mutants of the P22 tailspike protein. *J. Biol. Chem.* **264**: 10693-10698.
- Subramaniam, V., Bergenheim, N.C.H., Gafni, A., Steel, D. 1995. Phosphorescence reveals a continued slow annealing of the protein core following reactivation of *Escherichia coli* alkaline phosphatase. *Biochemistry* **34**: 1133-1136.
- Takemoto, L. 2001. Deamidation of Asn-143 of gamma S crystallin from protein aggregated of the human lens. *Curr. Eye Res.* **22**: 148-153.
- Takemoto, L., Fujii, N., and Boyle, D. 2001. Mechanism of asparagine deamidation during human senile cataractogenesis. *Exp. Eye Res.* **72**: 559-563.
- Takemoto, L. and Boyle, D. 2000. Specific glutamine and asparagine residues of γ S crystallin are resistant to *in vivo* deamidation. *J. Biol. Chem.* **275**: 26109-26112.
- Takemoto, L., Fujii, N., and Boyle, D. 1998. The possible role of alpha-crystallins in human senile cataractogenesis. *Int. J. Biol. Macromol.* **22**: 331-337.
- Tanksale, A., Ghatge, M., and Deshpande, V. 2002. Alpha-crystallin binds to the aggregation-prone molten-globule state of alkaline protease: Implications for preventing irreversible thermal denaturation. *Protein Sci.* **11**: 1720-1728.
- Tardieu, A., Veretout, F. Krop, B., and Slingsby, C. 1992. Protein interactions in the calf eye lens: interactions between β -crystallins are repulsive whereas in γ -crystallins they are attractive. *Eur. Biophys. J.* **21**: 1-21.
- Taylor, B.M., Sarver, R.W., Fici, G., Poorman, R.A., Lutzke, B.S., Molinari, A., Kawabe, T., Kappenman, K., Buhl, A.E., and Epps, D.E. 2003. Spontaneous aggregation and cytotoxicity of the beta-amyloid Abeta1-40: a kinetic model. *J. Protein Chem.* **22**: 31-40.
- Taylor, H.R., West, S.K., Rosenthal, F.S., Munoz, B., Newland, H.S., Abbey, H., and Emmet, E.A. 1988. Effects of ultraviolet radiation on cataract formation. *N. Engl. J. Med.* **319**: 1429-1433.
- Varley, P., Gronenborn, A.M., Christensen, H., Wingfield, P.T., Pain, R.H., and Clore, G.M. 1993. Kinetics of folding of the all β -sheet protein Interleukin-1 β . *Science* **260**: 1110-1113.
- Venter, J.C., et al. 2001. The sequence of the human genome. *Science* **291**:1304-51.

- Volles, M.J. and Lansbury, P.T.Jr. 2002. Vesicle permeabilization by protofibrillar alpha-synuclein is sensitive to Parkinson's disease-linked mutations and occurs by a pore-like mechanism. *Biochemistry* **41**: 4595-4602.
- Volles, M.J. and Lansbury, P.T. Jr. 2003. Zeroing in on the pathogenic form of α -synuclein and its mechanism of neurotoxicity in Parkinson's disease. *Biochemistry* **42**: 7871-7878.
- Volles, M.J., Lee, S.J., Rochet, J.C., Shtilerman, M.D., Ding, T.T., Kessler, J.C., and Lansbury, P.T. Jr. 2001. Vesicle permeabilization by protofibrillar alpha-synuclein: implications for pathogenesis and treatment of Parkinson's disease. *Biochemistry* **40**: 7812-7819.
- Weinreb, O., Van Rijk, A.F., Dovrat, A., and Bloemendal, H. 2000. *In vitro* filament-like formation upon interaction between lens α -crystallin and β L-crystallin promoted by stress. *Invest. Ophthalmol. Vis. Sci.* **41**: 3893-3897.
- Wenk, M., Baumgartner, R., Holak, T.A., Huber, R., Jaenicke, R., and Mayr, E.M. 1999. The domains of Protein S from *Myxococcus xanthus*: structure stability and interactions. *J. Mol. Biol.* **286**: 1533-1545.
- Wenk, M., Herbst, R., Hoeger, D., Kretschmar, M., Lubsen, N. H., and Jaenicke, R. 2000. Gamma S-crystallin of bovine and human eye lens: Solution structure, stability and folding of the intact two-domain protein and its separate domains. *Biophys. Chem.* **86**: 95-108.
- Wenk, M. and Mayr, E.M. 1998. *Myxococcus xanthus* spore coat protein S, a stress-induced member of the $\beta\gamma$ -crystallinn superfamily, gains stability from binding of calcium ions. *Eur. J. Biochem.* **255**: 604-610.
- Wetaufer, D.B. 1973. Nucleation, rapid folding, and globular intrachain interaction in proteins. *Proc. Natl. Acad. Sci.* **70**: 697-701.
- Wetzel, R. 1994. Mutations and off-pathway aggregation of proteins. *Trends Biotechnol.* **12**: 193-198.
- Wetzel, R. and Chrynk, B.A. 1994. Inclusion body formation by interleukin-1 beta depends on the thermal sensitivity of a folding intermediate. *FEBS Lett.* **350**: 245-248.
- Wieligmann, K., Mayr, E. M., Jaenicke, R. 1999. Folding and self-assembly of the domains of β B2-crystallin from rat eye lens. *J. Mol. Biol.* **286**: 989-994.
- Wistow, G., Turnell, B., Summers, L., Slingsby, C., Moss, D., Miller, L., Lindley, P., and Blundell, T. 1983. X-ray analysis of the eye lens protein γ -II crystallin at 1.9 Å resolution. *J. Mol. Biol.* **170**: 175-202.

