Insights into the structure and aggregation of lens crystallins and other aggregation-prone proteins

Australian National University

Aidan Bradley Grosas

December 2019

A thesis submitted for the degree of Doctor of Philosophy of

The Australian National University

© Copyright by Aidan Bradley Grosas

All Rights Reserved

Memories made provide delight and growth,

Memories lost cause despondence and decay,

As twilight begets the shining dawn,

We strive to pass on the lessons learnt,

From those seen in pictures, present yet vacant,

To those just beginning to recognise faces,

One burgeoning, one slowly evanescing,

Ships passing in the night,

Both loved in their own right.

For Oma and Avery

DECLARATION

This thesis is composed of my original work and contains no material previously published or written by another person, except where due reference is given in the text. The content of this thesis is the result of work I have carried out since the commencement of my research degree candidature and has not been previously submitted for another degree or diploma in any university or tertiary institution.

Aidan B. Grosas
December 2019

ACKNOWLEDGEMENTS

I am profoundly lucky to do what I love on a daily basis. This dictum certainly does not mean my PhD experience has been devoid of difficult, frustrating, and at times, harrowing moments, quite the contrary. Despite this, the irrevocable truth is that I simply feel blessed to be in a position to wholeheartedly satisfy my curiosity and follow my scientific ideas to fruition. It is something I will never take for granted. That said, my PhD has easily been the most tortuous journey of my life. However, the journey is rarely a solitary one, even if it may feel like it at times. The many experiences I have had throughout this time have permeated my life in a multitude of different ways. Now, unravelling the individual parts is no small feat of nostalgia. Many of the experiences did seem, at first, to be innocuous, but introspection and hindsight has ultimately highlighted their worth. Below, I would like to offer my sincere thanks to those who have been apart of my journey and whose influence have left an indelible impact.

To my supervisor Professor John Carver, firstly, thank you for taking on a relatively inexperienced kid who had never before expressed, purified or characterised a protein. While it proved a significant barrier to early progress, it ultimately unveiled to me the world of protein science in a very fundamental and visceral way. Your overall character and demeanour have made for an ease of collaboration throughout my time in your lab and I have very much enjoyed our discussions, both scientific and otherwise, particularly since you ceased your role as director! It takes a certain type of supervisor to make their PhD student feel like an equal rather than a subordinate and this dynamic really meant a great deal to me. Thank you for giving me the space to grow as a researcher and for your ever-stalwart professionalism peppered with some great humour.

To Dr David ‘Davey Daveo’ Thorn, I know my time in the Carver lab would have been quite different if you had not joined when you did. Having you around during my PhD really made coming to work fun, even if we were the only ones who knew why we were laughing most of the time! Despite the frivolity (or perhaps because of it), our conversations about techniques, results, and papers over lunch did more for furthering my scientific appreciation and understanding than any other

undertaking during my PhD. You are an excellent scientist and person who I now consider not only as a valued colleague, but as a brother. Truly, thanks for being you, Dave.

I have had the pleasure to work with and alongside many other excellent colleagues. To Dr Elmira Bahraminejad, thank you for being my PhD buddy and a quality person. It has been an absolute pleasure to share this experience with you and I am tremendously proud of your accomplishments. To Dr Joel Thevarajah, today is a far cry from our undergrad days in Western Sydney where I thought your hat was glued on your head! I certainly treasure our friendship (stirring and all) and I am glad we have had each other's back the entire time. Sincere thanks also to both Dr Nick Ray and Dr Manjeet Kumar who both had a hand in getting me off the ground at the beginning of my PhD. To my collaborators Dr Agata Rekas, Dr Jitendra Mata, Mar-dean Du Plessis, Dr Colin Jackson, Dr Peter Mabbitt, and thesis advisors Professor Gottfried Otting, and Professor Thomas Huber for their advice, scientific dialogue and experimental contributions throughout my candidacy. Thank you to the Carver lab members both past and present for their efforts in maintaining the lab and a collegial atmosphere. Thanks also to Professor Thomas Millar and Dr Christopher Jones for their interest in my progress and their counsel, particularly in the early days.

Many people outside the academic sphere deserve my humble gratitude for their role in my life while I have been under a PhD shroud. Firstly, to Win (Mumsie), Wing (Wing Wing), and Vinuri (Vin Vin), whose help, within the last year particularly, has been instrumental in allowing me to complete my PhD. The love and care you show Avery, Jen and I fills me with happiness and comfort. I cannot overstate how terribly grateful I am for all you do for us and I look forward to our future endeavours. My love to you all. To all my friends who I left back in Sydney, Nick, Andrew, Brent, Matthew, Anthony M and Anthony A, I miss you guys day to day. Thank you for the fun times when I returned for a visit and for coming to visit me on occasion, these were always highlights. I am looking forward to catching up some more and as always, I look forward to having you visit us! Also to Justin, thanks a lot for the company and gym sessions during my time in Canberra.

To my grandparents. Nanna, I look forward to sharing my graduation with you while also thinking of Poppy, who I know would be very proud. Pa, thank you for your unyielding interest in my 'scholastic' endeavours. While I know it is sometimes difficult to understand what I have been doing,

you have always provided me with uncompromising support. Given how long you have been waiting for me to complete my PhD, sharing the moment with you will no doubt be one of my fondest memories. Finally to my Oma, who I miss terribly. I know you would be overjoyed with our current lives, great grandkids and many other accomplishments to revel in, but your ability to appreciate it has been cruelly stripped from you. Despite my affinity for science, you are the reason I am doing the work I do. In the future, I will continue to strive to contribute towards solving neurodegenerative diseases in your loving memory.

To my family, Susan, Darren, and Ashley, thank you for constant support, passion, and unending love. Moving away from you all was one of the hardest things I have ever done. While there have been joyous events in the meantime, for instance becoming an uncle to beautiful Scarlett, I long to spend more time with you all now this process is complete. Thank you for being so supportive of me taking a, shall we say, less traditional route to finding a job. Your financial and emotional support have sustained me on many occasions and I will never forget this. Since I moved, and with due acknowledgement that I am indeed not too far away, I have continued to feel like my life is missing that one little thing that only you can provide. I really miss you all so much. While I am now blessed with my own little family that brings me so much joy, you are always in my thoughts. I really hope I have made you proud. I love you all so much.

To my love, Jen, where would I be without you? ‘In Sydney’, you would say – And you would probably be right! Never could I have imagined I would have met someone, seated just down the hall from me, who would end up forming the cornerstone of my life as I know it. You are brilliant, funny, and caring, and I am in awe of your ability to excel at anything you feel is worth your attention. You have propped me up, emotionally, financially, and physically more than anyone else outside of my family. The love, care, and gratitude I have for you is insurmountable. Thank you for combating my inane imposer syndrome tendencies, for looking after me when I’m down, for filling my plate, and for believing in me. You are my great advocate and I utterly cherish you. Avery and I are so very lucky to have you and I cannot wait to see what the future holds by your side. I love you, Jen.

Finally, to my bubba, my little sunshine, my Avery. Soon you will turn one, and I am still in disbelief! When I began my studies, I never would have guessed you would be around when I finished, and yet I now cannot imagine my life without you. While undertaking my PhD, I use to think of how proud I would be to have created my final thesis, my best and most creative work to date, a level of accomplishment I might never again achieve. Yet now, considering this idea ultimately seems farcical, and I cannot entertain it for a second. This is, quite simply, because nothing I could create will ever surpass what I see in you. Thank you for lighting up my life in ways I never considered possible. Your Daddy loves you more than you will ever know.

ABSTRACT

Cataract is the world's leading cause of blindness. The destabilisation, partial unfolding, and aggregation of lens crystallin proteins cause the loss of lens transparency (opacification) and cataract formation. Numerous congenital mutations and age-related changes to the long-lived α-, β- and γ-crystallins are associated with cataract and their study has provided insight into the molecular basis of this disease. In this thesis, α- and γ-crystallin isoforms have been characterised under crowded and oxidative conditions, respectively. In addition, the conformational heterogeneity of model proteins was studied by capillary electrophoresis as a prelude to such studies on the more complex crystallins.

Chapter 2 details the structural characterisation of the disulfide-linked γS-crystallin dimer, an oxidative product in the aging lens. X-ray crystallography revealed an intermolecular disulfide bond from C24-C24' and two intramolecular disulfides, one in each subunit, between C22 and C26. Small angle X-ray scattering confirmed the extended in-solution biological assembly in lieu of a compact state. It was demonstrated that the disulfide-linked dimer was stable at glutathione concentrations akin to those in aged and cataractous lenses. The dimer had a higher aggregation propensity compared to the monomeric form owing to uncooperative domain unfolding. These findings provide novel insight into the contributions of oxidative modification to the formation of age-related cataract. Finally, similarities noted upon comparison with cataract-associated mutants suggest both congenital and age-related forms may have similar developmental pathways.

Chapter 3 describes the impacts that a highly crowded environment comparable to the eye lens has on the structure and function of the molecular chaperone αB-crystallin. Macromolecular crowding using Ficoll 400 induces significant destabilisation, unfolding, an increase in size/oligomeric state, and a loss of chaperone function leading to kinetically distinct amorphous and fibrillar aggregation. These results are recapitulated in-principle using the biologically relevant crowding agent bovine γ-crystallin. Aggregation of αB-crystallin is prevented by its lens partner protein αA-crystallin at physiologically relevant ratios through an increase in the αA/αB-crystallin complex stability. These results complement multiple dilute *in vitro* and *in vivo* studies, providing a mechanistic basis for their findings and also support therapeutic approaches that prevent and reverse cataract via α-crystallin stabilisation.

Chapter 4 investigates capillary electrophoresis as a method for studying the conformational heterogeneity of a protein. Bovine serum albumin (BSA), yeast alcohol dehydrogenase (YADH), and bovine α-lactalbumin (BLA) were used to assess the application of this method towards various conformational aspects in comparison to size-exclusion chromatography coupled to multiple angle light scattering. The method distinguished between BSA oligomers and two different monomer populations, multiple YADH monomer and tetramer conformations, and apo- and holo-BLA. The 'dispersity of electrophoretic mobilities' allowed a relative comparison of the levels of conformational heterogeneity between unrelated proteins. Novel structural findings were made despite the extensive previous characterisation of these model proteins. This study allows for better interpretation of the heterogeneity of more complex proteins such as post-translationally modified crystallins from *in vivo* sources and oligomeric α-crystallin.

Overall, this thesis provides new insights into the impacts of post-translational and environmental changes to α- and γ-crystallin and provides a molecular basis for their aggregation and contribution to cataract. There is strong analogy to α- and γ-crystallin congenital mutants that exhibit similar physical and chemical changes noted in this thesis, an observation that is potentially of great benefit toward the elucidation of common structural deficiencies leading to crystallin aggregation. Lastly, capillary electrophoresis allows for a new view of protein conformational heterogeneity that may provide novel future insights into crystallin structure and function.

TABLE OF CONTENTS

CHAPTER 1:

INTRODUCTION

"The secret of getting ahead is getting started."

– Mark Twain

DECLARATION

The following literature review (Introduction) has formed the basis of a chapter in the upcoming book *Long-lived Proteins in Human Aging and Disease (Roger J.W. Truscott, ed.)*.

Grosas A.B. and Carver J.A. Eye lens crystallins: remarkable long-lived proteins in Long-lived Proteins in Human Aging and Disease (Roger J.W. Truscott, ed.), Weinham, Germany, Wiley-VCH, 2020.

The following literature review was researched and written by the author under the supervision of Professor John Carver.

Protein Folding, Misfolding, Aggregation and Molecular Chaperones

Protein sequence and structure are inextricably linked to molecular function and a protein's physical properties. Over 50 years of theoretical and experimental study has produced multiple theories to be ratified and problems to be overcome in the quest to understand how a protein folds (1, 2). Seminal work by Anfinsen and colleagues yielded a view that protein folding occurred because the native state is a protein's most thermodynamically stable conformation and that all the information needed for a protein to fold was inherent to its primary sequence (3, 4). Soon thereafter, Levinthal and colleagues highlighted an issue regarding the kinetics of protein folding. They indicated that due to the high degree of freedom that an unfolded polypeptide chain can exhibit, the number of possible conformations to be sampled means it would be impossible for a protein to fold on a biologically relevant timescale via random sampling (5, 6). This led to the postulation that proteins must fold via a specific and sequential folding pathway or a few pathways in order to restrict the kinetic sampling issue, thereby satisfying Levinthal's so called 'paradox' (7, 8). This process can be represented in a two-dimensional schematic whereby protein folding occurs on the on-folding pathway, usually via a high energy intermediate or intermediates, which ultimately leads to a low energy native protein (Figure 1. A, *left*).

However, advances in theory, simulation and experimental methodology have invariably built upon these formative ideas to suggest that macroscopic states, such as the native or unfolded state, should be viewed as a distribution or ensemble of many conformations rather than a single state (1, 9, 10). This 'new view' of protein folding has a strong basis in statistical mechanics and can be represented as a 'folding funnel' or 'free-energy landscape' (Figure 1. B, *left*) (11-13). This model does not ascribe to the idea of a finite folding pathway. Rather, it suggests that events leading to protein folding such as hydrophobic collapse, hydrogen bond and salt-bridge formation can occur in parallel for different protein molecules and in an order that is governed by the initial stochastic arrangement of the unfolded state (11, 14, 15). That is, some unfolded states might favour hydrophobic collapse first, then the formation of salt-bridges, or vice versa, depending on what is more favourable for the unfolded chain's conformation at the time folding is initiated. Ultimately, these enthalpy-based interactions trend toward the most favourable energy state as they continue down the folding funnel, fulfilling Anfinsen's

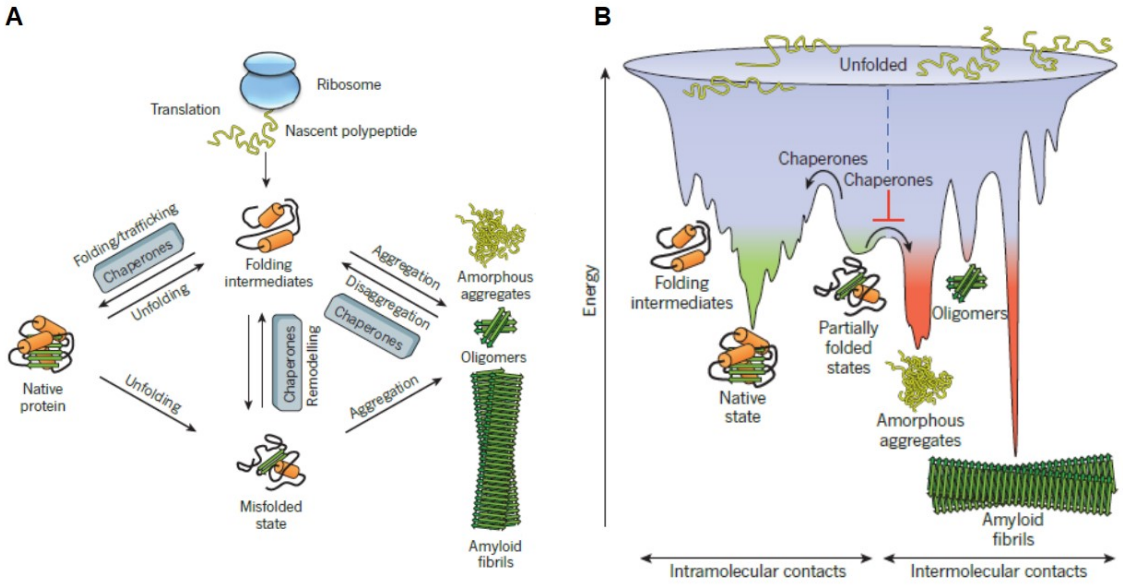


Figure 1. Protein folding, unfolding and aggregation. A. schematic of the protein on-folding pathway (*left*) and the off-folding pathway (*right*) showing the formation native protein and amorphous aggregates or amyloid fibrils from one or more folding intermediates. The off-folding pathway can also be reached via a misfolded state. The junctures where chaperone interaction assists protein folding or disaggregation is indicated. **B.** a depiction of the free-energy landscape that proteins explore upon folding and unfolding, via intramolecular contacts (*left*), and aggregation, via intermolecular contacts (*right*). The high free-energy peaks at which chaperones assist in refolding or discourage aggregation are indicated. Modified from Hartl *et al.*, 2011 (16).

criterion. Because of these newly formed interactions, the conformational entropy is diminished thereby limiting the polypeptide chain's available conformational search options and speeding up folding, thereby explaining Levinthal's paradox (9, 15, 17). As the protein navigates this free energy landscape, there are peaks and troughs leading to kinetic traps which are ultimately overcome by a gain of conformational entropy, i.e. partial unfolding, before refolding continues toward the global energy minimum, i.e. the native state (15, 18, 19).

The ruggedness of this landscape, i.e. the number of troughs between an unfolded protein and its native state, can dictate how many intermediates are formed on the folding journey and their longevity due to kinetically trapping (19-21). Partially folded state(s), formed as a folding intermediate or via aberrations in a protein's sequence, can provide access to additional conformational states which largely involves intermolecular contacts (16, 22, 23) (Figure 1. A and B, *right*). Amorphous aggregates and amyloid fibrils are protein conformations that are usually inherently more stable than the native

state, hence nature's desire to form them. Their formation, properties and structures can produce deleterious consequences for the cell and they are often associated with disease (*see below*). The formation of amyloid fibrillar aggregates is often preceded by high-energy soluble oligomers that can be considered as kinetically trapped intermediates between misfolded protein and amyloid fibrils (16, 22) and has been implicated as a species of high cytotoxicity (24, 25) (Figure 1. B, *right*).

While the protein folding energy landscape aims to funnel a protein from its unfolded to its native state, some kinetic traps may be difficult to overcome and thus assistance is provided by molecular chaperones (15, 16). This class of proteins interacts with, stabilizes or helps another protein to reach its native conformation but is not incorporated into the final conformation (16). Molecular chaperones involved in protein folding are generally ATP-dependent, such as heat-shock protein 70 (Hsp70) (26, 27) while those involved in preventing protein unfolding and aggregation are generally ATP-independent, such as the small heat-shock proteins (sHsps) like αB-crystallin (28, 29). Within the energy landscape, these molecular chaperones buffer the intramolecular contacts of kinetically trapped non-native conformations, be that folding intermediates or partially folded states, and assist with the necessary structural rearrangements needed to overcome the energy hurdle and continue folding to the native state (18, 30). Similarly, their role in preventing the formation of intermolecular contacts associated with the formation of aggregate states means moving a partially folded protein state from the off-pathway back to the on-pathway or, indeed, discouraging the transition to the off-folding pathway altogether (16, 22). The general mechanism of action for some ATP-dependent molecular chaperones, such as Hsp70, has been largely elucidated (31, 32). The ATP-independent sHsps are considered to interact with an abundance of surface-exposed hydrophobic residues that are yet to fully collapse in the partially folded or misfolded state however, the specifics of the interaction are yet to be fully elucidated (33, 34). In order to understand both the toxicity of protein aggregates and how molecular chaperones can interact with them, knowledge of protein aggregate formation, structures and cytotoxicity is prudent.

The Structure and Properties of Amyloid Fibrils and Amorphous Aggregates

The structure and properties of amyloid fibrils have been studied for decades however, it is only very recently, with the advent and improvement of solid-state NMR (35) and cryo-electron microscopic (36) techniques, that their structures have been elucidated at atomic level resolution (37). For instance, the structure of amyloid fibrils formed from amyloid-β (Aβ) 1-42 peptide, which is intimately associated with the extracellular plaque deposits in Alzheimer's disease, have been solved at atomic level resolution (38, 39) (Figure 2). An amyloid fibril's defining structural feature, as present in all amyloid fibrils no matter the precursor's primary sequence, is the cross-β-strand fold stacked perpendicular to the fibril axis with a distance of 4.7-4.8 Å between each strand (40, 41) (Figure 2. A). The Aβ(1-42) amyloid fibril consists of two twisted protofilaments (38, 39) (Figure 2. B). However, this is not the case for all amyloid fibrils such as those formed from α-synuclein (the putative causative agent in Parkinson's disease) which presents as a single protofilament (42). Steric zippers, which are regions of hydrophobic side chains tightly packed and devoid of bonded or intercalated water molecules, are thought to be unique to amyloid fibrils (43, 44). Steric zippers stabilize both intra- and inter-protofilament structure and the latter is also strengthened by salt bridges (Figure 2. A), although these features can differ between fibrils. Polymorphism in amyloid fibrils is becoming more evident with every structure solved and is even present when forming fibrils from the same protein (37, 45). The Aβ(1-42) amyloid fibril has been solved in two different conformations, the LS- (39) and S-shaped (38) (Figure 2. C, *left* and *right*, respectively), due to varied polypeptide chain packing likely induced by different solution conditions of formation. Potentially, the polymorphism among fibril structures for the same peptide or protein may manifest itself in different symptoms in the clinical context (37, 46).

Due to their architecture, amyloid fibrils resemble rod shaped structures which have the capacity to scatter visible light, albeit to a lesser extent than spherical aggregates (47, 48). The repeating cross-β sheet motif accommodates amyloid sensitive dyes that become highly fluorescent or exhibit birefringence in their presence (49, 50). Amyloid fibrils exhibit strength comparable to that of steel, a mechanical stiffness comparable to silk and varying levels of flexibility (51, 52), making the non-pathogenic forms attractive for the development of new lightweight materials.