Wistow, G., Summers, L., and Bundell, T. 1985. *Myxococcus xanthus* spore coat protein S may have a similar structure to vertebrate lens $\beta\gamma$ -crystallins. *Nature* **315**: 771-773.

Zhang, H., Kaneko, K., Nguyen, J.T., Livshits, T.L., Baldwin, M.L., Cohen, F.E., James, T.L., and Prusiner, S.B. 1995. Conformational transitions in peptides containing two putative α -helices of the prion protein. *J. Mol. Biol.* **250**: 514-526.

Zhang Z., Smith D.L., and Smith J.B. 2003. Human beta-crystallins modified by backbone cleavage, deamidation and oxidation are prone to associate. *Exp. Eye Res.* **77**: 259-72.

Zigman, S., Schultz, J., and Yulo, T. 1973. Possible roles of near ultraviolet light in the cataractous process. *Exp. Eye Res.* **15**: 201.

CHAPTER IX: APPENDIX

APPENDIX A: PROTEIN PARAMETERS

Protein parameters and constants calculated for human γ D crystallin and human γ S crystallin (calculated using ProtParam tool from <http://us.expasy.org/cgi-bin/protparam>).

Protein	Molecular Weight (Da)	E280 (1/nm)	Theoretical pI
WT HyD-Crys	20,634	41.04	8.25
WT* HyD-Crys	21,844	41.04	8.53
W42-only WT* HyD-Crys	21,727	23.97	8.53
W68-only WT* HyD-Crys	21,727	23.97	8.53
W130-only WT* HyD-Crys	21,727	23.97	8.53
W156-only WT* HyD-Crys	21,727	23.97	8.53
W42A WT* HyD-Crys	21,702	35.35	8.53
W68A WT* HyD-Crys	21,702	35.35	8.53
W130A WT* HyD-Crys	21,702	35.35	8.53
W156A WT* HyD-Crys	21,702	35.35	8.53
WT HyS-Crys	21,007	41.04	6.44
N143D HyS-Crys	21,007	41.04	6.11

APPENDIX B: MUTAGENIC PRIMER TABLE

Mutagenic primers utilized in site directed mutagenesis of human γ D crystallin to make triple tryptophan to phenylalanine proteins.

Mutation	Primers	T _m
W42F coding	5'-gcg tgg aca gcg gct gct tta tgc tct atg agc agc-3'	87°C
W42F noncoding	5'-gct gct cat aga gca taa agc agc cgc tgt cca cgc-3'	87°C
W68F coding	5'-c gac cac cag ttt atg ggc ctc agc gac tgc-3'	87°C
W68F noncoding	5'-cga gtc gct gag gcc cat aaa ctg ctg gtg gtc g-3'	87°C
W130F coding	5'-gct gga ggg ctc ctt tgt cct cta cga gct gtc c-3'	87°C
W130F noncoding	5'-g gac agc tag agg aca aag gag ccc tcc agc-3'	87°C
W156F coding	5'-ggc gct acc agg act ttg ggg cca cga atg cc-3'	87°C
W156F noncoding	5'-ggc att cgt ggc ccc aaa gtc ctg gta gcg cc-3'	87°C

APPENDIX C: UNFOLDING AND REFOLDING EQUILIBRIUM DATA ANALYSIS

I. Calculate the concentration of denaturant in each sample using the refractive index and the following equations:

$$[Urea]_A = 117.66 \cdot \Delta N + 29.753 \cdot \Delta N^2 + 185.56 \cdot \Delta N^3$$

$$[GdmCl]_A = 57.147 \cdot \Delta N + 38.68 \cdot \Delta N^2 - 91.60 \cdot \Delta N^3$$

II. Analyze samples using the 2-state folding curve analysis method described by Pace et al. (1989). The fluorescence light intensity at 350 nm or of 360 nm/320 nm was plotted versus the denaturant concentration.

A. The native and denatured baselines were calculated.

B. The fraction of protein unfolded (f_U) was calculated for each point in the curve using the equation:

$$f_U = \frac{(y_F - y)}{(y_F - y_U)}$$

In this equation, y_F was the value of y determined by the calculated fraction folded baseline, y was the observed light intensity at any given point, and y_U was the value of y determined by the calculated fraction unfolded baseline.

III. The equilibrium constant (K) and the free energy change (ΔG) was then calculated for each point in the curve using the following equations:

$$K = \frac{f_U}{(1 - f_U)}$$

$$\Delta G = -RT \ln K$$

In these equations, f_U was the fraction of denatured protein present in any given sample, R was the gas constant (1.987 calories/deg x mol) and T was the absolute temperature.

IV. The resulting data were analyzed using the method described by Fersht (1999). The light intensity of 360 nm/320 nm was plotted versus the concentration of denaturant and then fit using the following equations:

A. To calculate the fraction unfolded curve fit:

$$f_U = \frac{(f_U + f_N) \cdot \exp\left(\frac{\{m_{D-N} \cdot [Urea]_A - \Delta G_{H_2O}\}}{RT}\right)}{1 + \exp\left(\frac{\{m_{D-N} \cdot [Urea]_A - \Delta G_{H_2O}\}}{RT}\right)}$$

B. To analyze the raw data:

$$F = \frac{(a_N + b_N \cdot [Urea]_A) + (a_D + b_D \cdot [Urea]_A) \cdot \exp\left(m_{D-N} \cdot \frac{([Urea]_A - [Urea]_{1/2})}{RT}\right)}{1 + \exp\left(m_{D-N} \cdot \frac{([Urea]_A - [Urea]_{1/2})}{RT}\right)}$$

In these equations, ΔG is the Gibb's free energy of the protein at 0 M denaturant, f_u was the fraction unfolded baseline, f_n was the native baseline, a_N was the intercept of the fraction folded baseline, b_N was the slope of the fraction folded baseline, a_D was the intercept of the fraction unfolded baseline, b_D was the slope of the fraction unfolded baseline, $[Urea]_A$ was the actual urea concentration calculated by the refractive index in any given sample, m_{D-N} was the overall resistance of the protein to solvent denaturation

(cal/mol x M), and $[Urea]_{1/2}$ was the concentration of urea when half of the protein has been denatured (M).

For Kaledagraph curve fitting of fraction unfolded:

$$(1 * (\exp((m1 * m0 - m2) / 0.8))) / (1 + (\exp((m1 * m0 - m2) / 0.8)))$$

m0 = concentration of denaturant

m1 = m-value

m2 = ΔG_{H_2O}

For Kaledagraph curve fitting of raw data:

$$((m1 + m2 * m0) + (m3 + m4 * m0) * \exp(m5 * ((m0 - m6) / 0.8))) / (1 + \exp(m5 * ((m0 - m6) / 0.8)))$$

m0 = concentration of denaturant

m1 = intercept of native baseline

m2 = slope of native baseline

m3 = intercept of native baseline

m4 = slope of native baseline

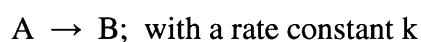
m5 = m-value

m6 = concentration of denaturant at transition midpoint

APPENDIX D: UNFOLDING AND REFOLDING KINETIC CURVE FIT CALCULATIONS

Data analysis of fluorescence changes during unfolding or refolding of human γ -crystallin proteins were calculated using the following parameters as described by Dr. Stephen Raso.