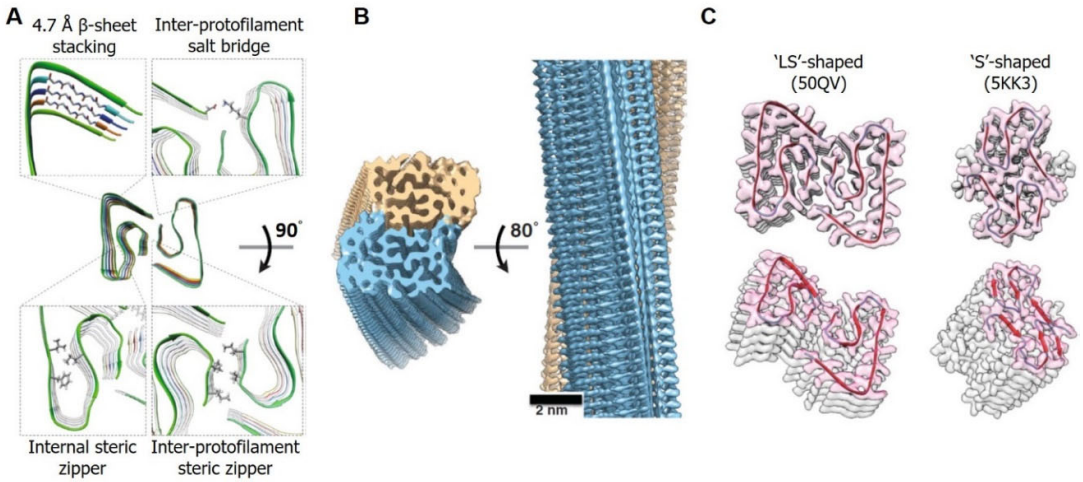


Figure 2. The structure of the amyloid-β (Aβ) amyloid fibril. A. The prominent structural features of an amyloid fibril including the β-sheet stacking/spacing, inter-protofilament salt bridges and intra- and inter-protofilament steric zippers of the Aβ(1-42) fibril (PDB: 50QV) as determined by cryo-electron microscopy (39). **B.** A cross-sectional (*left*) and side-on (*right*) view of the 3D reconstruction of the Aβ(1-42) fibril with the two protofilaments coloured brown and blue. **C.** Space-filling representations of two different Aβ(1-42) fibril models showing the inter-protofilament packing arrangement described as a 'LS-shaped' (*left*) (39) and a 'S-shaped' (*right*) (38) fibril. Modified from Gremer *et al.*, 2017 (39) and Iadanza *et al.*, 2018 (37).

By contrast, an amorphous protein aggregate consists of largely disordered clumps of unfolded or partially unfolded protein with no long-range order (53, 54). There are no atomic level resolution structures of amorphous protein aggregates due to their disordered nature and as such, compared to amyloid fibrils, as there is limited structural insight available (55, 56). However, these aggregates can be visualised as granular-like structures by techniques such as transmission electron microscopy (TEM) and atomic force microscopy (AFM) (56, 57) (Figure 3). Under the right conditions, most proteins can form amorphous aggregates including those that are prone to form amyloid fibrils, such as β_2-microglobulin which is responsible for dialysis-related amyloidosis (56, 58) (Figure 3). Despite the lack of atomic level resolution structures of amorphous aggregates, some insights at the atomic level have come from solid-state NMR. Eye lens γ-crystallin protein aggregates that appear amorphous-like via TEM have local order from solid-state NMR studies, suggesting that at least some amorphous aggregates may contain more order than originally thought (59-61). Generally, visible light is scattered far more effectively by amorphous aggregates relative to amyloid fibrils due to their spherical-like nature and larger cross-sectional area (47, 48).

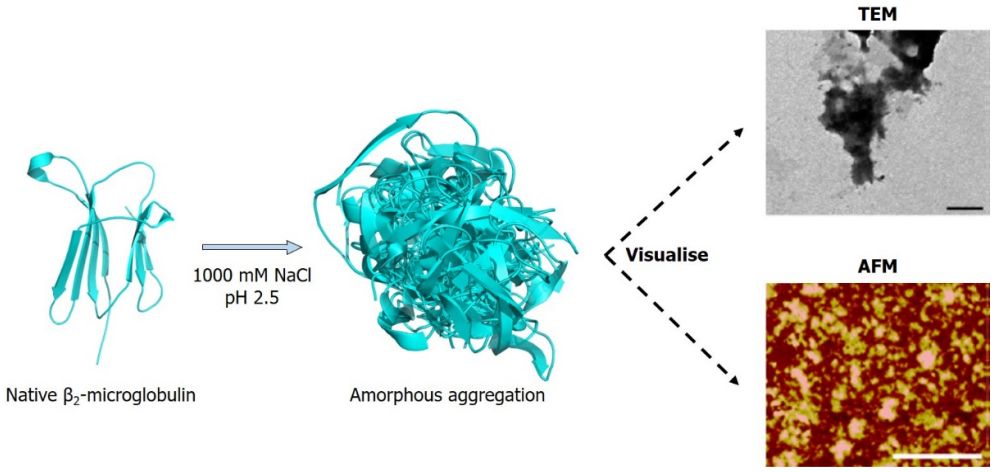


Figure 3. The formation, structure and visualisation of β2-microglobulin amorphous aggregates. The native structure of β2-microglobulin (PDB: 1LDS) from X-ray crystallography (62). Upon aggregation, this protein usually forms fibrils. However, under different conditions, e.g. high salt and low pH, it can undergo amorphous aggregation made up of multiple misfolded β2-microglobulin subunits and devoid of long-range order. The resultant aggregate can be visualised via TEM or AFM. It exhibits large clumps of protein without any definitive structure. Modified from Yoshimura *et al.*, 2012 (56).

Diseases Associated with Protein Aggregation

Investigation into the pathological nature of protein aggregates has been of particular interest for many years, particularly since the observation of plaques in the brain of dementia patients was made by Alois Alzheimer (37, 63). Currently, there are approximately 50 proteins or peptides associated with human disease that undergo aggregation (37), with approximately 6.5 % of deaths in Australia in 2017 being attributed to protein aggregation diseases (Table 1). However, in the majority of cases, the exact role these aggregates play in the etiology of the associated disease is not clear (64, 65). Some modes of cytotoxicity ascribed to protein aggregation diseases are the disruption of cellular and/or organelle membranes (66, 67), circumvention and inhibition of cellular degradation mechanisms including ubiquitination and autophagy processes (68-71), impairment of mitochondrial function (72, 73), sequestration of molecules essential to cellular function (74, 75), the production of reactive oxygen species (76, 77), and the interference of cellular transport mechanisms (78, 79). While the identification of these cytotoxic mechanisms associated with protein aggregation have provided great insight into the diseases' causes, it is still unclear what stage or stages of aggregate represents the most cytotoxic species. The challenge now lies in identifying the most cytotoxic aggregate species and elucidating the

Table 1: Selected protein aggregation diseases, their precursor and related percentage mortality in Australia (2017). Adapted from various sources (37, 80, 81).

Disease	Protein or peptide that exhibits aggregation	% of related deaths in Australia (2017)[a]
Alzheimer disease	Amyloid-β peptide	2.66[b]
Type II diabetes	Islet amyloid polypeptide (amylin)	1.49[c]
Parkinson disease	α-Synuclein	1.14[d]
Amyotrophic lateral sclerosis	Superoxide dismutase	0.41[e]
Spinal and bulbar muscular atrophy	Androgen receptor with polyQ expansion	0.37[f]
Dementia with Lewy bodies	α-Synuclein	
AL amyloidosis	Immunoglobulin light chains or fragments	0.11[g]
AH amyloidosis	Immunoglobulin heavy chains or fragments	
AA amyloidosis	Fragments of serum amyloid A protein	
Senile systemic amyloidosis	Wild-type transthyretin	
Familial amyloidotic polyneuropathy	Mutants of transthyretin	
Haemodialysis-related amyloidosis	β_2-microglobulin	
ApoAI amyloidosis	N-terminal fragments of ApoAI	
ApoAII amyloidosis	N-terminal fragments of ApoAII	
ApoAIV amyloidosis	N-terminal fragments of ApoAIV	
ApoCII amyloidosis	ApoCII	
ApoCIII amyloidosis	ApoCIII	
Lysozyme amyloidosis	Mutants of lysozyme	
Fibrinogen amyloidosis	Variants of fibrinogen α-chain	
Huntington	Huntingtin with polyQ expansion	0.07[h]
Spongiform encephalopathies	Prion protein or fragments thereof	0.04[i]
Frontotemporal dementia with Parkinsonism	Tau	0.01[j]
Spinocerebellar ataxias	Ataxins with polyQ expansion	0.01[k]
Cataract	Crystallins	0[l]

[a]As per the Australian Bureau of Statistics (ABS) 'Causes of Death, Australia, 2017' (82).
ABS cause of death and ICD-10 code for:
[b]Alzheimer disease (G30).
[c]non-insulin-dependent diabetes mellitus (E11).
[d]Parkinson disease (G20).
[e]spinal muscular atrophy and related syndromes (G12).
[f]other degenerative diseases of nervous system, not elsewhere classified (G31).
[g]amyloidosis (E85).
[h]Huntington disease (G10).
[i]slow virus infections of central nervous system (A81).
[j]secondary Parkinsonism (G21).
[k]hereditary ataxia (G11).
[l]senile cataract (H25).

mechanism by which they induce cytotoxicity and cell death.

Alzheimer's disease is associated with the aggregation of the Aβ peptide and has been thoroughly studied due to its prevalence among an aging population and its relatively high mortality rate (Table 1). The protein aggregates formed in Alzheimer's disease can be characterised as fibrillar, adopting the canonical features of an amyloid fibril (83) as previously described. In comparison, cataract is attributed to the aggregation of the crystallin proteins in the eye lens (84, 85) but, unlike Alzheimer's disease, cataract did not register a mortality rate in Australia in 2017 (Table 1). In cataract, crystallin aggregation is largely disordered and is characterised as amorphous in form (86), although a small amount of amyloid fibrillar aggregates has recently been identified in porcine and human cataract lenses (87, 88). While not directly contributing to mortality, cataract is highly correlated with aging (89, 90) and disproportionately affects disadvantaged individuals in isolated communities (91) and developing countries (92). An estimated 30 million people are blind due to catatract making it the leading cause of blindness in the world (93). While the role of protein aggregation in driving the cytotoxic mechanisms underlying diseases such as Alzheimer's is still unknown, the relationship is very clear for cataract as protein aggregation in the lens leads directly to light scattering which obscures vision (94). Therefore, an understanding of the aggregates formed in cataract is particularly important to elucidating their formation and subsequent prevention to decrease the incidence of blindness worldwide and reduce the economic burden of cataract.

The Structure and Function of the Eye Lens

The eye comprises several biological compartments used to facilitate its primary function as the organ of vision (95) (Figure 4. A). The cornea is situated on the front of the eye and both focusses visible light and filters potentially damaging ultra-violet (UV) light (96). The iris is made of coloured pigment and functions as a diaphragm to control the amount of light entering the eye (97). The lens refracts and focuses light via a mechanism termed accommodation to maintain a focal distance that allows images to be focused directly on to the retina, particularly from objects whose distance is varying (98). The photoreceptors in the retina capture the focused light which is translated into electro-chemical signals

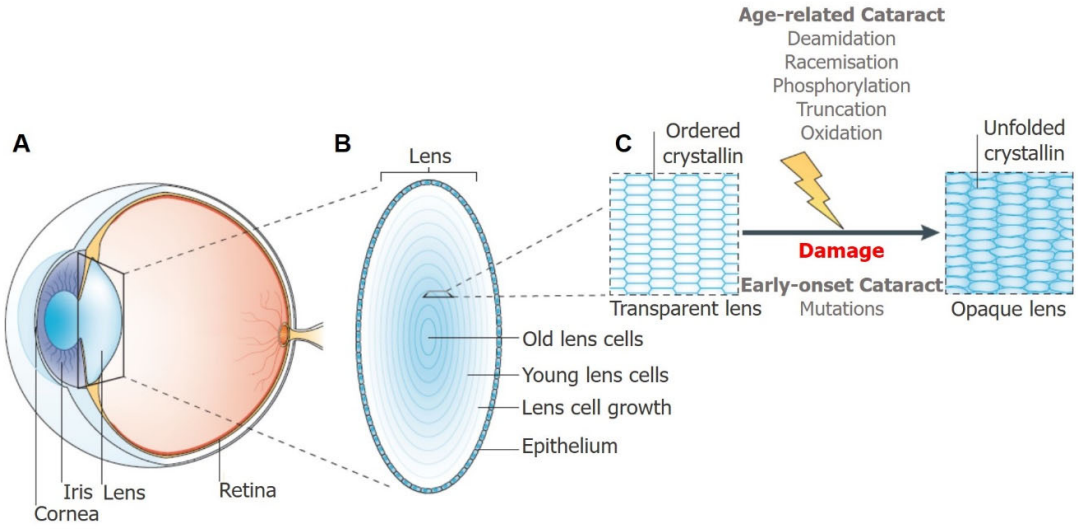


Figure 4. Eye lens structure and the development of cataract. A. A cross-section of the eye depicting its general structural features. **B.** A cross-section of the lens showing radial cell growth whereby old lens fibre cells are in and around the centre of the lens (the nucleus) while younger fibre cells are near the outside of the lens (the cortex) close to the epithelial layer that differentiate into new fibre cells. **C.** A schematic of lens fibre cell morphology and packing. Ordered crystallin in the transparent lens is damaged by mechanisms either associated with age-related or early-onset cataract to produce misfolded and aggregated crystallin in the opaque lens. Modified from Toyama and Hetzer, 2013 (99).

that are sent down the optic nerve and into the visual cortex of the brain where they are interpreted as an image (100).

Human lens organogenesis begins at approximately four weeks' gestation (101) and during embryonic development, two cell types form from surface ectoderm: epithelial and fibre cells (102). Morphogenetic changes lead to the differentiation of the surface ectoderm into epithelial cells that form a spheroid structure within the optic cup termed the lens vesicle (103, 104). The cells at the posterior of the lens vesicle elongate and form primary lens fibre cells that become the lens embryonic nucleus, while the cells at the anterior form a monolayer of simple cuboidal epithelial cells (84, 105). This structure, a compact spheroidal accretion of fibre cells surrounded by a layer of epithelial cells, forms the basis for lens growth. The epithelial cells at the anterior undergo mitosis and differentiation over a lifetime to form secondary lens fibre cells that are shuttled toward the embryonic nucleus and compact in concentric layers forming the foetal, juvenile and adult nucleus, and finally the lens cortex (102, 106, 107). Due to this process, the centre of the lens nucleus contains fibre cells that are as old as the

individual while the outer regions, including the adult nucleus and the cortex, contain progressively younger fibre cells out towards the epithelial layer (106, 108). As a result, a cellular age gradient forms in the lens from the centre to the periphery (Figure 4. B). The continued growth of the lens throughout life is noted at the macroscopic level as the human lens weighs approximately 65 mg at birth, grows rapidly in the first year of life and then slows significantly, before reaching about 250 mg at advanced age (109, 110).

Lens epithelial cells undergo specific changes during differentiation into lens fibre cells that are pivotal to the maintenance of lens transparency. Epithelial cell differentiation is characterised by two main events; the degradation and subsequent loss of cellular organelles and the large-scale synthesis of crystallin proteins that adopt a short-range array (103, 111, 112) (Figure 4. C). By preventing the scatter of visible light, these factors facilitate the high level of transparency exhibited by lens fibre cells (113, 114). Fibre cells are largely metabolically inactive, elongated, and consist of high concentrations of crystallin proteins (108). High lens crystallin concentrations: (i) provide the necessary refractive index to allow correct focusing of light onto the retina (94, 115), (ii) help to maintain the fibre cells structural integrity which is essential for tight cellular packing, (iii) contribute to maximising surface area between cells thereby limiting interstitial space, and (iv) allow the formation of efficient gap-junctions for cellular communication (106, 116, 117). Thus, any alteration in structure of the crystallins, e.g. unfolding or aggregation due to early-onset or age-related damage, has a significantly adverse effect on fibre cell transparency and morphology which can ultimately lead to the development of lens opacities and cataract (99, 104, 118) (Figure 4. C).

The Crystallin Proteins of the Eye Lens

The cytoplasm of lens fibre cells contains protein at concentrations of up to 300-400 mg/mL of which up to 90 % comprises crystallins (94, 119). There are three classes of human crystallins: α, β, and γ, each with numerous isotypes (115). The three classes have distinct differences in mass (oligomeric state), size, isoelectric point, and electrophoretic mobility (89, 120) and can be separated from lens cell cytoplasm via size-exclusion chromatography (SEC) (121). There is little protein turn-over in the lens

fibre cells meaning crystallins are long-lived proteins that must maintain their structural integrity throughout life to preserve lens transparency (99). The highly soluble, stable and dynamic nature of the crystallins combine to facilitate their structural longevity (122). The structure and linearised domain organisation of representative crystallin isoforms of each class, i.e. αB-crystallin, truncated βB1-crystallin, βB2-crystallin, γD-crystallin, and γS-crystallin (Figure 5. A, B, C, D, and E respectively), demonstrate the structural features of each crystallin class, which will be discussed.

The α-crystallins constitute 30-40 % of the lens crystallins (123) and are members of the small heat-shock protein (sHsp) family. They are unrelated to the other (β and γ) crystallins (124). Two isoforms exist in the lens, αA- (more acidic - pI ~ 5.5) and αB-crystallin (more basic - pI ~6.8) that share ~ 60 % sequence similarity (115). In the young human lens, there is approximately a 2:1 ratio of αA- to αB-crystallin which changes to a 3:2 ratio by ~ 55 years of age (125). While αA-crystallin is exclusively lenticular, αB-crystallin is found extensively in many extra-lenticular tissues where its expression is upregulated in times of cellular stress (126, 127). The α-crystallins have a monomeric mass of ~ 20 kDa but form large, polydisperse, oligomeric structures with a molar mass range of 300 kDa – 1 MDa (128, 129). Both αA- and αB-crystallin have a similar domain organisation, consisting of a largely disordered N-terminal region, a central ACD in an immunoglobulin-like fold (130), and a short unstructured C-terminal region consisting of a highly flexible extension of 16 residues for αA-crystallin and 12 residues for αB-crystallin (131, 132) (Figure 5. A). αA- and αB-crystallin associate together (forming α-crystallin in the lens) or, in the case of αB-crystallin, with other sHsps outside of the lens to form hetero-oligomers (133, 134). This association is facilitated by dynamic subunit exchange which occurs between both homo- and hetero-oligomers (135, 136). One of the major oligomeric species of αB-crystallin is the 24-mer, for which an atomic level model has been elucidated (137) (Figure 5. A). It shows αB-crystallin monomers form dimers via the central α-crystallin domain (ACD) which assemble into a trimer of dimers, i.e. a hexmeric ring, through interactions with the flanking C-terminal region and the ACD (136, 137). Finally, the N-terminal region is responsible for high-order assembly but the specifics of these interactions are not entirely clear.

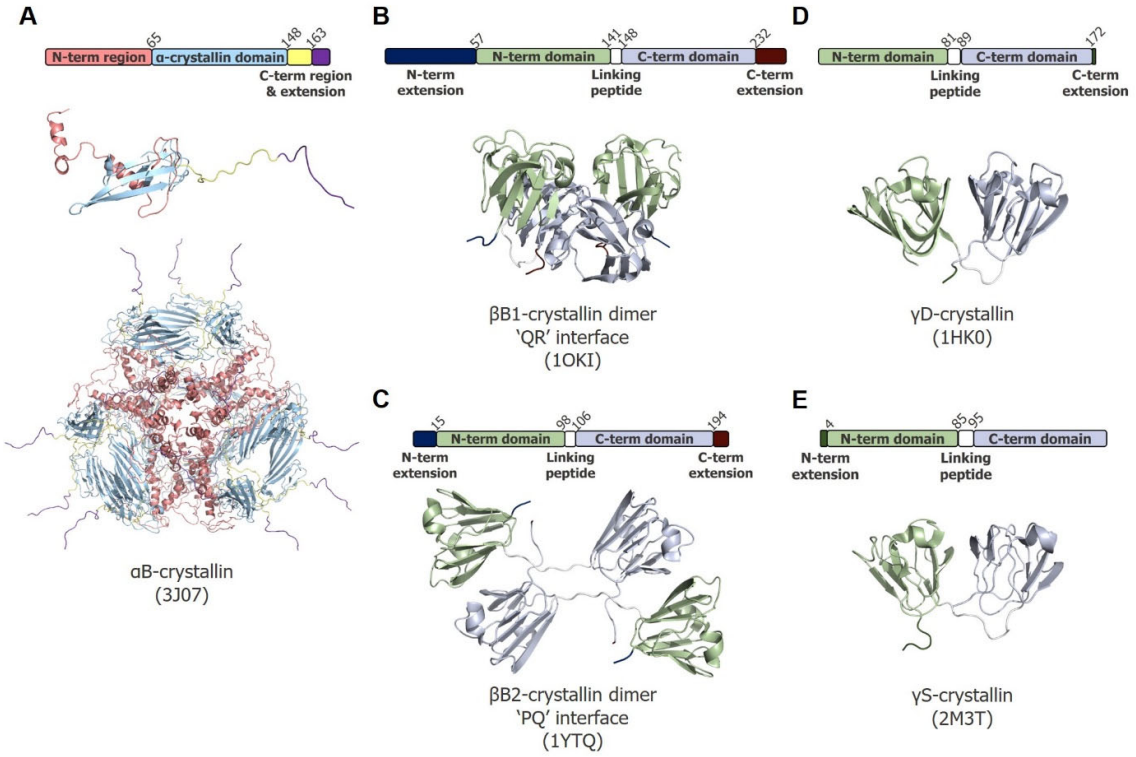


Figure 5. The crystallins of the eye lens. A. An atomic model of the αB-crystallin monomer and 24-mer (137). **B.** The crystal structure of the 'QR' truncated βB1-crystallin dimer (138) and **C.** the 'PQ' domain-swapped βB2-crystallin dimer (139). **D.** The crystal structure of monomeric γD-crystallin (140) and **E.** the NMR-derived solution structure of monomeric γS-crystallin (141). Linearised domain organisation coloured commensurate to the structure is included above each respective structure while the protein name and PDB accession code is included below.