I. For the reaction having no intermediate:



$$[A] = A_0 e^{-kt}$$

$$[B] = A_0(1 - e^{-kt})$$

$$Y_{\text{obs}} = Y_A[A] + Y_B[B]$$

$$Y_{\text{obs}} = Y_A(A_0 e^{-kt}) + Y_B(A_0(1 - e^{-kt}))$$

Because A_0 and Y_A or Y_B are simply relative fluorescence differences, they may be combined:

$$Y_{\text{obs}} = Y_A'(e^{-kt}) + Y_B'(1 - e^{-kt})$$

$$Y_{\text{obs}} = Y_A'(e^{-kt}) + Y_B' - Y_B'(e^{-kt})$$

$$Y_{\text{obs}} = (Y_A' - Y_B')(e^{-kt}) + Y_B'$$

For Kaliedagraph curve fitting:

$$(m2 - m3) * (\exp(-m1 * m0)) + m3$$

$$m1 = k1$$

$$m2 = Y_A \text{ (fluorescence of state A)}$$

$$m3 = Y_B \text{ (fluorescence of state B)}$$

II. For the reaction having one sequential intermediate:



For Kaledagraph curve fitting:

$$m_3 \exp(-m_1 m_0) + [(m_4 (m_1 \exp(-m_1 m_0) - \exp(-m_2 m_0))) / (m_2 - m_1)] + (m_5) (1 + (m_2 \exp(-m_1 m_0) - m_1 \exp(-m_2 m_0)) / (m_1 - m_2))$$

$$m_1 = k_1$$

$$m_2 = k_2$$

$$m_3 = Y_A \text{ (fluorescence of state A)}$$

$$m_4 = Y_B \text{ (fluorescence of state B)}$$

$$m_5 = Y_C \text{ (fluorescence of state C)}$$

**APPENDIX E: OLIGOMERIC RING STRUCTURES OF HUMAN γ D
CRYSTALLIN**

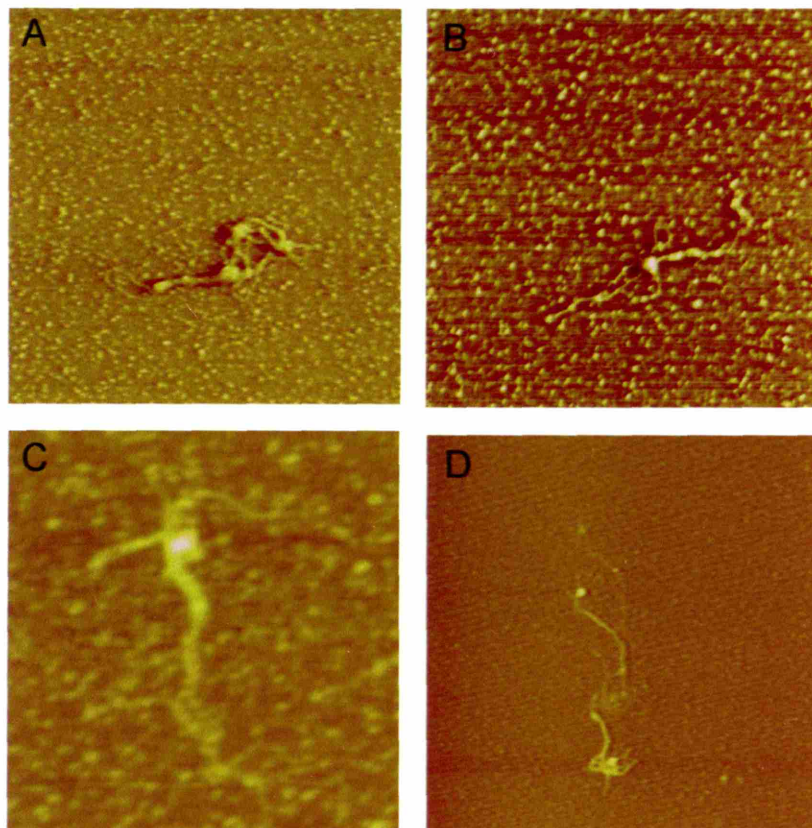


Figure 9-1

AFM height images of refolding HyD-Crys species absorbed to a mica surface. High surfaces are denoted as white (5 nm) and low surfaces are shown in black (0 nm). HyD-Crys was refolded for various times, extracted from solution, absorbed on mica, and observed using tapping mode AFM. Fibers of wild-type HyD-Crys refolded for 40 minutes are shown in all panels. Panel A-C are 1 μ m x 1 μ m scanning areas while D is a 2 μ m x 2 μ m scanning area.