In addition to contributing to the refractive power of the lens, the α-crystallins also function as molecular chaperone proteins that can prevent protein aggregation (28, 33). This property is particularly important in the lens where the long-lived nature of the crystallins means they are particularly susceptible to damage, subsequent unfolding, and aggregation occasioning cataract (90). The mechanism for α-crystallin (and sHsp) chaperone function is elusive. Numerous studies implicate interactions from multiple regions of the protein as being involved in chaperone action, which is consistent with the dynamic nature of α-crystallin structure and hence chaperone function (142, 143). Both αA- and αB-crystallin are distributed evenly throughout the lens; they are expressed at low levels in lens epithelial cells and then at high levels upon differentiation to lens fibre cells (124, 144). Due to the high concentrations of crystallins in the lens, the effects of macromolecular crowding on crystallin

structure and function are anticipated to be significant. Thus, macromolecular crowding could affect the structure and function of eye lens crystallins including the α-crystallins with consequent implications for lens structure and cataract formation (145). The effects of macromolecular crowding on αB-crystallin structure and function was investigated in Chapter 3 of this thesis.

The β- and γ-crystallins are structurally related proteins whose high concentrations, tight supramolecular packing and aromatic and sulphur-rich amino acid sequences provide a significant contribution to the refractive power of the lens (146, 147). They belong to a structurally homologous superfamily exhibiting a Greek-key motif (148). The β- and γ-crystallins consist of two structurally homologous domains which are connected by a flexible linker peptide (Figure 5. B-E). Within each domain, there are two Greek-key motifs of around 40 amino acids which are symmetrically intercalated forming a β-sheet sandwich about a set of hydrophobic residues (122, 149). Structural features such as interdomain interactions, hairpin loops, and tyrosine and tryptophan corners contribute to the high stability of these proteins (85, 150).

The β-crystallins constitute 30-40 % of the lens crystallins with monomeric masses ranging from approximately 22 to 28 kDa (115, 123). There are seven β-crystallin isoforms expressed in the human lens that share 45-60 % sequence identity and are designated as either acidic, βA1, βA2, βA3, and βA4 (pI: 5.7-6.4), or basic βB1, βB2, βB3 (pI: 5.9-8.6) (115, 123). The same gene encodes βA1- and βA3-crystallin but their respective gene synthesis is initiated at different start codons leading to an N-terminal extension that is 17 amino acids longer for βA3- than βA1-crystallin (151). The acidic and basic isotypes all have N-terminal extensions of varying lengths while only the basic β-crystallins, such as βB2-crystallin, also have a C-terminal extension (115, 152-154) (Figure 5. B and C). The extensions are unstructured and flexible. The β-crystallins share around 30 % sequence identity with the γ-crystallins and the most pronounced sequence difference is the presence of the aforementioned long N-terminal extensions in the β-crystallins (115, 122). Despite the structural relationship between the β- and γ-crystallins, the β-crystallins can form homo- or hetero-oligomeric complexes with oligomeric states ranging from dimers to octamers while γ-crystallins are largely monomeric (89, 122).

Structures exist for truncated βB1- and full-length βB3- and βA4-crystallin (the latter two are deposited in the PDB but are otherwise unpublished) which all form dimeric structures with an interface

classically described by crystallographic dyads 'QR' due to a bent linking peptide (138, 149) (Figure 5. B). The structure of βB3-crystallin also exhibits a putative trimer presenting a previously undescribed oligomeric interface (149). The structure of βB2-crystallin is domain swapped which is facilitated by an extended linking peptide (Figure 5. C). A dimer is formed around a PQ interface and two dimers come together to form a lattice tetramer in the QR orientation (139, 155). However, it has also been shown that the structure of βB2-crystallin in solution, using SAXS and NMR, is a compact 'face-en-face' (QR) conformation as opposed an extended domain-swapped one observed in the crystal lattice (156) (Figure 5. B and C).

Temporal and spatial differences in the expression patterns of β-crystallins throughout lens development lead to differences in the abundance of β-crystallin isotypes throughout the lens (157). βB1- and βB3-crystallin are expressed early in lens development and as such are located primarily in the lens nucleus (158, 159). In contrast, βB2-crystallin is expressed sightly later but to a greater extent relative to βB1- and βB3-crystallin, thereby increasing its spatial abundance throughout the lens (159, 160). Consequently, βB2-crystallin is the most abundant β-crystallin in the developed lens (123). Generally, the genes associated with the acidic β-crystallins are more widely expressed leading to a spatial distribution that spans both the nucleus and cortex of the lens (159, 160). The exception is βA2-crystallin of which only trace amounts are found in humans in comparison to other mammals despite high levels of the relevant mRNA having been detected (161). These differences in expression patterns lead to variation in the distribution of the different oligomeric forms of the β-crystallins throughout the lens with larger oligomers (octamer/hexamer) in the nucleus and smaller oligomers (dimers) in the cortex (108, 162, 163). It is postulated that this could be a mechanism to control water content in different regions of the lens due to osmotic pressure effects (115, 164). In addition, larger oligomers may also impart greater stability to older regions of the lens. The former postulate is consistent with a shift in the β-crystallin oligomeric population from dimers to tetramers in the aging and cataract-prone lens (125, 165).

The γ-crystallins constitute 20-30 % of the lens crystallins with monomeric masses ranging from approximately 20 to 21 kDa (115, 123). Six γ-crystallins, γA-F-crystallin, are closely linked on a gene cluster and exhibit 70-98 % sequence identity (149). In humans, γA-crystallin is expressed at low

levels while the genes encoding γE- and γF-crystallin are inactive due to premature stop codons that are not present in most other mammals (115, 149). The gene encoding the seventh γ-crystallin, γS-crystallin, is on a different chromosome and has greater sequence divergence exhibiting approximately 50 % identity compared to the other γ-crystallins (166). The γ-crystallins maintain the same general domain arrangement with the exception of differences in N- and C-terminal extensions and linking peptide length (Figure 5. D and E). The γA-F-crystallins contain a C-terminal extension consisting of two residues (Figure 5. D) in contrast to γS-crystallin which lacks this extension but instead has a flexible four-residue N-terminal extension (115, 167) (Figure 5. E). The linking peptides of γB- and γS-crystallin are one and two residues longer than the other γ-crystallins, respectively (115). For γS-crystallin, these differences originally it led to be characterised as a lens β-crystallin named βS-crystallin as it elutes between β- and γ-crystallin fractions during SEC separation but genetic sequence analysis suggested a closer relationship to the γ-crystallins (168, 169). Despite generally being considered monomeric proteins (170), γC-, γD-, and γS-crystallin form disulfide-linked dimers with the structure and stability for the latter being described and characterised in Chapter 2 of this thesis (171-173).

The γ-crystallins are the last of the lens crystallins to be expressed during fibre cell differentiation and as such are not observed in immature fibre cells of the lens cortex (144, 174). Due to relatively early induction of their expression in lens development, there is a significant abundance of γB- and γD-crystallin in the nuclear region of the lens (115, 144, 174). Further, the expression of γB- and γC-crystallin lingers for longer than the other γ-crystallins thereby adding these isoforms to the inner lens cortex in addition to the nucleus (115, 122). Conversely, γS-crystallin is induced later in lens development with levels increasing drastically in the postnatal lens placing a significant portion of γS-crystallin in the lens cortex (125, 175). As different species require different levels of accommodation from their eye lens, the abundance and variation in quantity of specific γ-crystallin isoforms contributes to the physical properties required to achieve this. Thus, soft human lenses contain an abundance of γC-, γD-, and γS-crystallin, the very soft avian lens contains γS-crystallin almost exclusively, and hard rodent lenses contain an abundance of γA-F-crystallin (123, 176-179).

Cataract and the Crystallins: the Development of Congenital, Early On-Set, and Age-Related Cataract

Cataract is defined as the loss of transparency and the opacification of the eye lens and can be classified as either congenital (present at birth), early on-set (present from birth to approximately 45 years of age) or age-related (onset from approximately 50 years of age) (180). Patients with congenital or early-onset cataract are generally genetically predisposed, largely due to deleterious mutations in lens crystallin proteins (181, 182) (Tables 2-4). Age-related cataract manifests due to accumulative damage caused to the long-lived lens crystallins by factors such as heat and time, leading crystallin destabilisation and spontaneous post-translational modifications (PTMs) resulting in crystallin aggregation (90, 183) (Table 5).

Cataract can be heterogeneous in morphology and spatial location which has led to the classification of several different types. While distinct types have been observed for different congenital or age-related cataracts, it is still not clear what link these morphologies have with the associated disease etiology (94). However, determination of a definitive link between cataract morphology and the underlying biochemical process could be important for future treatment options for cataract. The various cataract morphologies and their locations in the eye lens are schematically depicted in Figure 5. Nuclear, cortical and subcapsular cataracts (Figure 5. A-D) are largely age-related (180, 183). However, numerous other types such as polar, cerulean, lamellar, pulverulent, aculeiform, polymorphic and total cataract have been associated with different crystallin mutations in patients with congenital or early on-set cataract (Figure 5. E-L) (Tables 2-4) (182, 184). They can also be localised to a specific area of the lens, leading to a description that encompasses spatial and morphological features, e.g. nuclear pulverulent or cortical cerulean cataract. Other ocular abnormalities may also accompany cataract, some of which include microcornea (cornea < 10 mm in diameter), microphthalmia (abnormally small and anatomically malformed eye), nystagmus (uncontrolled and repetitive eye movements), and coloboma (missing pieces of eye tissue) (Tables 2-4) (185). Due to its extralenticular role as a molecular chaperone, some congenital αB-crystallin mutants can lead to muscular disorders such as myofibrillar myopathy in addition to cataract (186, 187) (Table 2).

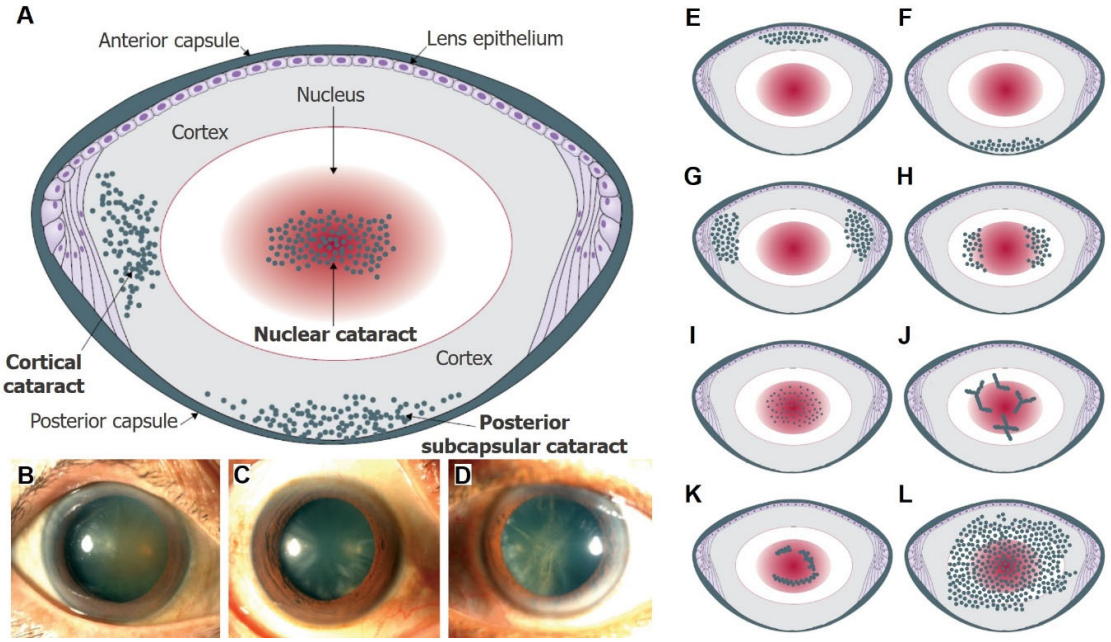


Figure 6. Various morphologies of cataract and their locations within the eye lens. A. A schematic of the lens showing its anatomical features (normal typeface) and the location of the main cataract morphologies associated with age-related cataract (bold typeface, blue-grey dots). Slit lamp microscopy photos showing **B.** nuclear cataract, **C.** wedge-shaped cortical cataract, and **D.** posterior subcapsular cataract (180). Other types of cataract, which can appear with congenital and early-onset cataract, and their general morphology and spatial position in the lens are schematically represented (blue-grey dots) (182, 184), including **E.** anterior polar cataract, **F.** posterior polar cataract, **G.** cerulean/blue dot cortical cataract, **H.** lamellar/zonular cataract, **I.** pulverulent/Coppock-like nuclear cataract, **J.** aculeiform/coralliform cataract, **K.** polymorphic cataract, and **L.** total cataract. Modified from Liu *et al.*, 2017 (180).

Congenital and early-on set cataracts are detected via genetic sequencing of a proband and their descendants to determine the mutation(s) that cosegregate with a cataract phenotype. A multitude of congenital mutations in humans have been identified across all α, β, and γ-crystallin isoforms (Tables 2-4). Biophysical characterisation of some of these mutants has provided insight into the potential molecular mechanisms of cataract. For example, the R120G mutant of αB-crystallin disrupts a vital salt-bridge across the ACD dimer interface (188), which leads to unfolding of the protein, an increase in oligomeric size and significantly reduced chaperone activity (189, 190). The Q155X mutant of βB2-crystallin results in a sequence truncation that leaves protein 51 amino acids short at its C-terminal end, causing the loss of the last Greek key motif (191) (Figure 5. C). Structural studies have demonstrated that this mutation results in a significant loss of ordered structure, reduced chemical stability,

Table 2: Documented mutations in human α-crystallins and their associated cataract types.

Crystallin Protein	Congenital Mutation	Type of Cataract
αA-crystallin	W9X[a]	Recessive cataract (no morphology documented) (192)
	R12C	Posterior polar and nuclear cataract with associated microcornea (193, 194)
	R21W	Lamellar cataract with associated microcornea (193, 194)
	R21L	Nuclear cataract (195)
	R49C	Nuclear cataract (196)
	R54C	Total cataract with associated microcornea (194, 197)
	G98R	Total cataract (198)
	R116C	Various cataract phenotypes (193, 199-202)
	R116H	Various cataract phenotypes (193, 203, 204)
αB-crystallin	R11H	Nuclear cataract (205)
	P20S	Posterior polar cataract (206)
	R56W	Nuclear and total cataract (207, 208)
	D109A	Posterior polar cataract with associated myofibrillar myopathy (209)
	D109H	Posterior polar cataract with associated myofibrillar myopathy and cardiomyopathy (210)
	R120G	Dominant cataract (no morphology documented) with associated myofibrillar myopathy (186)
	D140N	Lamellar cataract (206)
	K150fs[b]	Posterior polar cataract (211)
	A171T	Lamellar cataract (194)

[a]X indicates a change to a stop codon.
[b]fs indicates a 'frameshift' mutation in the reading frame beginning at the particular amino acid and position indicated.

diminished protein-protein interactions, and high surface hydrophobicity (212). The G18V mutation of γS-crystallin replaces a highly conserved glycine in the protein's first Greek-key motif with a valine residue (115). In comparison to the wild-type protein, G18V has an increased exposed hydrophobic surface area, a lower chemical and thermal stability, undergoes three-state unfolding, and has a perturbed protein hydration shell in crowded conditions (213-215).

In general, crystallin mutations can alter folding (212, 216), surface hydrophobicity (214, 217), oligomerisation (189, 218), conformational dynamics (219, 220), stability (213, 221), protein-protein interactions (222, 223), solubility (224, 225), and chaperone activity (221, 226) relative to the wild-type protein. However, examples have also been documented of seemingly innocuous alterations to the structural features of mutant crystallins compared to their wild-type counterparts (140, 227). Thus,

Table 3: Documented mutations in human β-crystallins and their associated cataract types.

Crystallin Protein	Congenital Mutation	Type of Cataract
βA1/A3-crystallin	G91del[a]	Nuclear and lamellar pulverulent cataract (228-231)
	Splice-donor[b]	Pulverulent nuclear and lamellar cataract (232, 233)
βA2-crystallin	V50M	Dominant cataract (no morphology documented) (234)
βA4-crystallin	G64W	Nuclear cataract with associated microcornea (235)
	F94S	Lamellar cataract with associated microphthalmia (236)
βB1-crystallin	M1K	Pulverulent nuclear cataract (237)
	N58fs	Nuclear cataract (238)
	S93R	Nuclear cataract with associated microcornea and microphthalmia (239)
	V96F	Dominant cataract (no morphology documented) with associated microcornea (234)
	L116P	Nuclear cataract (240)
	R123H	Sporadic cataract (no morphology documented) (241)
	S129R	Nuclear cataract with associated microcornea (242)
	Y206fs	Nuclear cataract (243)
	G220X	Pulverulent cataract (237)
	Q223X	Nuclear cataract (244)
	Q227X	Nuclear cataract with associated nystagmus (245)
	S228P	Nuclear cataract (246)
	R230C	Cataract (no morphology documented) with associated coloboma (247)
	R233H	Nuclear cataract with associated nystagmus (248)
	X253R	Nuclear, anterior, and posterior polar cataract with associated microcornea (249)
βB2-crystallin	A2V	Posterior subcapsular cataract (250)
	I21N	Nuclear cataract (248)
	P115T	Sporadic cataract (no morphology documented) (241)
	G119R	Nuclear cataract (240)
	V146N	Nuclear cataract with associated microcornea (248)
	G149V	Cataract (no morphology documented) with associated microcornea (247)
	Q155X	Polymorphic, pulverulent and cerulean cataract (191, 251, 252)
	S186P	Sporadic cataract (no morphology documented) (241)
βB3-crystallin	G165R	Nuclear cataract (253)
	V194E	Nuclear, cortical, anterior, and posterior polar cataract (234)

[a]del indicates at deletion mutation. In this instance, three base pairs are deleted resulting in the removal of a glycine residue at position 91.
[b]splice-donor indicates a mutation in the splice site at the beginning (5' end) of an intron that results in aberrant exon assembly and the translation of a non-functional protein.

Table 4: Documented mutations in human γ-crystallins and their associated cataract types.

Crystallin Protein	Congenital Mutation	Type of Cataract
γC-crystallin	T5P	Coppock-like cataract (254)
	M44fs	Cataract (no morphology documented) with associated microcornea (247)
	R48H	Nuclear and lamellar cataract (255)
	G62fs	Lamellar pulverulent cataract (256)
	C109X	Nuclear cataract (257)
	S119S	Nuclear cataract with associated microcornea (258)
	G129C	Nuclear cataract (259)
	Y144X	Cataract (no morphology documented) with associated microcornea (247)
	W157X	Nuclear cataract with associated microcornea (260)
	R168W	Nuclear and lamellar cataract (261, 262)
γD-crystallin	R14C	Nuclear and coralliform cataract (263, 264)
	R14S	Coralliform cataract (265)
	P24T	Cerulean, coralliform and lamellar cataract (261, 266-271)
	P24S	Polymorphic cortical cataract (272)
	A36P	Nuclear cataract (273)
	R37P	Nuclear cataract (274)
	R37S	Birefringent, pleiochroic crystals (275)
	W43R	Nuclear cataract (276)
	M44V	Blue dot cataract (277)
	Y56X	Nuclear cataract (258)
	R58H	Aculeiform cataract (254)
	G61C	Coralliform cataract (278)
	R77S	Anterior polar cataract (279)
	E107A	Nuclear cataract (280)
	Y134X	Dominant cataract (no morphology documented) with associated microcornea (193)
	R140X	Nuclear cataract (194)
	W157X	Nuclear cataract (261)
	G165fs	Nuclear cataract (281)
γS-crystallin	G18D	Cortical and sutural cataract (282)
	G18V	Polymorphic cortical cataract (283)
	D26G	Coppock cataract (273)
	S39C	Lamellar and sutural cataract (194)
	V42M	Nuclear cataract (284)
	G57W	Pulverulent cataract (285)
	Y67N	Nuclear cataract (286)
	G75V	Lamellar cataract (282)

Table 5: Major PTMs of human lens crystallins.