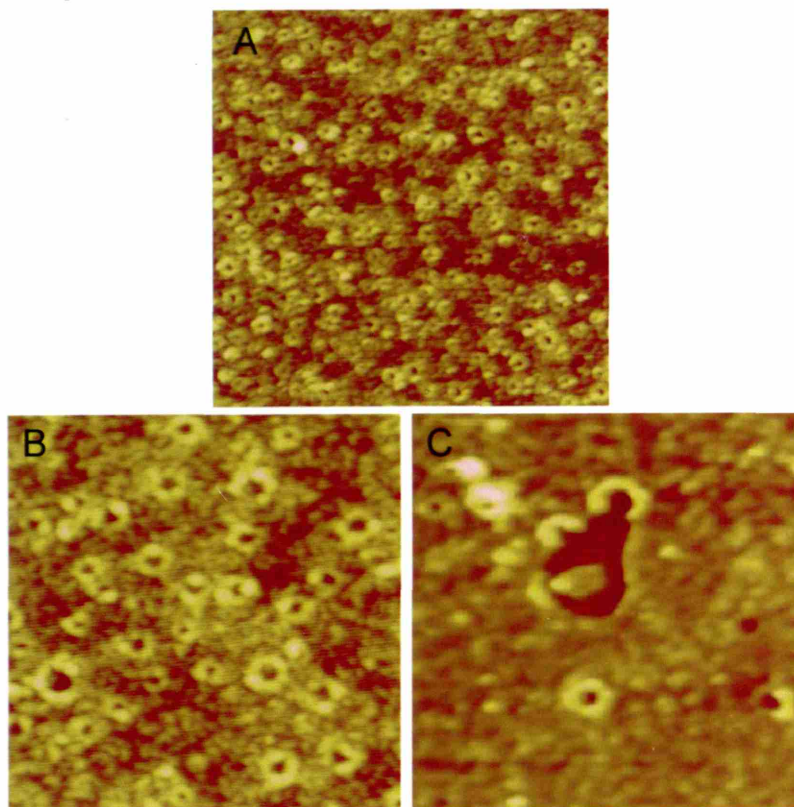
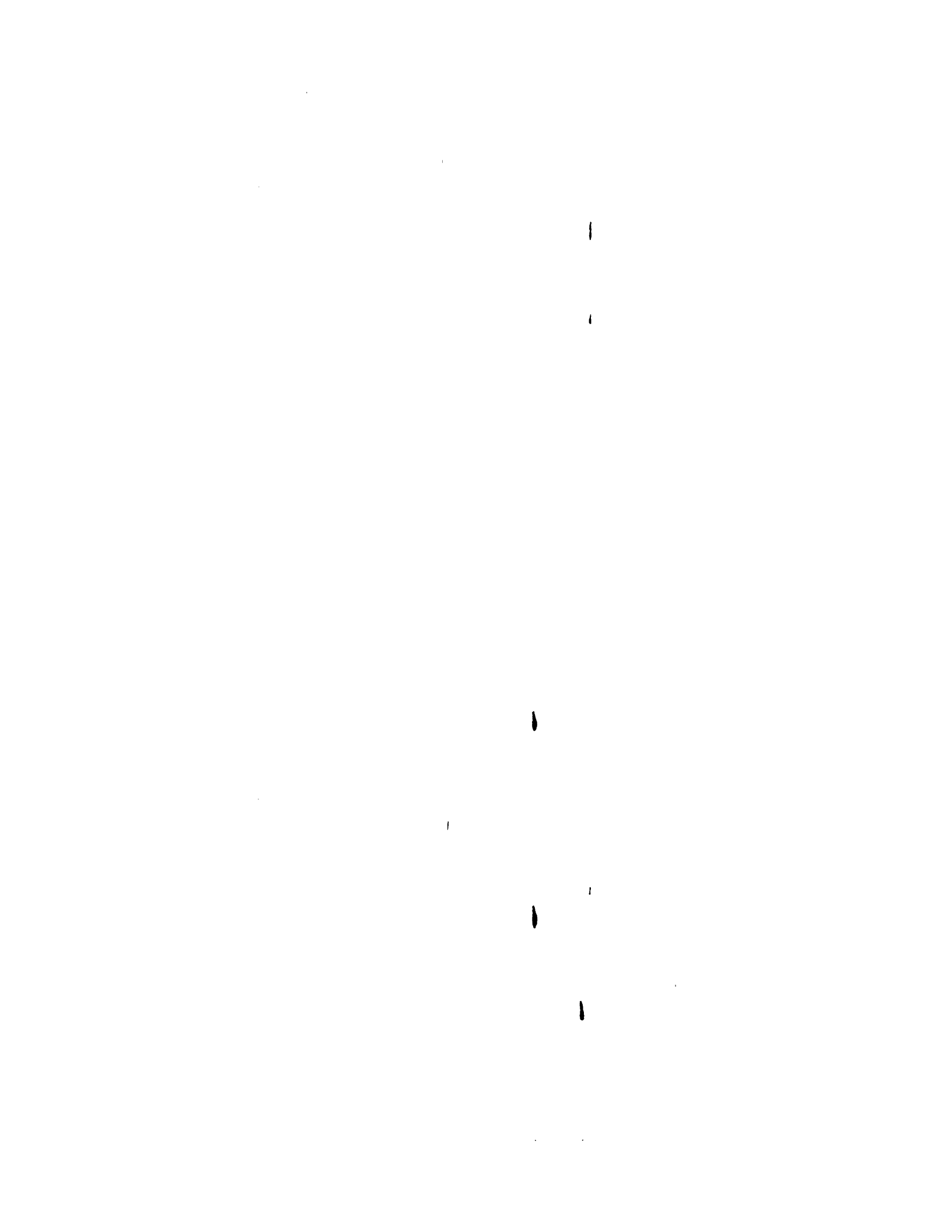