Post-translational modification	Chemical Process Result of Modification Residues Commonly Affected	Examples Identified in Human Lens Crystallins
Deamidation	Elimination of a side-chain amide and formation and hydrolysis of a succinimide intermediate, typically converting the affected residue to its corresponding carboxylic acid derivative	Q147 αA-crystallin (287) N82 βA4-crystallin (288) N76 γS-crystallin (288)
	Conversion of neutrally charged to negatively charged residue at physiological pH	
	Asparagine and glutamine	
Racemisation	Interconversion of an amino acid between its L- and D-forms, via deprotonation and protonation of the amino acid α-carbon or through a succinimide intermediate	D109 αB-crystallin (289) D4 βB2-crystallin (290) D153 γS-crystallin (291)
	Conversion of amino acid from L- to D-enantiomer	
	Asparagine, aspartic acid, and serine	
Phosphorylation	Covalent addition of a phosphoryl group, likely via a protein kinase	S45 αB-crystallin (292) T117 βB2-crystallin (293) Y62 γC-crystallin (293)
	Addition of negative charge and bulkier amino acid R-group	
	Serine, threonine and tyrosine	
Truncation	Protease-mediated and spontaneous cleavage of peptide bonds and loss of amino acids from the N- or C-termini	S172[a] αA-crystallin (287) N22[a] βA3-crystallin (294) S87[b] γD-crystallin (295)
	Loss of one, several or many amino acids	
	Asparagine and serine	
Oxidation	Interaction with oxygen, a loss of hydrogen or electron(s).	W9 αB-crystallin (296) M226 βB1-crystallin (293) C24 γS-crystallin (297)
	Intramolecular, intermolecular and mixed disulfides, sulfoxide formation, scission of indole ring	
	Cysteine, methionine, and tryptophan	

[a]truncation occurs after the residue indicated.
[b]truncation occurs before the residue indicated.

while biophysical studies of specific congenital crystallin mutations might be able to rationalise their role in certain types of cataract, the overall role of these mutations in the structural and biochemical changes that lead to the development of cataract are less clear. However, it is important to note that despite the identification and study of congenital crystallin mutations associated with cataract being important to our fundamental understanding of crystallin structure and function and cataract

development, the majority of human cataracts occur as a result of aging and do not arise via mutations (94, 183).

Extensive proteomic analyses of lenses at different ages, both normal and cataractous, have identified numerous PTMs associated with the long-lived lens crystallins (287, 288, 298, 299). Some of the major PTMs are deamidation, racemisation, phosphorylation, truncation, and oxidation which have all, to some extent, been associated with age-related cataract formation (287, 288, 298) (Table 5). Some of these post-translationally modified crystallins have been biophysically characterised either via mutagenesis studies, e.g. deamidation and truncation studies, or via modification mimics, e.g. of racemisation and phosphorylation. As a result, these studies have enhanced our understanding of how these modifications might affect crystallin structure and function and contribute to the development of age-related cataract.

Specifically, deamidation is arguably the most abundant PTM in the aged human lens (288). It can be studied *in vitro* via mutagenesis of relevant neutrally changed asparagine or glutamine residues into negatively charged aspartic acid or glutamic acid, respectively. Across all types of crystallins, deamidation has a mostly deleterious effect causing conformational change and destabilisation (300), altered protein dynamics (301), partial loss of chaperone activity (302), and an increase in protein-protein interactions (303). However, the magnitude and nature of the changes relative to the wild-type crystallin are dependent on the amino acid sequence position of deamidation (302). For instance, the N76D γS-crystallin and Q143E αA-crystallin deamidation mutants had small and localised structural differences in comparison to their wild-type counterparts, however both mutants exhibited an increased tendency to oligomerise (304, 305).

Racemisation is a prolific PTM in the aging and cataractous eye lens (306, 307). The impacts of this PTM has been difficult to study in full-length crystallins because the incorporation of an unnatural D-form amino acid has not yet been achieved. While biophysical studies on crystallins that exhibit racemisation are sparse, some insight has been provided from mutational studies on racemisation prone crystallin residues to understand how changes in that amino acid affect crystallin structure and function. Changes to crystallin residues that exhibit an increased abundance of racemisation between the water-soluble and water-insoluble fractions of the lens have been investigated e.g. D151

αA-crystallin and D109 αB-crystallin (308, 309). Multiple mutations at D151 in αA-crystallin lead to a loss of stability and chaperone activity (310) while experimental studies on D109A αB-crystallin led to a loss of stability and solubility of the protein, recapitulating the results of molecular dynamics calculations which predicted the breakage of an important inter-subunit salt bridge due to D109 racemisation (289). These results have provided a general understanding of the potential biophysical consequences of racemisation and its relation to aging and cataract formation.

The incorporation of a phosphorylated amino acid is possible with unnatural amino acid incorporation methodology (311). However, this approach has not been applied to crystallin proteins. Previous studies have used phosphomimics and mutated residues at sites that have been identified to be phosphorylated *in vivo* with either a negatively charged aspartic or glutamic acid. The studies of crystallin phosphomimics have been confined to the α-crystallins as it is a major PTM for this crystallin type (34), including occurring to a significant degree early in life (312). Using either an aspartic or glutamic acid to substitute specific serine residues, the triple phosphomimic in the N-terminal region of αB-crystallin (S19/S45/S59) has an altered oligomeric state (313-315), increased subunit exchange rate (314), a more solvent exposed and flexible N-terminal region (315), and an increase in its *in vitro* chaperone activity for target proteins under stress conditions (313, 315). The consensus from these phosphomimic studies is that phosphorylation appears to be a 'gain of function' modification that regulates sHsp chaperone function (34). Currently, there has been no biophysical characterization of the effects of phosphorylation on β- or γ-crystallins despite this PTM being detected in the eye lens (Table 5).

Truncation of the polypeptide from both termini has been identified across all crystallin types and particularly within unstructured and flexible terminal extensions which are a characteristic of all three crystallin types (316) (Figure 5). Truncation mutants can be generated via addition of either a late start or early stop codon in the gene sequence of the crystallin of interest. Significant (more than one amino acid) C-terminal truncation of αA-crystallin leads to lower oligomeric average mass (317, 318), a decreased rate of heterooligomeric subunit exchange (317), impaired structure and solubility (319), and reduced chaperone activity (318). The N-terminal truncation of βB1-crystallin resulted in phase

separation (320), impaired higher order heterooligomer formation (165, 321), and a reduced heterooligomeric subunit exchange rate and solubility changes (321).

Despite the diversity of PTMs identified in the human lens, oxidation can be considered the hallmark modification of age-related cataract as its abundance is highly correlated with age and the onset of age-related cataract, particularly nuclear cataract (90, 297, 322-325). Advanced age-related nuclear cataract leads to oxidation of more than 90% of cysteine thiol groups and approximately 50% methionine sulfhydryl groups (297, 322, 326). The major reason lens protein oxidation becomes more pronounced with age is due to a loss of the lens's principal reducing agent glutathione, which is approximately 6 mM in concentration in young normal lenses but decreases to close to 1 mM or less in old and cataractous lenses (327-329). The *in vitro* oxidation of cysteine residues is possible through passive means, e.g. oxygenated solution, high crystallin protein concentration, and time (173), or active means, e.g. addition of a strong chemical oxidant (e.g. diamide or a redox active metal like copper) (330). Photo-oxidation of tryptophan can form the derivative kynurenine due to scission of tryptophan's pyrrole ring, producing a far more polar side chain by introducing amino and carbonyl groups into the side chain (331). This oxidation can be mimicked via mutagenesis of tryptophan to glutamine (332). It has been demonstrated that copper(II) treated γD-crystallin, which produces a disulfide-linked γD-crystallin dimer, and tryptophan oxidation mimic W42Q γD-crystallin both exhibit conformational change, instability, and increased aggregation propensity (330, 332). Interestingly, it has also been found that the presence of an intramolecular disulfide in the W42Q γD-crystallin tryptophan oxidation mimic, which further distorts the Greek-key fold, causes the oxidation mimic to be aggregation-prone while the same intramolecular disulfide is not deleterious to wild-type γD-crystallin (333). Disulfide-linked γS-crystallin (via C24) is destabilised and has an increased aggregation propensity, aspects which are explored in-depth in Chapter 2 of this thesis (173). These examples provide a fascinating, albeit small, view into how oxidative changes can affect lens crystallin proteins structure and biophysical properties and lead to aggregation associated with cataract formation.

Biophysical Methods for the Characterisation of Crystallin Proteins

Due to the long-lived nature of the crystallins, there has been intense interest in elucidating their structure and physical properties to understand the mechanisms behind their function, dysfunction and aggregation. However, the investigation of these properties has been hindered in some instances due to the highly soluble and dynamic nature of the lens crystallins. To counter this, the development of traditional and novel techniques has enabled a closer look at these proteins. Below, some of the prominent biophysical methodologies and associated findings that have contributed to our increased understanding of crystallin structure and function are described.

X-ray crystallography was used to derive the first structure of a crystallin bovine γB-crystallin (isolated from the bovine lens and previously referred to as γII-crystallin) nearly 40 years ago (334). Nearly a decade later, the first structure of a β-crystallin, bovine βB2-crystallin, was elucidated as a domain-swapped dimer (155). Since then, structures of the human versions of γB-crystallin (335), γD-crystallin (140) (Figure 5. D), truncated βB1-crystallin (138) (Figure 5. B), βB2-crystallin (139) (Figure 5. C), βB3-crystallin, and βA4-crystallin have also been solved. There are still a number of β- and γ-crystallins that have not been crystallised, however progress is being made. One such example is γS-crystallin whose bovine (336) and human (150) forms of the isolated C-terminal domain have been crystallised but attempts to crystallise the full-length protein only came almost two decades later. The crystal structure of chicken γS-crystallin has now been solved (337) while the first full-length crystal structure of human γS-crystallin was determined of an oxidised form as a disulfide-linked dimer (173). The latter is described in detail in Chapter 2 of this thesis. These recent crystallographic achievements suggest that the remaining β- and γ-crystallins will likely be amenable to crystallisation, if only upon mild modification. Crystallography of the two full-length α-crystallins is unlikely due to their polydisperse, dynamic and flexible nature (338). To overcome this, the largely disordered N-terminal region and part of the C-terminal region were removed leaving the structured ACD which has been crystallised for multiple species and mutants of both αA- and αB-crystallin (130, 188, 339-342).

X-ray crystallography has provided an unparalleled understanding of the structural homology of related β- and γ-crystallins previously inferred via sequence homology and gene duplication (343).

Examination of other structural features such as tyrosine corners, relative solvent exposure of residues and the positions and interactions of residues that are mutated or modified in association with cataract formation have been made possible by X-ray crystal structures (150). Crystallography has also provided significant insight into the interfacial interactions of the oligomeric α- and β-crystallins (139, 188). However, a shortfall of X-ray crystallography is that it deals with a static, solid-state structure. Therefore, experimental validation with the oligomeric structure formed in-solution is judicious. A prudent example is the domain-swapped conformation of the βB2-crystallin dimer which was shown to be a crystallographic artefact as small angle X-ray scattering and NMR methods determined the dimer formed around the more compact QR interface like the crystal structure of truncated βB1-crystallin (156) (Figure 5. B and C). While the structures that have been solved have provided relevant precursors for molecular dynamics simulations (289, 344), experimental validation of flexibility and dynamism is usually the domain of NMR spectroscopy.

NMR spectroscopy has provided much information on crystallin structure and dynamics. The α-crystallins were thought to not be amendable to conventional solution-phase NMR due to their large oligomeric size. However, it was observed that 12-16 residues at the C-terminal end of αA- and αB-crystallin are readily observable via NMR spectroscopy (131) (Figure 5. A), arising from a solvent-exposed, highly flexible and dynamic C-terminal extension that protrudes from the high-mass oligomer (345). The C-terminal extension of α-crystallins and other mammalian sHsps has been a focus of further studies since its discovery approximately 25 years ago providing insight into its role in oligomerisation (346), chaperone action (132), proline isomerisation (347), and the identification of different conformers in solution (142). In addition to the α-crystallins, solution-state NMR has also been used to discern flexible N- and C-terminal extensions in β-crystallins (152, 153) (Figure 5. B and C) and the short flexible N-terminal extension in γS-crystallin (167) (Figure 5. E).

Advances in solution- and solid-state NMR hardware and pulse sequences as well as isotopic protein labelling have been applied to both full-length αB-crystallin and the isolated ACD to elucidate residues involved in oligomerisation (137, 348-350), chaperone action (141, 142, 349), structural dynamics (351, 352), metal binding (353) and derive structures of the wild-type ACD (137, 349, 352). While NMR studies of β-crystallin calcium binding (354) and structure (156) have provided insight into

their structural and physical properties, the ~40 kDa mass of the β-crystallin dimer limits the resolution available from solution-state NMR. However, NMR studies on the ~ 20 kDa monomeric γ-crystallins have been numerous providing complementary and orthogonal information to that derived from X-ray crystallography. The structures of γS- and γC-crystallin and some of their cataract-associated mutants, which have proven recalcitrant to crystallisation, have been solved using NMR spectroscopy (141, 355-357) (Figure 5. E). NMR spectroscopy has also been used to investigate sparsely populated conformers and the unfolding mechanism of a cataract-associated variant of γS-crystallin providing insight into structural changes that precede aggregation (358, 359). Furthermore, solid-state NMR has shown that aggregates of a cataract-associated mutant of γD-crystallin that appear largely amorphous in form via TEM still retained a significant portion of its native structure, with potential implications for the targeting and prevention of such aggregates in cataract (60).

While size limitations currently preclude the structural determination of the dimeric β-crystallins and all monomeric γ-crystallins via negative stain TEM or cryo-electron microscopy (cryo-EM), major insight into the quaternary structure of the large α-crystallins has been obtained for over two decades. The first structural model via cryo-EM data of αB-crystallin's quaternary structure was from an asymmetric 32-subunit oligomer with a large central cavity (360), confirming a previous hypothesis that the α-crystallin oligomer contained a central cavity (361). Later, negative stain transmission electron microscopy data were used to build a three-dimensional model of an αB-crystallin 24-mer showing a similar central cavity but an overall symmetrical oligomer (362). In an effort to incorporate more detail into these models, hybrid approaches have been adopted to derive pseudo-atomic models (363) utilising NMR and structural modelling in combination with either small angle X-ray scattering (349), negative stain TEM (137), or cryo-EM (364). Cryo-EM also allowed for other oligomeric quaternary structures of αB-crystallin to be determined from a 12-mer to a 48-mer with some oligomers also exhibiting different conformers (364). Despite the usefulness of these quaternary structures and pseudo-atomic models, it is difficult to depict αB-crystallin as a single structure given the wide distribution and conformational heterogeneity of the oligomers (365). Indeed, the conformational heterogeneity of αB-crystallin and probably most, if not all, sHsps is an integral feature of their structure and chaperone action.

Mass spectrometry has provided detailed information on the continuum of oligomeric states in the α-crystallins with technological advances allowing structural models of the α-crystallin oligomeric arrangements to be proposed. Mass spectrometry has also been important in obtaining structural data for the smaller β- and γ-crystallins. Mass spectrometry has been used extensively for investigating crystallin PTMs including phosphorylation (293), truncation (160), isomerisation (291), oxidation (287), acetylation (366), methylation (367), and deamidation (288). Structural studies have also used hydrogen-deuterium exchange, cross-linking, and ion mobility mass spectrometry techniques. They have been applied to β-crystallins to discern crystallin-crystallin interactions and solvent accessibility (368, 369), size and stability (370), and dynamic conformational changes (371). Native mass spectrometry of αA- and αB-crystallin has been particularly useful in elucidating the mass and number of oligomeric states (128, 372), subunit stoichiometry (128, 373), chaperone action (373, 374), oligomer and ACD dimer conformation (188, 375), and subunit exchange (317, 376) of these proteins. Mass spectrometry data have also been used to build a model of αB-crystallin, in conjunction with NMR data, which rationalises the underlying thermodynamics and kinetics of oligomerisation and subunit exchange (350, 351, 372, 376). While native mass spectrometry has provided an unprecedented insight into the polydisperse nature of the α-crystallins, the analysis occurs in the gas phase (377). No technique currently exists to resolve oligomeric species of α-crystallins in-solution under native-like conditions (378).

Solution separation techniques enable characterisation of crystallins in a native-like environment (aqueous solution at physiological temperature and pH) which is important given the aforementioned techniques either do not involve separation and therefore rely on averaging or do separate but under non-native-like conditions. Techniques such as SEC and analytical ultracentriguation (AUC) have been used to study β- (154, 379) and γ-crystallins (173, 304) and provided great insight. However, for the α-crystallins, SEC and AUC have only been able to provide information on shifts in oligomeric population due to changes in solution conditions (e.g. pH changes) or mutations (e.g. phosphomimics), otherwise giving poor resolution for individual oligomeric species (338, 362).
These methods also rely on separation via hydrodynamic size and thereby miss information related to charge and conformational differences of these proteins, which is particularly important for some post-

translationally modified crystallins (293) and the structurally heterogeneous α-crystallins (142). Any alternative method would need a separation mechanism that is particularly sensitive to conformation and, in the case of resolving α-crystallin's oligomeric states, a separation timescale that is less than that of α-crystallin subunit exchange, i.e. seconds to minutes depending on temperature and solution conditions. Very recently, it has been observed that oligomeric states of αB-crystallin could be discerned using a microfluidic device and second-derivative signal analysis (378). The separation was effected by electrophoretic mobility on a time-scale shorter than subunit exchange, providing the first evidence that individual oligomeric states of native αB-crystallin can be resolved in solution (378).

Free-solution capillary electrophoresis (CE) also separates on the basis of electrophoretic mobility, and is far more common and less specialised than in-house fabricated microfluidic devices. The CE instrumentation consists of a capillary which is usually made of fused silica but can be coated with an inert surface to minimise analyte adsorption (380) (Figure 7, *left*). Commonly, capillaries will range in diameter from 25 to 500 μm and be anywhere between 7 to 100 cm long (381). The sample is introduced into the capillary before the start of each run via hydrodynamic pressure at the end opposite to the detector. Once the sample is in the capillary, small reservoirs of buffer are used to submerge both ends of the capillary along with two electrodes that then produce a potential difference across the capillary from 5 to 30 kV using direct current (Figure 7, *left*). Detection is usually spectrophotometric (UV-visible or fluorescence) with coupling to mass spectrometry also becoming a common feature (381). CE can create an electroosmotic flow (EOF) which is due, for a fused silica capillary in a solution above pH 3, to negatively charged deprotonated silanol groups on the capillary surface (380) (Figure 7, *top right*). A mass transfer of buffer cations towards the negatively charged capillary surface occurs, forming a fixed layer. A second layer, called the diffuse layer, is free to move laterally along the capillary while remaining close to the capillary wall. Given there are predominately cations in the diffuse layer, their migration toward the negatively charged cathode drags the bulk solution toward this electrode, thereby rationalising where the detector is placed (381, 382) (Figure 7, *top right*).

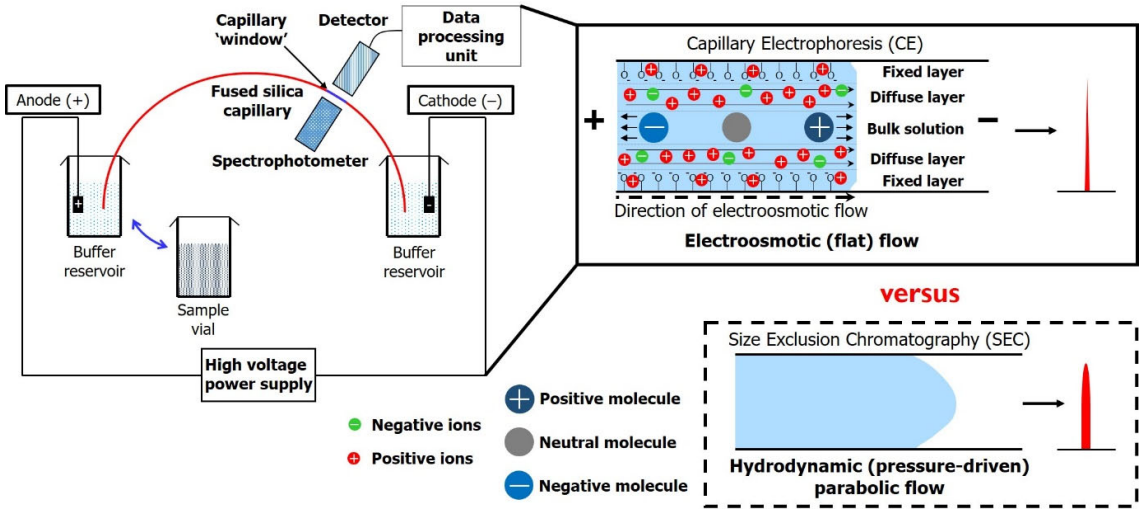


Figure 7. A schematic of a CE apparatus, the origin of the electroosmotic flow (EOF), separation of positive, neutral and negative molecules, and the corresponding flow profile compared with SEC. The features of a CE apparatus include the anodic and cathodic electrodes, buffer reservoirs, sample vial, capillary and spectrophotometer (*left*). Positive buffer ions form an electric double layer (the fixed and diffuse layers together) and pull the solution towards the negative cathode generating an EOF (*top right*). As there is a high concentration of positive ions in the diffuse layer near the capillary wall, application of a potential difference across the capillary causes the solution to flow towards the cathode from the edge of the capillary wall resulting in a flat-flow profile. Positive ions are detected first as they migrate faster than the EOF due to attraction to the cathode, neutral ions are detected second as they migrate at the same rate as the EOF, and negative ions are detected last as they migrate slower than the EOF due to attraction toward the anode (*top right*). Due to the EOF induced flat-flow profile in CE, this leads to sharp peaks and higher resolution separation in comparison to parabolic flow profiles from hydrodynamic methods, such as SEC, that lead to peak broadening and lower resolution separation (*bottom right*).

Due to the solution uniformly moving toward the cathode, positively charged species migrate faster due to electrostatic attraction toward the cathode and are thus detected first. Neutrally changed species are detected second as they are unaffected by the electric field and travel at the same rate as the EOF. As such, neutral molecules are usually used as markers to determine the EOF rate. Finally, negatively charged molecules travel the slowest as they are retarded due to their electrophoretic mobility counteracting the direction of the EOF as they are attracted toward the anode (Figure 7, *top right*). As a consequence of the solution being 'dragged' from the edges of the capillary, the EOF creates a 'flat' flow profile leading to very sharp CE peaks greatly increasing peak resolution in separation (382). This is in contrast to hydrodynamic parabolic flow profiles, such as in SEC, which consequently leads to broad peaks/signals (383) (Figure 7, *top left*).

Above a protein mass of ~10 kDa, CE separation becomes independent of mass (384, 385), a situation that is referred to as CE in the 'critical conditions' (386). The same behaviour has also been noted for different polyelectrolytes such as DNA (387) and charged polymers (388). CE instead separates based on conformational or compositional features of a polyelectrolyte. For a protein, this is largely reflected in conformational differences inherent to species of the same oligomer. This includes different oligomeric states e.g. dimer and hexamer (389), open and closed forms (390), and proteins that have been modified at the amino acid level e.g. via PTM (391). Conformational separation can result in three possible outcomes, i.e. the complete separation of distinct conformations resulting in resolved peaks, peak broadening due to high conformational heterogeneity and a narrow peak due to low conformational heterogeneity (Figure 8). When this CE separation is plotted as a distribution of electrophoretic mobilities (392), the dispersity value for a protein peak can be calculated (386) akin to the 'polydispersity index' or PDI from SEC coupled to multi-angle light scattering. As such, separation can yield complementary and analogous information to that of the common size-based separation technique SEC. This aspect is experimentally examined in Chapter 4 of this thesis.

CE could provide significant advantages for the conformational characterisation of crystallin proteins as its mechanism of separation is underscored by the ability to separate via conformation and charge (393, 394). CE separation could be particularly useful for separating and characterising post-translationally modified crystallins and conformationally heterogeneous proteins such as the α-crystallins. As a prelude to such crystallin studies, model proteins with varying degrees of conformational heterogeneity have been examined in Chapter 4.

Aims of this Thesis

Protein misfolding and aggregation disorders, such as eye lens cataract, cause numerous health, social, and economic issues. However, despite the already substantial body of knowledge on crystallin proteins and their potential contributions to the development of cataract, there is no clear understanding of the biochemical or molecular reason(s) as to why cataract develops. This thesis uses *in vitro* biophysical characterisation to provide a molecular understanding of how oxidation and the crowded lens

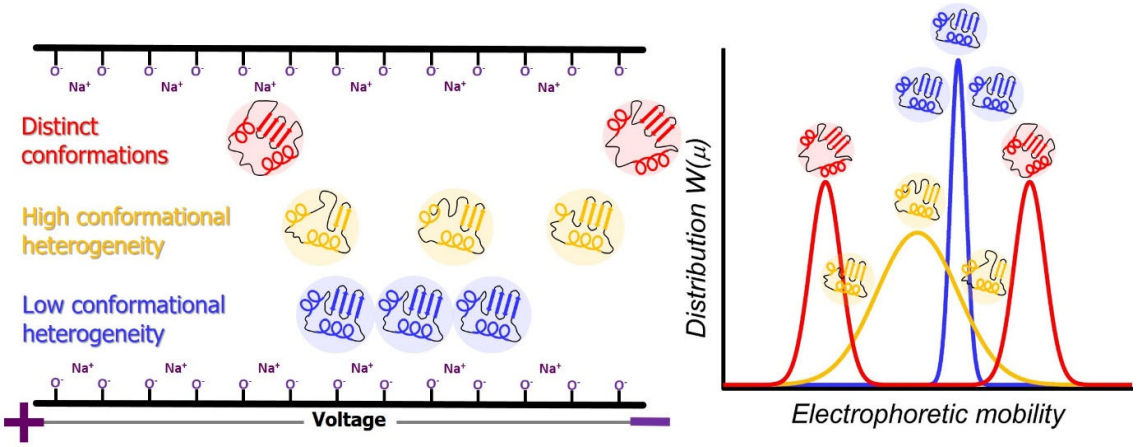


Figure 8. A schematic of a CE separation and resultant distribution of electrophoretic mobilities. Provided a protein has distinct conformational species that exist on a timescale amenable to CE separation, these species will be resolved as two separate peaks (*red*). A protein that exhibits high conformational heterogeneity but does not have conformations distinct enough to be resolved as two separate peaks will present as a broadened CE peak (*yellow*). A protein that exhibits low conformational heterogeneity will present with a sharp CE peak (*blue*). The corresponding separations from each of these samples are compared as a distribution of electrophoretic mobilities which allows calculation of their dispersity value (*right*).

environment impacts crystallin structure, function, and aggregation. It also demonstrates the use of a technique that has not yet been used to characterise crystallin proteins which may provide new information on crystallin heterogeneity in the future. The specific aims are to:

1. Investigate the structural features, biophysical properties and physiological relevance of an oxidised form of γS-crystallin.
2. Use multiple means of mimicking the highly crowded environment of the eye lens and thereby elucidate its effect on the structural and functional features of αB-crystallin.
3. Demonstrate the ability of capillary electrophoresis to characterise the conformational heterogeneity of model proteins as a proof-of-concept for its application towards the conformationally and post-translationally heterogeneous crystallins.

References

1. Baldwin RL. Matching speed and stability. Nature. **1994**;369(6477):183-4.

2. Dobson CM, Šali A, Karplus M. Protein folding: a perspective from theory and experiment. Angew. Chem. Int. Ed. **1998**;37(7):868-93.

3. Anfinsen CB, Haber E, Sela M, White FH. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. Proc. Natl. Acad. Sci. U.S.A. **1961**;47(9):1309-14.

4. Anfinsen CB. Principles that govern the folding of protein chains. Science. **1973**;181(4096):223-30.

5. Levinthal C. Are there pathways for protein folding? J. Chim. Phys. **1968**;65:44-5.

6. Levinthal C. Mossbauer spectroscopy in biological systems: University of Illinois Press, Urbana; **1969**.

7. Dill KA. Theory for the folding and stability of globular proteins. Biochemistry. **1985**;24(6):1501-9.

8. Zwanzig R, Szabo A, Bagchi B. Levinthal's paradox. Proc. Natl. Acad. Sci. U.S.A. **1992**;89(1):20-2.

9. Šali A, Shakhnovich E, Karplus M. How does a protein fold? Nature. **1994**;369(6477):248-51.

10. Radford SE. Protein folding: progress made and promises ahead. Trends Biochem. Sci. **2000**;25(12):611-8.

11. Wolynes P, Onuchic J, Thirumalai D. Navigating the folding routes. Science. **1995**;267(5204):1619-20.

12. Onuchic JN, Wolynes PG, Luthey-Schulten Z, Socci ND. Toward an outline of the topography of a realistic protein-folding funnel. Proc. Natl. Acad. Sci. U.S.A. **1995**;92(8):3626-30.

13. Noé F, Schütte C, Vanden-Eijnden E, Reich L, Weikl TR. Constructing the equilibrium ensemble of folding pathways from short off-equilibrium simulations. Proc. Natl. Acad. Sci. U.S.A. **2009**;106(45):19011-6.

14. Socci ND, Onuchic JN, Wolynes PG. Diffusive dynamics of the reaction coordinate for protein folding funnels. J. Chem. Phys. **1996**;104(15):5860-8.

15. Dill KA, Chan HS. From levinthal to pathways to funnels. Nat. Struct. Biol. **1997**;4(1):10-9.

16. Hartl FU, Bracher A, Hayer-Hartl M. Molecular chaperones in protein folding and proteostasis. Nature. **2011**;475(7356):324-32.

17. Onuchic JN, and ZL-S, Wolynes PG. Theory of protein folding: the energy landscape perspective. Annu. Rev. Phys. Chem. **1997**;48(1):545-600.

18. Dobson CM. Protein folding and misfolding. Nature. **2003**;426(6968):884-90.

19. Dill KA, Ozkan SB, Shell MS, Weikl TR. The protein folding problem. Annu. Rev. Biophys. **2008**;37(1):289-316.

20. Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG. Funnels, pathways, and the energy landscape of protein folding: a synthesis. Proteins: Struct., Funct., Bioinf. **1995**;21(3):167-95.

21. Onuchic JN, Wolynes PG. Theory of protein folding. Curr. Opin. Struct. Biol. **2004**;14(1):70-5.

22. Jahn TR, Radford SE. The yin and yang of protein folding. FEBS J. **2005**;272(23):5962-70.

23. Ecroyd H, Carver JA. Unraveling the mysteries of protein folding and misfolding. IUBMB Life. **2008**;60(12):769-74.

24. Klein WL, Krafft GA, Finch CE. Targeting small Aβ oligomers: the solution to an Alzheimer's disease conundrum? Trends Neurosci. **2001**;24(4):219-24.

25. Benilova I, Karran E, De Strooper B. The toxic Aβ oligomer and Alzheimer's disease: an emperor in need of clothes. Nat. Neurosci. **2012**;15(3):349-57.

26. Bukau B, Horwich AL. The Hsp70 and Hsp60 chaperone machines. Cell. **1998**;92(3):351-66.

27. Kampinga HH, Craig EA. The HSP70 chaperone machinery: J proteins as drivers of functional specificity. Nat. Rev. Mol. Cell Biol. **2010**;11(8):579-92.

28. Horwitz J. Alpha-crystallin can function as a molecular chaperone. Proc. Natl. Acad. Sci. U.S.A. **1992**;89(21):10449-53.

29. Jakob U, Gaestel M, Engel K, Buchner J. Small heat shock proteins are molecular chaperones. J. Biol. Chem. **1993**;268(3):1517-20.

30. Hartl FU, Hayer-Hartl M. Converging concepts of protein folding *in vitro* and *in vivo*. Nat. Struct. Mol. Biol. **2009**;16(6):574-81.

31. Mayer MP, Bukau B. Hsp70 chaperones: cellular functions and molecular mechanism. Cell. Mol. Life Sci. **2005**;62(6):670-84.

32. Mayer MP. Hsp70 chaperone dynamics and molecular mechanism. Trends Biochem. Sci. **2013**;38(10):507-14.

33. Haslbeck M, Franzmann T, Weinfurtner D, Buchner J. Some like it hot: the structure and function of small heat-shock proteins. Nat. Struct. Mol. Biol. **2005**;12(10):842-6.

34. Treweek TM, Meehan S, Ecroyd H, Carver JA. Small heat-shock proteins: important players in regulating cellular proteostasis. Cell. Mol. Life Sci. **2015**;72(3):429-51.

35. Loquet A, El Mammeri N, Stanek J, Berbon M, Bardiaux B, Pintacuda G, *et al.* 3D structure determination of amyloid fibrils using solid-state NMR spectroscopy. Methods. **2018**;138-139:26-38.

36. Fitzpatrick AWP, Saibil HR. Cryo-EM of amyloid fibrils and cellular aggregates. Curr. Opin. Struct. Biol. **2019**;58:34-42.

37. Iadanza MG, Jackson MP, Hewitt EW, Ranson NA, Radford SE. A new era for understanding amyloid structures and disease. Nat. Rev. Mol. Cell Biol. **2018**;19(12):755-73.

38. Colvin MT, Silvers R, Ni QZ, Can TV, Sergeyev I, Rosay M, *et al.* Atomic resolution structure of monomorphic Aβ42 amyloid fibrils. J. Am. Chem. Soc. **2016**;138(30):9663-74.

39. Gremer L, Schölzel D, Schenk C, Reinartz E, Labahn J, Ravelli RBG, *et al.* Fibril structure of amyloid-β(1-42) by cryo-electron microscopy. Science. **2017**;358(6359):116-9.

40. Sunde M, Serpell LC, Bartlam M, Fraser PE, Pepys MB, Blake CCF. Common core structure of amyloid fibrils by synchrotron X-ray diffraction. J. Mol. Biol. **1997**;273(3):729-39.

41. Nelson R, Sawaya MR, Balbirnie M, Madsen AØ, Riekel C, Grothe R, *et al.* Structure of the cross-β spine of amyloid-like fibrils. Nature. **2005**;435(7043):773-8.

42. Tuttle MD, Comellas G, Nieuwkoop AJ, Covell DJ, Berthold DA, Kloepper KD, *et al.* Solid-state NMR structure of a pathogenic fibril of full-length human α-synuclein. Nat. Struct. Mol. Biol. **2016**;23(5):409-15.

43. Eisenberg D, Jucker M. The amyloid state of proteins in human diseases. Cell. **2012**;148(6):1188-203.

44. Sawaya MR, Sambashivan S, Nelson R, Ivanova MI, Sievers SA, Apostol MI, *et al.* Atomic structures of amyloid cross-β spines reveal varied steric zippers. Nature. **2007**;447(7143):453-7.

45. Close W, Neumann M, Schmidt A, Hora M, Annamalai K, Schmidt M, *et al.* Physical basis of amyloid fibril polymorphism. Nature Commun. **2018**;9(1):699.

46. Fändrich M, Nyström S, Nilsson KPR, Böckmann A, LeVine III H, Hammarström P. Amyloid fibril polymorphism: a challenge for molecular imaging and therapy. J. Intern. Med. **2018**;283(3):218-37.

47. Wyatt PJ. Measurement of special nanoparticle structures by light scattering. Anal. Chem. **2014**;86(15):7171-83.

48. Hall D, Zhao R, Dehlsen I, Bloomfield N, Williams SR, Arisaka F, *et al.* Protein aggregate turbidity: simulation of turbidity profiles for mixed-aggregation reactions. Anal. Biochem. **2016**;498:78-94.

49. Lorenzo A, Yankner BA. Beta-amyloid neurotoxicity requires fibril formation and is inhibited by congo red. Proc. Natl. Acad. Sci. U.S.A. **1994**;91(25):12243-7.

50. Taylor CG, Meisl G, Horrocks MH, Zetterberg H, Knowles TPJ, Klenerman D. Extrinsic amyloid-binding dyes for detection of individual protein aggregates in solution. Anal. Chem. **2018**;90(17):10385-93.

51. Smith JF, Knowles TPJ, Dobson CM, MacPhee CE, Welland ME. Characterization of the nanoscale properties of individual amyloid fibrils. Proc. Natl. Acad. Sci. U.S.A. **2006**;103(43):15806-11.

52. Knowles TP, Fitzpatrick AW, Meehan S, Mott HR, Vendruscolo M, Dobson CM, *et al.* Role of intermolecular forces in defining material properties of protein nanofibrils. Science. **2007**;318(5858):1900-3.

53. Taubes G. Protein chemistry: misfolding the way to disease. Science. **1996**;271(5255):1493-5.

54. Buxbaum JN. Diseases of protein conformation: what do *in vitro* experiments tell us about *in vivo* diseases? Trends Biochem. Sci. **2003**;28(11):585-92.

55. Kundu B, Guptasarma P. Use of a hydrophobic dye to indirectly probe the structural organization and conformational plasticity of molecules in amorphous aggregates of carbonic anhydrase. Biochem. Biophys. Res. Commun. **2002**;293(1):572-7.

56. Yoshimura Y, Lin Y, Yagi H, Lee Y-H, Kitayama H, Sakurai K, *et al.* Distinguishing crystal-like amyloid fibrils and glass-like amorphous aggregates from their kinetics of formation. Proc. Natl. Acad. Sci. U.S.A. **2012**;109(36):14446-51.

57. Ross CA, Poirier MA. What is the role of protein aggregation in neurodegeneration? Nat. Rev. Mol. Cell Biol. **2005**;6(11):891-8.

58. Yamamoto S, Gejyo F. Historical background and clinical treatment of dialysis-related amyloidosis. Biochim. Biophys. Acta. **2005**;1753(1):4-10.

59. Rousseau F, Schymkowitz J, Serrano L. Protein aggregation and amyloidosis: confusion of the kinds? Curr. Opin. Struct. Biol. **2006**;16(1):118-26.

60. Boatz JC, Whitley MJ, Li M, Gronenborn AM, van der Wel PCA. Cataract-associated P23T γD-crystallin retains a native-like fold in amorphous-looking aggregates formed at physiological pH. Nature Commun. **2017**;8(1):15137.

61. van der Wel PCA. Insights into protein misfolding and aggregation enabled by solid-state NMR spectroscopy. Solid State Nucl. Magn. Reson. **2017**;88:1-14.

62. Trinh CH, Smith DP, Kalverda AP, Phillips SEV, Radford SE. Crystal structure of monomeric human β-2-microglobulin reveals clues to its amyloidogenic properties. Proc. Natl. Acad. Sci. U.S.A. **2002**;99(15):9771-6.

63. Sipe JD, Cohen AS. Review: history of the amyloid fibril. J. Struct. Biol. **2000**;130(2):88-98.

64. Lansbury PT, Lashuel HA. A century-old debate on protein aggregation and neurodegeneration enters the clinic. Nature. **2006**;443(7113):774-9.

65. Haass C, Selkoe DJ. Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid β-peptide. Nat. Rev. Mol. Cell Biol. **2007**;8(2):101-12.

66. McLaurin J, Chakrabartty A. Membrane disruption by Alzheimer β-amyloid peptides mediated through specific binding to either phospholipids or gangliosides: implications for neurotoxicity. J. Biol. Chem. **1996**;271(43):26482-9.

67. Sciacca Michele FM, Kotler Samuel A, Brender Jeffrey R, Chen J, Lee D-k, Ramamoorthy A. Two-step mechanism of membrane disruption by Aβ through membrane fragmentation and pore formation. Biophys. J. **2012**;103(4):702-10.

68. Bence NF, Sampat RM, Kopito RR. Impairment of the ubiquitin-proteasome system by protein aggregation. Science. **2001**;292(5521):1552-5.

69. Hara T, Nakamura K, Matsui M, Yamamoto A, Nakahara Y, Suzuki-Migishima R, *et al.* Suppression of basal autophagy in neural cells causes neurodegenerative disease in mice. Nature. **2006**;441(7095):885-9.

70. Salminen A, Kaarniranta K, Kauppinen A, Ojala J, Haapasalo A, Soininen H, *et al.* Impaired autophagy and APP processing in Alzheimer's disease: the potential role of Beclin 1 interactome. Prog. Neurobiol. **2013**;106-107:33-54.

71. Ciechanover A, Kwon YT. Degradation of misfolded proteins in neurodegenerative diseases: therapeutic targets and strategies. Exp. Mol. Med. **2015**;47(3):e147.

72. Hashimoto M, Rockenstein E, Crews L, Masliah E. Role of protein aggregation in mitochondrial dysfunction and neurodegeneration in Alzheimer's and Parkinson's diseases. Neuromolecular Med. **2003**;4(1):21-35.

73. Johri A, Beal MF. Mitochondrial dysfunction in neurodegenerative diseases. J. Pharmacol. Exp. Ther. **2012**;342(3):619-30.

74. Olzscha H, Schermann SM, Woerner AC, Pinkert S, Hecht MH, Tartaglia GG, *et al.* Amyloid-like aggregates sequester numerous metastable proteins with essential cellular functions. Cell. **2011**;144(1):67-78.

75. Han S, Kollmer M, Markx D, Claus S, Walther P, Fändrich M. Amyloid plaque structure and cell surface interactions of β-amyloid fibrils revealed by electron tomography. Sci. Rep. **2017**;7(1):43577.

76. Andersen JK. Oxidative stress in neurodegeneration: cause or consequence? Nat. Med. **2004**;10(7):18-25.

77. Niedzielska E, Smaga I, Gawlik M, Moniczewski A, Stankowicz P, Pera J, *et al.* Oxidative stress in neurodegenerative diseases. Mol. Neurobiol. **2016**;53(6):4094-125.

78. Crosby AH. Disruption of cellular transport: a common cause of neurodegeneration? Lancet Neurol. **2003**;2(5):311-6.

79. Woerner AC, Frottin F, Hornburg D, Feng LR, Meissner F, Patra M, *et al.* Cytoplasmic protein aggregates interfere with nucleocytoplasmic transport of protein and RNA. Science. **2016**;351(6269):173-6.

80. Chiti F, Dobson CM. Protein misfolding, functional amyloid, and human disease. Annu. Rev. Biochem. **2006**;75(1):333-66.

81. Sipe JD, Benson MD, Buxbaum JN, Ikeda S-i, Merlini G, Saraiva MJM, *et al.* Amyloid fibril proteins and amyloidosis: chemical identification and clinical classification international society of amyloidosis 2016 nomenclature guidelines. Amyloid. **2016**;23(4):209-13.

82. Australian Bureau of Statistics. 3303.0: causes of death, Australia, 2017. Australian Bureau of Statistics, Canberra, Australia; **2018**.

83. Levine III H. Thioflavine T interaction with synthetic Alzheimer's disease β-amyloid peptides: detection of amyloid aggregation in solution. Protein Sci. **1993**;2(3):404-10.

84. Bloemendal H. The vertebrate eye lens. Science. **1977**;197(4299):127-38.

85. Serebryany E, King JA. The βγ-crystallins: native state stability and pathways to aggregation. Prog. Biophys. Mol. Biol. **2014**;115(1):32-41.

86. Ecroyd H, Carver JA. Crystallin proteins and amyloid fibrils. Cell. Mol. Life Sci. **2009**;66(1):62-81.

87. Zhang TO, Alperstein AM, Zanni MT. Amyloid β-sheet secondary structure identified in UV-induced cataracts of porcine lenses using 2D IR spectroscopy. J. Mol. Biol. **2017**;429(11):1705-21.

88. Alperstein AM, Ostrander JS, Zhang TO, Zanni MT. Amyloid found in human cataracts with two-dimensional infrared spectroscopy. Proc. Natl. Acad. Sci. U.S.A. **2019**;116(14):6602-7.

89. Harding JJ, Dilley KJ. Structural proteins of the mammalian lens: a review with emphasis on changes in development, aging and cataract. Exp. Eye Res. **1976**;22(1):1-73.