Figure 9-2

AFM height images of refolding his-tagged HyD-Crys species absorbed to a mica surface. High surfaces are denoted as white (5 nm) and low surfaces are shown in black (0 nm). HyD-Crys was refolded for various times, extracted from solution, absorbed on mica, and observed using tapping mode AFM. Fibers of wild-type HyD-Crys refolded for 40 minutes are shown in all panels A and B while panel C shows images collected after 30 minutes of refolding. Panel A shows protein in a 1 μm x 1 μm scanning area, panel B shows protein in a 400 nm x 400 nm scanning area, and panel C is a 2.5 μm x 2.5 μm scanning area.



Wild-type or His-tagged HyD-Crys was unfolded in 5.5 M GdnHCl 37°C. The protein was then refolded by dilution in 10 mM NaPO₄, 5 mM DTT, 1 mM EDTA, pH 7.0 at 37°C for various times. Ten µL of protein was spotted onto a cleaved mica surface. Samples were dried for two hours by evaporation in a high vacuum desiccator. Samples were then rinsed with 500 µL of milli-Q H₂O and dried for an additional two hours by desiccation. The height and phase of the samples were analyzed by atomic force microscopy using the tapping mode. Mica disks containing adhered protein were stored under high vacuum at room temperature.

Long fibrous species of wild-type HyD-Crys were observed after 40 minutes of refolding (Figure 9-1). The fiber length ranged from 300 to 700 nm. The fibers tended to be localized in clusters where multiple thin filaments had associated and wound around one another. In addition, the fibers had a central “core” from which the individual filaments seemed to emerge. It is unclear as to whether the core was a thickly bundled area of fibers or a nucleation center consisting of an alternate conformation of the molecule.

His-tagged HyD-Crys was also observed after 20 to 40 minutes of refolding. Instead of thin fibrils, however, circular oligomeric species were observed (Figure 9-2). All of the rings observed had areas of low height at their center that were presumed to be open cavities. The protein edge of all of the oligomers remained constant at approximately 25 nm despite large differences in overall diameter.

The ring structures generally fell into one of four diameter size ranges. The species were typically 60 nm, 80 nm, or 120 nm in diameter. A few oligomers, however, were larger than 120 nm and we have categorized them as a class of their own due to their size variability. The larger pores tended to create “puddles” absent of protein on the mica within their central cavities. Of the over two hundred rings measured, most fell into the 60 nm or 80 nm category. A more detailed analysis of the aggregates formed by wild-type HyD-Crys lacking the His-tag confirmed that circular oligomers were present amongst those aggregates as well.

“Did it ever get funny, because the part I saw was just weird.”