90. Truscott RJW, Friedrich MG. The etiology of human age-related cataract. Proteins don't last forever. Biochim. Biophys. Acta, Gen. Subj. **2016**;1860(1, Part B):192-8.

91. Landers J, Henderson T, Craig J. Prevalence and associations of cataract in indigenous Australians within central Australia: the central Australian ocular health study. Clin. Exp. Ophthalmol. **2010**;38(4):387-92.

92. Harding JJ, van Heyningen R. Epidemiology and risk factors for cataract. Eye. **1987**;1(5):537-41.

93. Wang W, Yan W, Fotis K, Prasad NM, Lansingh VC, Taylor HR, *et al.* Cataract surgical rate and socioeconomics: a global study. Invest. Ophthalmol. Vis. Sci. **2017**;57(14):5872-81.

94. Moreau KL, King JA. Protein misfolding and aggregation in cataract disease and prospects for prevention. Trends Mol. Med. **2012**;18(5):273-82.

95. Kels BD, Grzybowski A, Grant-Kels JM. Human ocular anatomy. Clin. Dermatol. **2015**;33(2):140-6.

96. Meek KM, Knupp C. Corneal structure and transparency. Prog. Retin. Eye Res. **2015**;49:1-16.

97. Malhotra A, Minja FJ, Crum A, Burrowes D. Ocular anatomy and cross-sectional imaging of the eye. Semin. Ultrasound CT MRI. **2011**;32(1):2-13.

98. Banh A, Bantseev V, Choh V, Moran KL, Sivak JG. The lens of the eye as a focusing device and its response to stress. Prog. Retin. Eye Res. **2006**;25(2):189-206.

99. Toyama BH, Hetzer MW. Protein homeostasis: live long, won't prosper. Nat. Rev. Mol. Cell Biol. **2013**;14(1):55-61.

100. Stenkamp DL. Development of the vertebrate eye and retina. Prog. Mol. Biol. Transl. Sci. **2015**;134:397-414.

101. Snell RS, Lemp MA. Clinical anatomy of the eye. Second Edition. Wiley-Blackwell; **2013**.

102. Hejtmancik JF, Shiels A. Overview of the lens. Prog. Mol. Biol. Transl. Sci. **2015**;134:119-27.

103. McAvoy JW. Induction of the eye lens. Differentiation. **1980**;17(1):137-49.

104. Francis PJ, Berry V, Moore AT, Bhattacharya S. Lens biology: Development and human cataractogenesis. Trends Genet. **1999**;15(5):191-6.

105. Martinez G, de Iongh RU. The lens epithelium in ocular health and disease. Int. J. Biochem. Cell Biol. **2010**;42(12):1945-63.

106. Wride MA. Lens fibre cell differentiation and organelle loss: many paths lead to clarity. Philos. Trans. R. Soc. B Biol. Sci. **2011**;366(1568):1219-33.

107. Bassnett S, Šikić H. The lens growth process. Prog. Retin. Eye Res. **2017**;60:181-200.

108. Augusteyn RC. On the growth and internal structure of the human lens. Exp. Eye Res. **2010**;90(6):643-54.

109. Harding JJ, Rixon KC, Marriott FHC. Men have heavier lenses than women of the same age. Exp. Eye Res. **1977**;25(6):651.

110. Augusteyn RC. Growth of the human eye lens. Mol. Vis. **2007**;13:252-7.

111. Delaye M, Tardieu A. Short-range order of crystallin proteins accounts for eye lens transparency. Nature. **1983**;302(5907):415-7.

112. Bassnett S. Lens organelle degradation. Exp. Eye Res. **2002**;74(1):1-6.

113. Ponce A, Sorensen C, Takemoto L. Role of short-range protein interactions in lens opacifications. Mol. Vis. **2006**;12:879-84.

114. Bassnett S. On the mechanism of organelle degradation in the vertebrate lens. Exp. Eye Res. **2009**;88(2):133-9.

115. Bloemendal H, de Jong W, Jaenicke R, Lubsen NH, Slingsby C, Tardieu A. Ageing and vision: structure, stability and function of lens crystallins. Prog. Biophys. Mol. Biol. **2004**;86(3):407-85.

116. Bassnett S, Shi Y, Vrensen GFJM. Biological glass: structural determinants of eye lens transparency. Philos. Trans. R. Soc. B Biol. Sci. **2011**;366(1568):1250-64.

117. Bassnett S, Costello MJ. The cause and consequence of fiber cell compaction in the vertebrate lens. Exp. Eye Res. **2017**;156:50-7.

118. Benedek GB, Pande J, Thurston GM, Clark JI. Theoretical and experimental basis for the inhibition of cataract. Prog. Retin. Eye Res. **1999**;18(3):391-402.

119. Piatigorsky J. Lens differentiation in vertebrates. Differentiation. **1981**;19(1-3):134-53.

120. Bours J. Isoelectric focusing of lens crystallins in thin-layer polyacrylamide gels: a method for detection of soluble proteins in eye lens extracts. J. Chromatogr. **1971**;60:225-33.

121. Spector A. Methods of isolation of alpha, beta, and gamma crystallins and their subgroups. Invest. Ophthalmol. Vis. Sci. **1964**;3(2):182-93.

122. Slingsby C, Clout NJ. Structure of the crystallins. Eye. **1999**;13(3):395-402.

123. Lampi KJ, Ma Z, Shih M, Shearer TR, Smith JB, Smith DL, *et al.* Sequence analysis of βA3, βB3, and βA4 crystallins completes the identification of the major proteins in young human lens. J. Biol. Chem. **1997**;272(4):2268-75.

124. Graw J. Genetics of crystallins: cataract and beyond. Exp. Eye Res. **2009**;88(2):173-89.

125. Ma Z, Hanson SRA, Lampi KJ, David LL, Smith DL, Smith JB. Age-related changes in human lens crystallins identified by HPLC and mass spectrometry. Exp. Eye Res. **1998**;67(1):21-30.

126. Bhat SP, Nagineni CN. αB subunit of lens-specific protein α-crystallin is present in other ocular and non-ocular tissues. Biochem. Biophys. Res. Commun. **1989**;158(1):319-25.

127. Piatigorsky J, Wistow GJ. The recruitment of crystallins: new functions precede gene duplication. Science. **1991**;252(5009):1078-9.

128. Aquilina JA, Benesch JLP, Bateman OA, Slingsby C, Robinson CV. Polydispersity of a mammalian chaperone: mass spectrometry reveals the population of oligomers in αB-crystallin. Proc. Natl. Acad. Sci. U.S.A. **2003**;100(19):10611-6.

129. Horwitz J. Alpha-crystallin. Exp. Eye Res. **2003**;76(2):145-53.

130. Laganowsky A, Benesch JLP, Landau M, Ding L, Sawaya MR, Cascio D, *et al.* Crystal structures of truncated alphaA and alphaB crystallins reveal structural mechanisms of polydispersity important for eye lens function. Protein Sci. **2010**;19(5):1031-43.

131. Carver JA, Aquilina JA, Truscott RJW, Ralston GB. Identification by 1H NMR spectroscopy of flexible C-terminal extensions in bovine lens α-crystallin. FEBS Lett. **1992**;311(2):143-9.

132. Treweek TM, Rekas A, Walker MJ, Carver JA. A quantitative NMR spectroscopic examination of the flexibility of the C-terminal extensions of the molecular chaperones, αA- and αB-crystallin. Exp. Eye Res. **2010**;91(5):691-9.

133. Horwitz J. Alpha crystallin: The quest for a homogeneous quaternary structure. Exp. Eye Res. **2009**;88(2):190-4.

134. Mymrikov EV, Seit-Nebi AS, Gusev NB. Heterooligomeric complexes of human small heat shock proteins. Cell Stress Chaperon. **2012**;17(2):157-69.

135. Bova MP, Ding L-L, Horwitz J, Fung BK-K. Subunit exchange of αA-crystallin. J. Biol. Chem. **1997**;272(47):29511-7.

136. Delbecq SP, Rosenbaum JC, Klevit RE. A mechanism of subunit recruitment in human small heat shock protein oligomers. Biochemistry. **2015**;54(28):4276-84.

137. Jehle S, Vollmar BS, Bardiaux B, Dove KK, Rajagopal P, Gonen T, *et al.* N-terminal domain of αB-crystallin provides a conformational switch for multimerization and structural heterogeneity. Proc. Natl. Acad. Sci. U.S.A. **2011**;108(16):6409-14.

138. van Montfort RLM, Bateman OA, Lubsen NH, Slingsby C. Crystal structure of truncated human βB1-crystallin. Protein Sci. **2003**;12(11):2606-12.

139. Smith MA, Bateman OA, Jaenicke R, Slingsby C. Mutation of interfaces in domain-swapped human βB2-crystallin. Protein Sci. **2007**;16(4):615-25.

140. Basak A, Bateman O, Slingsby C, Pande A, Asherie N, Ogun O, *et al.* High-resolution X-ray crystal structures of human γD crystallin (1.25 Å) and the R58H mutant (1.15 Å) associated with aculeiform cataract. J. Mol. Biol. **2003**;328(5):1137-47.

141. Kingsley Carolyn N, Brubaker William D, Markovic S, Diehl A, Brindley Amanda J, Oschkinat H, *et al.* Preferential and specific binding of human αB-crystallin to a cataract-related variant of γS-crystallin. Structure. **2013**;21(12):2221-7.

142. Mainz A, Peschek J, Stavropoulou M, Back KC, Bardiaux B, Asami S, *et al.* The chaperone αB-crystallin uses different interfaces to capture an amorphous and an amyloid client. Nat. Struct. Mol. Biol. **2015**;22(11):898-905.

143. Cox D, Selig E, Griffin MDW, Carver JA, Ecroyd H. Small heat-shock proteins prevent α-synuclein aggregation via transient interactions and their efficacy is affected by the rate of aggregation. J. Biol. Chem. **2016**;291(43):22618-29.

144. McAvoy JW. Cell division, cell elongation and distribution of α-, β- and γ-crystallins in the rat lens. J. Embryol. Exp. Morphol. **1978**;44(1):149-65.

145. Roman SG, Chebotareva NA, Eronina TB, Kleymenov SY, Makeeva VF, Poliansky NB, *et al.* Does the crowded cell-like environment reduce the chaperone-like activity of α-crystallin? Biochemistry. **2011**;50(49):10607-23.

146. Takemoto L, Sorensen CM. Protein–protein interactions and lens transparency. Exp. Eye Res. **2008**;87(6):496-501.

147. Zhao H, Brown PH, Magone MT, Schuck P. The molecular refractive function of lens γ-crystallins. J. Mol. Biol. **2011**;411(3):680-99.

148. Driessen HPC, Herbrink P, Bloemendal H, de Jong WW. Primary structure of the bovine β-crystallin bp chain. Eur. J. Biochem. **1981**;121(1):83-91.

149. Slingsby C, Wistow GJ, Clark AR. Evolution of crystallins for a role in the vertebrate eye lens. Protein Sci. **2013**;22(4):367-80.

150. Purkiss AG, Bateman OA, Goodfellow JM, Lubsen NH, Slingsby C. The X-ray crystal structure of human γS-crystallin C-terminal domain. J. Biol. Chem. **2002**;277(6):4199-205.

151. Lubsen NH, Aarts HJM, Schoenmakers JGG. The evolution of lenticular proteins: the β- and γ-crystallin super gene family. Prog. Biophys. Mol. Biol. **1988**;51(1):47-76.

152. Cooper PG, Carver JA, Truscott RJW. 1H-NMR spectroscopy of bovine lens β-crystallin. Eur. J. Biochem. **1993**;213(1):321-8.

153. Carver JA, Cooper PG, Truscott RJW. 1H-NMR spectroscopy of βB2-crystallin from bovine eye lens. Eur. J. Biochem. **1993**;213(1):313-20.

154. Werten PJL, Lindner RA, Carver JA, de Jong WW. Formation of βA3/βB2-crystallin mixed complexes: involvement of N- and C-terminal extensions. Biochim. Biophys. Acta. **1999**;1432(2):286-92.

155. Bax B, Lapatto R, Nalini V, Driessen H, Lindley PF, Mahadevan D, *et al.* X-ray analysis of βB2-crystallin and evolution of oligomeric lens proteins. Nature. **1990**;347(6295):776-80.

156. Xi Z, Whitley MJ, Gronenborn AM. Human βB2-crystallin forms a face-en-face dimer in solution: an integrated NMR and SAXS study. Structure. **2017**;25(3):496-505.

157. Cvekl A, Piatigorsky J. Lens development and crystallin gene expression: many roles for Pax-6. Bioessays. **1996**;18(8):621-30.

158. Aarts HJM, Lubsen NH, Schoenmakers JGG. Crystallin gene expression during rat lens development. Eur. J. Biochem. **1989**;183(1):31-6.

159. Ueda Y, Duncan MK, David LL. Lens proteomics: the accumulation of crystallin modifications in the mouse lens with age. Invest. Ophthalmol. Vis. Sci. **2002**;43(1):205-15.

160. Lampi KJ, Ma Z, Hanson SRA, Azuma M, Shih M, Shearer TR, *et al.* Age-related changes in human lens crystallins identified by two-dimensional electrophoresis and mass spectrometry. Exp. Eye Res. **1998**;67(1):31-43.

161. Lapko VN, Smith DL, Smith JB. Expression of βA2-crystallin in human lenses. Exp. Eye Res. **2003**;77(3):383-5.

162. Bindels JG, Koppers A, Hoenders HJ. Structural aspects of bovine β-crystallins: physical characterization including dissociation-association behavior. Exp. Eye Res. **1981**;33(3):333-43.

163. Berbers GAM, Boerman OC, Bloemendal H, de Jong WW. Primary gene products of bovine β-crystallin and reassociation behavior of its aggregates. Eur. J. Biochem. **1982**;128(2-3):495-502.

164. Vérétout F, Tardieu A. The protein concentration gradient within eye lens might originate from constant osmotic pressure coupled to differential interactive properties of crystallins. Eur. Biophys. J. **1989**;17(2):61-8.

165. Ajaz MS, Ma Z, Smith DL, Smith JB. Size of human lens β-crystallin aggregates are distinguished by N-terminal truncation of βB1. J. Biol. Chem. **1997**;272(17):11250-5.

166. Zarina S, Abbasi A, Zaidi ZH. Primary structure of βS-crystallin from human lens. Biochem. J. **1992**;287(2):375-81.

167. Cooper PG, Carver JA, Aquilina JA, Ralston GB, Truscott RJW. A 1H NMR spectroscopic comparison of γS- and γB-crystallins. Exp. Eye Res. **1994**;59(2):211-20.

168. Quax-Jeuken Y, Driessen H, Leunissen J, Quax W, de Jong W, Bloemendal H. βS-crystallin: structure and evolution of a distinct member of the βγ-superfamily. EMBO J. **1985**;4(10):2597-602.

169. Smith J, Yang Z, Lin P, Zaidi Z, Abbasi A, Russell P. The complete sequence of human lens γS-crystallin. Biochem. J. **1995**;307(2):407-10.

170. Vendra VPR, Khan I, Chandani S, Muniyandi A, Balasubramanian D. Gamma crystallins of the human eye lens. Biochim. Biophys. Acta, Gen. Subj. **2016**;1860(1, Part B):333-43.

171. Quintanar L, Domínguez-Calva JA, Serebryany E, Rivillas-Acevedo L, Haase-Pettingell C, Amero C, *et al.* Copper and zinc ions specifically promote nonamyloid aggregation of the highly stable human γ-D crystallin. ACS Chem. Biol. **2016**;11(1):263-72.

172. Domínguez-Calva JA, Pérez-Vázquez ML, Serebryany E, King JA, Quintanar L. Mercury-induced aggregation of human lens γ-crystallins reveals a potential role in cataract disease. J. Biol. Inorg. Chem. **2018**;23(7):1105-18.

173. Thorn DC, Grosas AB, Mabbitt PD, Ray NJ, Jackson CJ, Carver JA. The structure and stability of the disulfide-linked γS-crystallin dimer provide insight into oxidation products associated with lens cataract formation. J. Mol. Biol. **2019**;431(3):483-97.

174. Peek R, McAvoy JW, Lubsen NH, Schoenmakers JGG. Rise and fall of crystallin gene messenger levels during fibroblast growth factor induced terminal differentiation of lens cells. Dev. Biol. **1992**;152(1):152-60.

175. Thomson JA, Augusteyn RC. Ontogeny of human lens crystallins. Exp. Eye Res. **1985**;40(3):393-410.

176. Siezen RJ, Thomson JA, Kaplan ED, Benedek GB. Human lens γ-crystallins: isolation, identification, and characterization of the expressed gene products. Proc. Natl. Acad. Sci. U.S.A. **1987**;84(17):6088-92.

177. Siezen RJ, Wu E, Kaplan ED, Thomson JA, Benedek GB. Rat lens γ-crystallins: characterization of the six gene products and their spatial and temporal distribution resulting from differential synthesis. J. Mol. Biol. **1988**;199(3):475-90.

178. Van Rens GLM, De Jong WW, Bloemendal H. One member of the γ-crystallin gene family, γS, is expressed in birds. Exp. Eye Res. **1991**;53(1):135-8.

179. Hanson SRA, Smith DL, Smith JB. Deamidation and disulfide bonding in human lens γ-crystallins. Exp. Eye Res. **1998**;67(3):301-12.

180. Liu Y-C, Wilkins M, Kim T, Malyugin B, Mehta JS. Cataracts. Lancet. **2017**;390(10094):600-12.

181. Wirth MG, Russell-Eggitt IM, Craig JE, Elder JE, Mackey DA. Aetiology of congenital and paediatric cataract in an Australian population. Br. J. Ophthalmol. **2002**;86(7):782-6.

182. Reddy MA, Francis PJ, Berry V, Bhattacharya SS, Moore AT. Molecular genetic basis of inherited cataract and associated phenotypes. Surv. Ophthalmol. **2004**;49(3):300-15.

183. Asbell PA, Dualan I, Mindel J, Brocks D, Ahmad M, Epstein S. Age-related cataract. Lancet. **2005**;365(9459):599-609.

184. Ionides A, Francis P, Berry V, Mackay D, Bhattacharya S, Shiels A, *et al.* Clinical and genetic heterogeneity in autosomal dominant cataract. Br. J. Ophthalmol. **1999**;83(7):802-8.

185. Guercio JR, Martyn LJ. Congenital malformations of the eye and orbit. Otolaryng. Clin. N. Am. **2007**;40(1):113-40.

186. Vicart P, Caron A, Guicheney P, Li Z, Prévost M-C, Faure A, *et al.* A missense mutation in the αB-crystallin chaperone gene causes a desmin-related myopathy. Nat. Genet. **1998**;20(1):92-5.

187. Selcen D, Engel AG. Myofibrillar myopathy caused by novel dominant negative αB-crystallin mutations. Ann. Neurol. **2003**;54(6):804-10.

188. Hochberg GKA, Ecroyd H, Liu C, Cox D, Cascio D, Sawaya MR, *et al.* The structured core domain of αB-crystallin can prevent amyloid fibrillation and associated toxicity. Proc. Natl. Acad. Sci. U.S.A. **2014**;111(16):1562-70.

189. Bova MP, Yaron O, Huang Q, Ding L, Haley DA, Stewart PL, *et al.* Mutation R120G in αB-crystallin, which is linked to a desmin-related myopathy, results in an irregular structure and defective chaperone-like function. Proc. Natl. Acad. Sci. U.S.A. **1999**;96(11):6137-42.

190. Treweek TM, Rekas A, Lindner RA, Walker MJ, Aquilina JA, Robinson CV, *et al.* R120G αB-crystallin promotes the unfolding of reduced α-lactalbumin and is inherently unstable. FEBS J. **2005**;272(3):711-24.

191. Yao K, Tang X, Shentu X, Wang K, Rao H, Xia K. Progressive polymorphic congenital cataract caused by a CRYBB2 mutation in a chinese family. Mol. Vis. **2005**;11(6):758-63.

192. Pras E, Frydman M, Levy–Nissenbaum E, Bakhan T, Raz J, Assia EI, *et al.* A nonsense mutation (W9X) in CRYAA causes autosomal recessive cataract in an inbred Jewish Persian family. Invest. Ophthalmol. Vis. Sci. **2000**;41(11):3511-5.

193. Hansen L, Yao W, Eiberg H, Kjaer KW, Baggesen K, Hejtmancik JF, *et al.* Genetic heterogeneity in microcornea-cataract: five novel mutations in CRYAA, CRYGD, and GJA8. Invest. Ophthalmol. Vis. Sci. **2007**;48(9):3937-44.

194. Devi RR, Yao W, Vijayalakshmi P, Sergeev YV, Sundaresan P, Hejtmancik JF. Crystallin gene mutations in indian families with inherited pediatric cataract. Mol. Vis. **2008**;14:1157-70.

195. Graw J, Klopp N, Illig T, Preising MN, Lorenz B. Congenital cataract and macular hypoplasia in humans associated with a *de novo* mutation in CRYAA and compound heterozygous mutations in P. Graefe's Arch. Clin. Exp. Ophthalmol. **2006**;244(8):912-9.

196. Mackay DS, Andley UP, Shiels A. Cell death triggered by a novel mutation in the alphaA-crystallin gene underlies autosomal dominant cataract linked to chromosome 21q. Europ. J. Hum. Genet. **2003**;11(10):784-93.

197. Khan AO, Aldahmesh MA, Meyer B. Recessive congenital total cataract with microcornea and heterozygote carrier signs caused by a novel missense CRYAA mutation (R54C). Am. J. Ophthalmol. **2007**;144(6):949-52.

198. Santhiya ST, Soker T, Klopp N, Illig T, Prakash M, Selvaraj B, *et al.* Identification of a novel, putative cataract-causing allele in CRYAA (G98R) in an Indian family. Mol. Vis. **2006**;12(6):768-73.

199. Litt M, Kramer P, LaMorticella DM, Murphey W, Lovrien EW, Weleber RG. Autosomal dominant congenital cataract associated with a missense mutation in the human alpha crystallin gene CRYAA. Hum. Mol. Genet. **1998**;7(3):471-4.

200. Vanita V, Singh JR, Hejtmancik JF, Nürnberg P, Hennies HC, Singh D, *et al.* A novel fan-shaped cataract-microcornea syndrome caused by a mutation of CRYAA in an Indian family. Mol. Vis. **2006**;12:518-22.

201. Beby F, Commeaux C, Bozon M, Denis P, Edery P, Morlé L. New phenotype associated with an Arg116Cys mutation in the CRYAA gene: nuclear cataract, iris coloboma, and microphthalmia. Arch. Ophthalmol. **2007**;125(2):213-6.

202. Richter L, Flodman P, Barria von-Bischhoffshausen F, Burch D, Brown S, Nguyen L, *et al.* Clinical variability of autosomal dominant cataract, microcornea and corneal opacity and novel mutation in the alpha A crystallin gene (CRYAA). Am. J. Med. Genet. A. **2008**;146A(7):833-42.

203. Gu F, Luo W, Li X, Wang Z, Lu S, Zhang M, *et al.* A novel mutation in alphaA-crystallin (CRYAA) caused autosomal dominant congenital cataract in a large Chinese family. Hum. Mutat. **2008**;29(5):769-76.

204. Zhang L, Zhang Y, Liu P, Cao W, Tang X, Su S. Congenital anterior polar cataract associated with a missense mutation in the human alpha crystallin gene CRYAA. Mol. Vis. **2011**;17:2693-7.

205. Chen Q, Ma J, Yan M, Mothobi ME, Liu Y, Zheng F. A novel mutation in CRYAB associated with autosomal dominant congenital nuclear cataract in a Chinese family. Mol. Vis. **2009**;15:1359-65.

206. Liu M, Ke T, Wang Z, Yang Q, Chang W, Jiang F, *et al.* Identification of a CRYAB mutation associated with autosomal dominant posterior polar cataract in a Chinese family. Invest. Ophthalmol. Vis. Sci. **2006**;47(8):3461-6.

207. Safieh LA, Khan A, Alkuraya F. Identification of a novel CRYAB mutation associated with autosomal recessive juvenile cataract in a Saudi family. Mol. Vis. **2009**;15:980-4.

208. Khan AO, Safieh LA, Alkuraya FS. Later retinal degeneration following childhood surgical aphakia in a family with recessive CRYAB mutation (p.R56W). Ophthalmic Genet. **2010**;31(1):30-6.

209. Fichna JP, Potulska-Chromik A, Miszta P, Redowicz MJ, Kaminska AM, Zekanowski C, *et al.* A novel dominant D109A CRYAB mutation in a family with myofibrillar myopathy affects αB-crystallin structure. BBA Clin. **2017**;7:1-7.

210. Sacconi S, Féasson L, Antoine JC, Pécheux C, Bernard R, Cobo AM, *et al.* A novel CRYAB mutation resulting in multisystemic disease. Neuromuscular Disord. **2012**;22(1):66-72.

211. Berry V, Francis P, Reddy MA, Collyer D, Vithana E, MacKay I, *et al.* Alpha-B crystallin gene (CRYAB) mutation causes dominant congenital posterior polar cataract in humans. Am. J. Hum. Genet. **2001**;69(5):1141-5.

212. Liu B-F, Liang J. Interaction and biophysical properties of human lens Q155* βB2-crystallin mutant. Mol Vis. **2005**;11:321-7.

213. Ma Z, Piszczek G, Wingfield PT, Sergeev YV, Hejtmancik JF. The G18V CRYGS mutation associated with human cataracts increases γS-crystallin sensitivity to thermal and chemical stress. Biochemistry. **2009**;48(30):7334-41.

214. Khago D, Wong EK, Kingsley CN, Alfredo Freites J, Tobias DJ, Martin RW. Increased hydrophobic surface exposure in the cataract-related G18V variant of human γS-crystallin. Biochim. Biophys. Acta, Gen. Subj. **2016**;1860(1, Part B):325-32.

215. Huang K-Y, Kingsley CN, Sheil R, Cheng C-Y, Bierma JC, Roskamp KW, *et al.* Stability of protein-specific hydration shell on crowding. J. Am. Chem. Soc. **2016**;138(16):5392-402.

216. Ji F, Jung J, Koharudin LMI, Gronenborn AM. The human W42R γD-crystallin mutant structure provides a link between congenital and age-related cataracts. J. Biol. Chem. **2013**;288(1):99-109.

217. Bharat SV, Shekhtman A, Pande J. The cataract-associated V41M mutant of human γS-crystallin shows specific structural changes that directly enhance local surface hydrophobicity. Biochem. Biophys. Res. Commun. **2014**;443(1):110-4.

218. Xu J, Wang S, Zhao W-J, Xi Y-B, Yan Y-B, Yao K. The congenital cataract-linked A2V mutation impairs tetramer formation and promotes aggregation of βB2-crystallin. PLoS One. **2012**;7(12):e51200.

219. Jung J, Byeon I-JL, Wang Y, King J, Gronenborn AM. The structure of the cataract-causing P23T mutant of human γD-crystallin exhibits distinctive local conformational and dynamic changes. Biochemistry. **2009**;48(12):2597-609.

220. Brubaker WD, Freites JA, Golchert KJ, Shapiro RA, Morikis V, Tobias Douglas J, *et al.* Separating instability from aggregation propensity in γS-crystallin variants. Biophys. J. **2011**;100(2):498-506.

221. Singh D, Raman B, Ramakrishna T, Rao CM. Mixed oligomer formation between human αA-crystallin and its cataract-causing G98R mutant: structural, stability and functional differences. J. Mol. Biol. **2007**;373(5):1293-304.

222. Fu L, Liang JJN. Alteration of protein–protein interactions of congenital cataract crystallin mutants. Invest. Ophthalmol. Vis. Sci. **2003**;44(3):1155-9.

223. Banerjee PR, Pande A, Patrosz J, Thurston GM, Pande J. Cataract-associated mutant E107A of human γD-crystallin shows increased attraction to α-crystallin and enhanced light scattering. Proc. Natl. Acad. Sci. U.S.A. **2011**;108(2):574-9.

224. Pande A, Annunziata O, Asherie N, Ogun O, Benedek GB, Pande J. Decrease in protein solubility and cataract formation caused by the Pro23 to Thr mutation in human γD-crystallin. Biochemistry. **2005**;44(7):2491-500.

225. Zhao W-J, Xu J, Chen X-J, Liu H-H, Yao K, Yan Y-B. Effects of cataract-causing mutations W59C and W151C on βB2-crystallin structure, stability and folding. Int. J. Biol. Macromol. **2017**;103:764-70.

226. Li H, Li C, Lu Q, Su T, Ke T, Li DW-C, *et al.* Cataract mutation P20S of αB-crystallin impairs chaperone activity of αA-crystallin and induces apoptosis of human lens epithelial cells. Biochim. Biophys. Acta, Mol. Basis Dis. **2008**;1782(5):303-9.

227. Ji F, Jung J, Gronenborn AM. Structural and biochemical characterization of the childhood cataract-associated R76S mutant of human γD-crystallin. Biochemistry. **2012**;51(12):2588-96.

228. Ferrini W, Schorderet DF, Othenin-Girard P, Uffer S, Héon E, Munier FL. CRYBA3/A1 gene mutation associated with suture-sparing autosomal dominant congenital nuclear cataract: a novel phenotype. Invest. Ophthalmol. Vis. Sci. **2004**;45(5):1436-41.

229. Reddy MA, Bateman OA, Chakarova C, Ferris J, Berry V, Lomas E, *et al.* Characterization of the G91del CRYBA1/3-crystallin protein: a cause of human inherited cataract. Hum. Mol. Genet. **2004**;13(9):945-53.

230. Qi Y, Jia H, Huang S, Lin H, Gu J, Su H, *et al.* A deletion mutation in the βA1/A3 crystallin gene (CRYBA1/A3) is associated with autosomal dominant congenital nuclear cataract in a Chinese family. Hum. Genet. **2004**;114(2):192-7.

231. Lu S, Zhao C, Jiao H, Kere J, Tang X, Zhao F, *et al.* Two Chinese families with pulverulent congenital cataracts and ΔG91 CRYBA1 mutations. Mol. Vis. **2007**;13:1154-60.

232. Kannabiran C, Rogan PK, Olmos L, Basti S, Rao GN, Kaiser-Kupfer M, *et al.* Autosomal dominant zonular cataract with sutural opacities is associated with a splice mutation in the βA3/A1-crystallin gene. Mol. Vis. **1998**;4:21-6.

233. Bateman JB, Geyer DD, Flodman P, Johannes M, Sikela J, Walter N, *et al.* A new βA1-crystallin splice junction mutation in autosomal dominant cataract. Invest. Ophthalmol. Vis. Sci. **2000**;41(11):3278-85.

234. Reis LM, Tyler RC, Muheisen S, Raggio V, Salviati L, Han DP, *et al.* Whole exome sequencing in dominant cataract identifies a new causative factor, CRYBA2, and a variety of novel alleles in known genes. Hum. Genet. **2013**;132(7):761-70.

235. Zhou G, Zhou N, Hu S, Zhao L, Zhang C, Qi Y. A missense mutation in CRYBA4 associated with congenital cataract and microcornea. Mol. Vis. **2010**;16:1019-24.

236. Billingsley G, Santhiya ST, Paterson AD, Ogata K, Wodak S, Hosseini SM, *et al.* CRYBA4, a novel human cataract gene, is also involved in microphthalmia. Am. J. Hum. Genet. **2006**;79(4):702-9.

237. Mackay DS, Boskovska OB, Knopf HLS, Lampi KJ, Shiels A. A nonsense mutation in CRYBB1 associated with autosomal dominant cataract linked to human chromosome 22q. Am. J. Hum. Genet. **2002**;71(5):1216-21.

238. Cohen D, Bar-Yosef U, Levy J, Gradstein L, Belfair N, Ofir R, *et al.* Homozygous CRYBB1 deletion mutation underlies autosomal recessive congenital cataract. Invest. Ophthalmol. Vis. Sci. **2007**;48(5):2208-13.

239. Jin A, Zhang Y, Xiao D, Xiang M, Jin K, Zeng M. A novel mutation p.S93R in CRYBB1 associated with dominant congenital cataract and microphthalmia. Curr. Eye Res. **2019**:1-7.

240. Chen P, Chen H, Pan X-J, Tang S-Z, Xia Y-J, Zhang H. Novel mutations in CRYBB1/CRYBB2 identified by targeted exome sequencing in Chinese families with congenital cataract. Int. J. Ophthalmol. **2018**;11(10):1577-82.

241. Ma AS, Grigg JR, Ho G, Prokudin I, Farnsworth E, Holman K, *et al.* Sporadic and familial congenital cataracts: mutational spectrum and new diagnoses using next-generation sequencing. Hum. Mutat. **2016**;37(4):371-84.

242. Wang KJ, Wang S, Cao N-Q, Yan Y-B, Zhu SQ. A novel mutation in CRYBB1 associated with congenital cataract-microcornea syndrome: The p.Ser129Arg mutation destabilizes the βB1/βA3-crystallin heteromer but not the βB1-crystallin homomer. Hum. Mutat. **2011**;32(3):2050-60.

243. Siggs OM, Javadiyan S, Sharma S, Souzeau E, Lower KM, Taranath DA, *et al.* Partial duplication of the CRYBB1-CRYBA4 locus is associated with autosomal dominant congenital cataract. Europ. J. Hum. Genet. **2017**;25(6):711-8.

244. Yang J, Zhu Y, Gu F, He X, Cao Z, Li X, *et al.* A novel nonsense mutation in CRYBB1 associated with autosomal dominant congenital cataract. Mol. Vis. **2008**;14:727-31.

245. Rao Y, Dong S, Li Z, Yang G, Peng C, Yan M, *et al.* A novel truncation mutation in CRYBB1 associated with autosomal dominant congenital cataract with nystagmus. Mol. Vis. **2017**;23:624-37.

246. Jun W, Xu M, Feng G, Liu N-p, Hao X-l, Wang K-j, *et al.* A missense mutation S228P in the CRYBB1 gene causes autosomal dominant congenital cataract. Chinese Med. J. **2007**;120(9):820-4.

247. Sun Z, Zhou Q, Li H, Yang L, Wu S, Sui R. Mutations in crystallin genes result in congenital cataract associated with other ocular abnormalities. Mol. Vis. **2017**;23:977-86.

248. Wang KJ, Wang BB, Zhang F, Zhao Y, Ma X, Zhu SQ. Novel β-crystallin gene mutations in Chinese families with nuclear cataracts. Arch. Ophthalmol. **2011**;129(3):337-43.

249. Willoughby CE, Shafiq A, Ferrini W, Chan L, Billingsley G, Priston M, *et al.* CRYBB1 mutation associated with congenital cataract and microcornea. Mol. Vis. **2005**;11:587-93.

250. Yao K, Li J, Jin C, Wang W, Zhu Y, Shentu X, *et al.* Characterization of a novel mutation in the CRYBB2 gene associated with autosomal dominant congenital posterior subcapsular cataract in a Chinese family. Mol. Vis. **2011**;17:144-52.

251. Bateman JB, von-Bischhoffshaunsen FRB, Richter L, Flodman P, Burch D, Spence MA. Gene conversion mutation in crystallin, β-B2 (CRYBB2) in a Chilean family with autosomal dominant cataract. Ophthalmology. **2007**;114(3):425-32.

252. Wang L, Lin H, Gu J, Su H, Huang S, Qi Y. Autosomal-dominant cerulean cataract in a Chinese family associated with gene conversion mutation in Beta-B2-crystallin. Ophthalmic Res. **2009**;41(3):148-53.

253. Riazuddin SA, Yasmeen A, Yao W, Sergeev YV, Zhang Q, Zulfiqar F, *et al.* Mutations in βB3-crystallin associated with autosomal recessive cataract in two Pakistani families. Invest. Ophthalmol. Vis. Sci. **2005**;46(6):2100-6.

254. Héon E, Priston M, Schorderet DF, Billingsley GD, Girard PO, Lubsen N, *et al.* The γ-crystallins and human cataracts: a puzzle made clearer. Am. J. Hum. Genet. **1999**;65(5):1261-7.

255. Kumar M, Agarwal T, Khokhar S, Kumar M, Kaur P, Roy TS, *et al.* Mutation screening and genotype phenotype correlation of α-crystallin, γ-crystallin and GJA8 gene in congenital cataract. Mol. Vis. **2011**;17:693-707.

256. Ren Z, Li A, Shastry BS, Padma T, Ayyagari R, Scott MH, *et al.* A 5-base insertion in the γC-crystallin gene is associated with autosomal dominant variable zonular pulverulent cataract. Hum. Genet. **2000**;106(5):531-7.

257. Yao K, Jin C, Zhu N, Wang W, Wu R, Jiang J, *et al.* A nonsense mutation in CRYGC associated with autosomal dominant congenital nuclear cataract in a Chinese family. Mol. Vis. **2008**;14:1272-6.

258. Santana A, Waiswol M, Arcieri ES, de Vasconcellos JPC, de Melo MB. Mutation analysis of CRYAA, CRYGC, and CRYGD associated with autosomal dominant congenital cataract in Brazilian families. Mol. Vis. **2009**;15:793-800.

259. Li X-Q, Cai H-C, Zhou S-Y, Yang J-H, Xi Y-B, Gao X-B, *et al.* A novel mutation impairing the tertiary structure and stability of γC-crystallin (CRYGC) leads to cataract formation in humans and zebrafish lens. Hum. Mutat. **2012**;33(2):391-401.

260. Zhang L, Fu S, Ou Y, Zhao T, Su Y, Liu P. A novel nonsense mutation in CRYGC is associated with autosomal dominant congenital nuclear cataracts and microcornea. Mol. Vis. **2009**;15:276-82.

261. Santhiya ST, Shyam Manohar M, Rawlley D, Vijayalakshmi P, Namperumalsamy P, Gopinath PM, *et al.* Novel mutations in the γ-crystallin genes cause autosomal dominant congenital cataracts. J. Med. Genet. **2002**;39(5):352-8.

262. Gonzalez-Huerta LM, Messina-Baas OM, Cuevas-Covarrubias SA. A family with autosomal dominant primary congenital cataract associated with a CRYGC mutation: evidence of clinical heterogeneity. Mol. Vis. **2007**;13(10):1333-8.

263. Stephan DA, Gillanders E, Vanderveen D, Freas-Lutz D, Wistow G, Baxevanis AD, *et al.* Progressive juvenile-onset punctate cataracts caused by mutation of the γD-crystallin gene. Proc. Natl. Acad. Sci. U.S.A. **1999**;96(3):1008-12.

264. Gu F, Li R, Ma XX, Shi LS, Huang SZ, Ma X. A missense mutation in the γD-crystallin gene CRYGD associated with autosomal dominant congenital cataract in a Chinese family. Mol. Vis. **2006**;12(1):26-31.

265. Zhang L-Y, Gong B, Tong J-P, Fan DS-P, Chiang SW-Y, Lou D, *et al.* A novel γD-crystallin mutation causes mild changes in protein properties but leads to congenital coralliform cataract. Mol. Vis. **2009**;15:1521-9.

266. Hilal L, Nandrot E, Belmekki M, Chefchaouni M, Bacha SE, Benazzouz B, *et al.* Evidence of clinical and genetic heterogeneity in autosomal dominant congenital cerulean cataracts. Ophthalmic Genet. **2002**;23(4):199-208.

267. Nandrot E, Slingsby C, Basak A, Cherif-Chefchaouni M, Benazzouz B, Hajaji Y, *et al.* Gamma-D crystallin gene (CRYGD) mutation causes autosomal dominant congenital cerulean cataracts. J. Med. Genet. **2003**;40(4):262-7.

268. Mackay DS, Andley UP, Shiels A. A missense mutation in the γD crystallin gene (CRYGD) associated with autosomal dominant "coral-like" cataract linked to chromosome 2q. Mol. Vis. **2004**;10:155-62.

269. Shentu X, Yao K, Xu W, Zheng S, Hu S, Gong X. Special fasciculiform cataract caused by a mutation in the γD-crystallin gene. Mol. Vis. **2004**;10(4):233-9.

270. Khan AO, Aldahmesh MA, Ghadhfan FE, Al-Mesfer S, Alkuraya FS. Founder heterozygous P23T CRYGD mutation associated with cerulean (and coralliform) cataract in 2 Saudi families. Mol. Vis. **2009**;15:1407-11.

271. Yang G, Xiong C, Li S, Wang Y, Zhao J. A recurrent mutation in CRYGD is associated with autosomal dominant congenital coralliform cataract in two unrelated Chinese families. Mol. Vis. **2011**;17:1085-9.

272. Plotnikova OV, Kondrashov FA, Vlasov PK, Grigorenko AP, Ginter EK, Rogaev EI. Conversion and compensatory evolution of the γ-crystallin genes and identification of a cataractogenic mutation that reverses the sequence of the human CRYGD gene to an ancestral state. Am. J. Hum. Genet. **2007**;81(1):32-43.

273. Sun W, Xiao X, Li S, Guo X, Zhang Q. Mutation analysis of 12 genes in Chinese families with congenital cataracts. Mol. Vis. **2011**;17:2197-206.

274. Wang L, Chen X, Lu Y, Wu J, Yang B, Sun X. A novel mutation in γD-crystallin associated with autosomal dominant congenital cataract in a Chinese family. Mol. Vis. **2011**;17:804-9.

275. Kmoch S, Brynda J, Asfaw B, Bezouška K, Novák P, Řezáčová P, *et al.* Link between a novel human γD-crystallin allele and a unique cataract phenotype explained by protein crystallography. Hum. Mol. Genet. **2000**;9(12):1779-86.

276. Wang B, Yu C, Xi Y-B, Cai H-C, Wang J, Zhou S, *et al.* A novel CRYGD mutation (p.Trp43Arg) causing autosomal dominant congenital cataract in a Chinese family. Hum. Mutat. **2011**;32(1):1939-47.

277. Santhiya ST, Kumar GS, Sudhakar P, Gupta N, Klopp N, Illig T, *et al.* Molecular analysis of cataract families in India: new mutations in the CRYBB2 and GJA3 genes and rare polymorphisms. Mol. Vis. **2010**;16:1837-47.

278. Li F, Wang S, Gao C, Liu S, Zhao B, Zhang M, *et al.* Mutation G61C in the CRYGD gene causing autosomal dominant congenital coralliform cataracts. Mol. Vis. **2008**;14:378-86.

279. Roshan M, Vijaya PH, Lavanya GR, Shama PK, Santhiya S, Graw J, *et al.* A novel human CRYGD mutation in a juvenile autosomal dominant cataract. Mol. Vis. **2010**;16:887-96.

280. Messina-Baas OM, Gonzalez-Huerta LM, Cuevas-Covarrubias SA. Two affected siblings with nuclear cataract associated with a novel missense mutation in the CRYGD gene. Mol. Vis. **2006**;12:995-1000.

281. Zhang L-Y, Yam G, Fan D, Tam P, Lam D, Pang C-P. A novel deletion variant of γD-crystallin responsible for congenital nuclear cataract. Mol. Vis. **2007**;13:2096-104.

282. Zhai Y, Li J, Yu W, Zhu S, Yu Y, Wu M, *et al.* Targeted exome sequencing of congenital cataracts related genes: broadening the mutation spectrum and genotype–phenotype correlations in 27 Chinese Han families. Sci. Rep. **2017**;7(1):1219-27.

283. Sun H, Ma Z, Li Y, Liu B, Li Z, Ding X, *et al.* Gamma-S crystallin gene (CRYGS) mutation causes dominant progressive cortical cataract in humans. J. Med. Genet. **2005**;42(9):706-10.

284. Vanita V, Singh JR, Singh D, Varon R, Sperling K. Novel mutation in the γ-S crystallin gene causing autosomal dominant cataract. Mol. Vis. **2009**;15:476-81.

285. Yang Z, Li Q, Zhu S, Ma X. A G57W mutation of CRYGS associated with autosomal dominant pulverulent cataracts in a Chinese family. Ophthalmic Genet. **2015**;36(3):281-3.

286. Zhang T, Yan L, Leng Y, Chen C, Ma L, Wang Q, *et al.* A novel missense mutation of CRYGS underlies congenital cataract in a Chinese family. Gene. **2018**;675:9-14.

287. Hanson SRA, Hasan A, Smith DL, Smith JB. The major *in vivo* modifications of the human water-insoluble lens crystallins are disulfide bonds, deamidation, methionine oxidation and backbone cleavage. Exp. Eye Res. **2000**;71(2):195-207.

288. Wilmarth PA, Tanner S, Dasari S, Nagalla SR, Riviere MA, Bafna V, *et al.* Age-related changes in human crystallins determined from comparative analysis of post-translational modifications in young and aged lens: does deamidation contribute to crystallin insolubility? J. Proteome Res. **2006**;5(10):2554-66.

289. Lyon YA, Collier MP, Riggs DL, Degiacomi MT, Benesch JLP, Julian RR. Structural and functional consequences of age-related isomerization in α-crystallins. J. Biol. Chem. **2019**;294(19):7546-55.

290. Fujii N, Kawaguchi T, Sasaki H, Fujii N. Simultaneous stereoinversion and isomerization at the Asp-4 residue in βB2-crystallin from the aged human eye lenses. Biochemistry. **2011**;50(40):8628-35.

291. Fujii N, Sakaue H, Sasaki H, Fujii N. A rapid, comprehensive liquid chromatography-mass spectrometry (LC-MS)-based survey of the Asp isomers in crystallins from human cataract lenses. J. Biol. Chem. **2012**;287(47):39992-40002.

292. Miesbauer LR, Zhou X, Yang Z, Yang Z, Sun Y, Smith DL, *et al.* Post-translational modifications of water-soluble human lens crystallins from young adults. J. Biol. Chem. **1994**;269(17):12494-502.

293. MacCoss MJ, McDonald WH, Saraf A, Sadygov R, Clark JM, Tasto JJ, *et al.* Shotgun identification of protein modifications from protein complexes and lens tissue. Proc. Natl. Acad. Sci. U.S.A. **2002**;99(12):7900-5.

294. Zhang Z, Smith DL, Smith JB. Human β-crystallins modified by backbone cleavage, deamidation and oxidation are prone to associate. Exp. Eye Res. **2003**;77(3):259-72.

295. Srivastava OP, McEntire JE, Srivastava K. Identification of a 9 kDa γ-crystallin fragment in human lenses. Exp. Eye Res. **1992**;54(6):893-901.

296. Searle BC, Dasari S, Wilmarth PA, Turner M, Reddy AP, David LL, *et al.* Identification of protein modifications using MS/MS de novo sequencing and the OpenSea alignment algorithm. J. Proteome Res. **2005**;4(2):546-54.

297. Hains PG, Truscott RJW. Proteomic analysis of the oxidation of cysteine residues in human age-related nuclear cataract lenses. Biochim. Biophys. Acta. **2008**;1784(12):1959-64.

298. Hains PG, Truscott RJW. Post-translational modifications in the nuclear region of young, aged, and cataract human lenses. J. Proteome Res. **2007**;6(10):3935-43.

299. Sharma KK, Santhoshkumar P. Lens aging: effects of crystallins. Biochim. Biophys. Acta, Gen. Subj. **2009**;1790(10):1095-108.

300. Lampi KJ, Amyx KK, Ahmann P, Steel EA. Deamidation in human lens βB2-crystallin destabilizes the dimer. Biochemistry. **2006**;45(10):3146-53.

301. Forsythe HM, Vetter CJ, Jara KA, Reardon PN, David LL, Barbar EJ, *et al.* Altered protein dynamics and increased aggregation of human γS-crystallin due to cataract-associated deamidations. Biochemistry. **2019**;58(40):4112-24.

302. Gupta R, Srivastava OP. Effect of deamidation of asparagine 146 on functional and structural properties of human lens αB-crystallin. Invest. Ophthalmol. Vis. Sci. **2004**;45(1):206-14.

303. Pande A, Mokhor N, Pande J. Deamidation of human γS-crystallin increases attractive protein interactions: implications for cataract. Biochemistry. **2015**;54(31):4890-9.

304. Ray NJ, Hall D, Carver JA. Deamidation of N76 in human γS-crystallin promotes dimer formation. Biochim. Biophys. Acta, Gen. Subj. **2016**;1860(1, Part B):315-24.

305. Ray NJ, Hall D, Carver JA. A structural and functional study of Gln147 deamidation in αA-crystallin, a site of modification in human cataract. Exp. Eye Res. **2017**;161:163-73.

306. Masters PM, Bada JL, Samuel Zigler J. Aspartic acid racemisation in the human lens during ageing and in cataract formation. Nature. **1977**;268(5615):71-3.

307. Hooi MYS, Truscott RJW. Racemisation and human cataract. D-Ser, D-Asp/Asn and D-Thr are higher in the lifelong proteins of cataract lenses than in age-matched normal lenses. AGE. **2011**;33(2):131-41.

308. Fujii N, Satoh K, Harada K, Ishibashi Y. Simultaneous stereoinversion and isomerization at specific aspartic acid residues in αA-crystallin from human lens. J. Biochem. **1994**;116(3):663-9.

309. Lyon YA, Sabbah GM, Julian RR. Differences in α-crystallin isomerization reveal the activity of protein isoaspartyl methyltransferase (PIMT) in the nucleus and cortex of human lenses. Exp. Eye Res. **2018**;171:131-41.

310. Takata T, Matsubara T, Nakamura-Hirota T, Fujii N. Negative charge at aspartate 151 is important for human lens αA-crystallin stability and chaperone function. Exp. Eye Res. **2019**;182:10-8.

311. Rogerson DT, Sachdeva A, Wang K, Haq T, Kazlauskaite A, Hancock SM, *et al.* Efficient genetic encoding of phosphoserine and its nonhydrolyzable analog. Nat. Chem. Biol. **2015**;11(7):496-503.

312. Carver JA, Nicholls KA, Aquilina AJ, Truscott RJW. Age-related changes in bovine α-crystallin and high-molecular-weight protein. Exp. Eye Res. **1996**;63(6):639-47.

313. Ecroyd H, Meehan S, Horwitz J, Aquilina JA, Benesch Justin LP, Robinson Carol V, *et al.* Mimicking phosphorylation of αB-crystallin affects its chaperone activity. Biochem. J. **2006**;401(1):129-41.

314. Ahmad MF, Raman B, Ramakrishna T, Rao CM. Effect of phosphorylation on αB-crystallin: differences in stability, subunit exchange and chaperone activity of homo and mixed oligomers of αB-crystallin and its phosphorylation-mimicking mutant. J. Mol. Biol. **2008**;375(4):1040-51.

315. Peschek J, Braun N, Rohrberg J, Back KC, Kriehuber T, Kastenmüller A, *et al.* Regulated structural transitions unleash the chaperone activity of αB-crystallin. Proc. Natl. Acad. Sci. U.S.A. **2013**;110(40):3780-9.

316. Lampi KJ, Wilmarth PA, Murray MR, David LL. Lens β-crystallins: the role of deamidation and related modifications in aging and cataract. Prog. Biophys. Mol. Biol. **2014**;115(1):21-31.

317. Aquilina JA, Benesch JLP, Ding LL, Yaron O, Horwitz J, Robinson CV. Subunit exchange of polydisperse proteins: mass spectrometry reveals consequences of αA-crystallin truncation. J. Biol. Chem. **2005**;280(15):14485-91.

318. Aziz A, Santhoshkumar P, Sharma KK, Abraham EC. Cleavage of the C-terminal serine of human αA-crystallin produces Aa 1-172 with increased chaperone activity and oligomeric size. Biochemistry. **2007**;46(9):2510-9.

319. Chaves JM, Srivastava K, Gupta R, Srivastava OP. Structural and functional roles of deamidation and/or truncation of N- or C-termini in human αA-crystallin. Biochemistry. **2008**;47(38):10069-83.

320. Bateman OA, Lubsen NH, Slingsby C. Association behaviour of human βB1-crystallin and its truncated forms. Exp. Eye Res. **2001**;73(3):321-31.

321. Srivastava K, Gupta R, Chaves JM, Srivastava OP. Truncated human βB1-crystallin shows altered structural properties and interaction with human βA3-crystallin. Biochemistry. **2009**;48(30):7179-89.

322. Truscott RJW, Augusteyn RC. The state of sulphydryl groups in normal and cataractous human lenses. Exp. Eye Res. **1977**;25(2):139-48.

323. Truscott RJW, Augusteyn RC. Changes in human lens proteins during nuclear cataract formation. Exp. Eye Res. **1977**;24(2):159-70.

324. Truscott RJW. Age-related nuclear cataract - oxidation is the key. Exp. Eye Res. **2005**;80(5):709-25.

325. Fan X, Zhou S, Wang B, Hom G, Guo M, Li B, *et al.* Evidence of highly conserved β-crystallin disulfidome that can be mimicked by *in vitro* oxidation in age-related human cataract and glutathione depleted mouse lens. Mol. Cell. Proteom. **2015**;14(12):3211-23.

326. Garner MH, Spector A. Selective oxidation of cysteine and methionine in normal and senile cataractous lenses. Proc. Natl. Acad. Sci. U.S.A. **1980**;77(3):1274-7.

327. Harding JJ. Free and protein-bound glutathione in normal and cataractous human lenses. Biochem. J. **1970**;117(5):957-60.

328. Kamei A. Glutathione levels of the human crystalline lens in aging and its antioxidant effect against the oxidation of lens proteins. Biol. Pharm. Bull. **1993**;16(9):870-5.

329. Bova LM, Sweeney MHJ, Jamie JF, Truscott RJW. Major changes in human ocular UV protection with age. Invest. Ophthalmol. Vis. Sci. **2001**;42(1):200-5.

330. Ramkumar S, Fan X, Wang B, Yang S, Monnier VM. Reactive cysteine residues in the oxidative dimerization and Cu2+ induced aggregation of human γD-crystallin: implications for age-related cataract. Biochim. Biophys. Acta, Mol. Basis Dis. **2018**;1864(11):3595-604.

331. Xia Z, Yang Z, Huynh T, King JA, Zhou R. UV-radiation induced disruption of dry-cavities in human γD-crystallin results in decreased stability and faster unfolding. Sci. Rep. **2013**;3(1):1560-8.

332. Serebryany E, King JA. Wild-type human γD-crystallin promotes aggregation of its oxidation-mimicking, misfolding-prone W42Q mutant. J. Biol. Chem. **2015**;290(18):11491-503.

333. Serebryany E, Woodard JC, Adkar BV, Shabab M, King JA, Shakhnovich EI. An internal disulfide locks a misfolded aggregation-prone intermediate in cataract-linked mutants of human γD-crystallin. J. Biol. Chem. **2016**;291(36):19172-83.

334. Blundell T, Lindley P, Miller L, Moss D, Slingsby C, Tickle I, *et al.* The molecular structure and stability of the eye lens: X-ray analysis of γ-crystallin II. Nature. **1981**;289(5800):771-7.

335. Ebersbach H, Fiedler E, Scheuermann T, Fiedler M, Stubbs MT, Reimann C, *et al.* Affilin–novel binding molecules based on human γ-B-crystallin, an all β-sheet protein. J. Mol. Biol. **2007**;372(1):172-85.

336. Basak AK, Kroone RC, Lubsen NH, Naylor CE, Jaenicke R, Slingsby C. The C-terminal domains of γS-crystallin pair about a distorted twofold axis. Protein Eng., Des. Sel. **1998**;11(5):337-44.

337. Sagar V, Chaturvedi SK, Schuck P, Wistow G. Crystal structure of chicken γS-crystallin reveals lattice contacts with implications for function in the lens and the evolution of the βγ-crystallins. Structure. **2017**;25(7):1068-78

338. Hochberg GKA, Benesch JLP. Dynamical structure of αB-crystallin. Prog. Biophys. Mol. Biol. **2014**;115(1):11-20.

339. Bagnéris C, Bateman OA, Naylor CE, Cronin N, Boelens WC, Keep NH, *et al.* Crystal structures of α-crystallin domain dimers of αB-crystallin and Hsp20. J. Mol. Biol. **2009**;392(5):1242-52.

340. Laganowsky A, Eisenberg D. Non-3D domain swapped crystal structure of truncated zebrafish alphaA crystallin. Protein Sci. **2010**;19(10):1978-84.

341. Clark AR, Naylor CE, Bagnéris C, Keep NH, Slingsby C. Crystal structure of R120G disease mutant of human αB-crystallin domain dimer shows closure of a groove. J. Mol. Biol. **2011**;408(1):118-34.

342. Sluchanko NN, Beelen S, Kulikova AA, Weeks SD, Antson AA, Gusev NB, *et al.* Structural basis for the interaction of a human small heat shock protein with the 14-3-3 universal signaling regulator. Structure. **2017**;25(2):305-16.

343. Driessen HPC, Herbrink P, Bloemendal H, de Jong WW. The β-crystallin bp chain is internally duplicated and homologous with γ-crystallin. Exp. Eye Res. **1980**;31(2):243-6.

344. Wong EK, Prytkova V, Freites JA, Butts CT, Tobias DJ. Molecular mechanism of aggregation of the cataract-related γD-crystallin W42R variant from multiscale atomistic simulations. Biochemistry. **2019**;58(35):3691-9.

345. Carver JA, Lindner RA. NMR spectroscopy of α-crystallin. Insights into the structure, interactions and chaperone action of small heat-shock proteins. Int. J. Biol. Macromol. **1998**;22(3):197-209.

346. Lindner RA, Carver JA, Ehrnsperger M, Buchner J, Esposito G, Behlke J, *et al.* Mouse Hsp25, a small heat shock protein. Eur. J. Biochem. **2000**;267(7):1923-32.

347. Alderson TR, Benesch JLP, Baldwin AJ. Proline isomerization in the C-terminal region of Hsp27. Cell Stress Chaperon. **2017**;22(4):639-51.

348. Jehle S, van Rossum B, Stout JR, Noguchi SM, Falber K, Rehbein K, *et al.* αB-crystallin: a hybrid solid-state/solution-state NMR investigation reveals structural aspects of the heterogeneous oligomer. J. Mol. Biol. **2009**;385(5):1481-97.

349. Jehle S, Rajagopal P, Bardiaux B, Markovic S, Kühne R, Stout JR, *et al.* Solid-state NMR and SAXS studies provide a structural basis for the activation of αB-crystallin oligomers. Nat. Struct. Mol. Biol. **2010**;17(9):1037-42.

350. Baldwin AJ, Hilton GR, Lioe H, Bagnéris C, Benesch JLP, Kay LE. Quaternary dynamics of αB-crystallin as a direct consequence of localised tertiary fluctuations in the C-terminus. J. Mol. Biol. **2011**;413(2):310-20.

351. Baldwin AJ, Walsh P, Hansen DF, Hilton GR, Benesch JLP, Sharpe S, *et al.* Probing dynamic conformations of the high-molecular-weight αB-crystallin heat shock protein ensemble by NMR spectroscopy. J. Am. Chem. Soc. **2012**;134(37):15343-50.

352. Rajagopal P, Tse E, Borst AJ, Delbecq SP, Shi L, Southworth DR, *et al.* A conserved histidine modulates HSPB5 structure to trigger chaperone activity in response to stress-related acidosis. eLife. **2015**;4:e07304.

353. Mainz A, Bardiaux B, Kuppler F, Multhaup G, Felli IC, Pierattelli R, *et al.* Structural and mechanistic implications of metal binding in the small heat-shock protein αB-crystallin. J. Biol. Chem. **2012**;287(2):1128-38.

354. Jobby MK, Sharma Y. Calcium-binding to lens βB2- and βA3-crystallins suggests that all β-crystallins are calcium-binding proteins. FEBS J. **2007**;274(16):4135-47.

355. Wu Z, Delaglio F, Wyatt K, Wistow G, Bax A. Solution structure of γS-crystallin by molecular fragment replacement NMR. Protein Sci. **2005**;14(12):3101-14.

356. Dixit K, Pande A, Pande J, Sarma SP. Nuclear magnetic resonance structure of a major lens protein, human γC-crystallin: role of the dipole moment in protein solubility. Biochemistry. **2016**;55(22):3136-49.

357. Bari KJ, Sharma S, Chary KVR. Structure of G57W mutant of human γS-crystallin and its involvement in cataract formation. J. Struct. Biol. **2019**;205(3):72-8.

358. Lee S, Mahler B, Toward J, Jones B, Wyatt K, Dong L, *et al.* A single destabilizing mutation (F9S) promotes concerted unfolding of an entire globular domain in γS-crystallin. J. Mol. Biol. **2010**;399(2):320-30.

359. Mahler B, Doddapaneni K, Kleckner I, Yuan C, Wistow G, Wu Z. Characterization of a transient unfolding intermediate in a core mutant of γS-crystallin. J. Mol. Biol. **2011**;405(3):840-50.

360. Haley DA, Horwitz J, Stewart PL. The small heat-shock protein, αB-crystallin, has a variable quaternary structure. J. Mol. Biol. **1998**;277(1):27-35.

361. Carver JA, Aquilina JA, Truscott RJW. A possible chaperone-like quaternary structure for α-crystallin. Exp. Eye Res. **1994**;59(2):231-4.

362. Peschek J, Braun N, Franzmann TM, Georgalis Y, Haslbeck M, Weinkauf S, *et al.* The eye lens chaperone α-crystallin forms defined globular assemblies. Proc. Natl. Acad. Sci. U.S.A. **2009**;106(32):13272-7.

363. Haslbeck M, Peschek J, Buchner J, Weinkauf S. Structure and function of α-crystallins: traversing from *in vitro* to *in vivo*. Biochim. Biophys. Acta, Gen. Subj. **2016**;1860(1, Part B):149-66.

364. Braun N, Zacharias M, Peschek J, Kastenmüller A, Zou J, Hanzlik M, *et al.* Multiple molecular architectures of the eye lens chaperone αB-crystallin elucidated by a triple hybrid approach. Proc. Natl. Acad. Sci. U.S.A. **2011**;108(51):20491-6.

365. Delbecq SP, Klevit RE. One size does not fit all: the oligomeric states of αB crystallin. FEBS Lett. **2013**;587(8):1073-80.

366. Park Z-Y, Sadygov R, Clark JM, Clark JI, Yates JR. Assigning *in vivo* carbamylation and acetylation in human lens proteins using tandem mass spectrometry and database searching. Int. J. Mass Spectrom. **2007**;259(1):161-73.

367. Lapko VN, Smith DL, Smith JB. Methylation and carbamylation of human γ-crystallins. Protein Sci. **2003**;12(8):1762-74.

368. Takata T, Smith JP, Arbogast B, David LL, Lampi KJ. Solvent accessibility of βB2-crystallin and local structural changes due to deamidation at the dimer interface. Exp. Eye Res. **2010**;91(3):336-46.

369. Lampi KJ, Fox CB, David LL. Changes in solvent accessibility of wild-type and deamidated βB2-crystallin following complex formation with αA-crystallin. Exp. Eye Res. **2012**;104:48-58.

370. Akashi S, Maleknia SD, Saikusa K, Downard KM. Stability of the βB2B3 crystallin heterodimer to increased oxidation by radical probe and ion mobility mass spectrometry. J. Struct. Biol. **2015**;189(1):20-7.

371. Lampi KJ, Murray MR, Peterson MP, Eng BS, Yue E, Clark AR, *et al.* Differences in solution dynamics between lens β-crystallin homodimers and heterodimers probed by hydrogen–deuterium exchange and deamidation. Biochim. Biophys. Acta, Gen. Subj. **2016**;1860(1, Part B):304-14.

372. Baldwin AJ, Lioe H, Robinson CV, Kay LE, Benesch JLP. αB-crystallin polydispersity is a consequence of unbiased quaternary dynamics. J. Mol. Biol. **2011**;413(2):297-309.

373. Benesch JLP, Ayoub M, Robinson CV, Aquilina JA. Small heat shock protein activity is regulated by variable oligomeric substructure. J. Biol. Chem. **2008**;283(42):28513-7.

374. Aquilina JA, Benesch JLP, Ding LL, Yaron O, Horwitz J, Robinson CV. Phosphorylation of αB-crystallin alters chaperone function through loss of dimeric substructure. J. Biol. Chem. **2004**;279(27):28675-80.

375. Benesch JLP, Aquilina JA, Ruotolo BT, Sobott F, Robinson CV. Tandem mass spectrometry reveals the quaternary organization of macromolecular assemblies. Chem. Biol. **2006**;13(6):597-605.

376. Baldwin Andrew J, Lioe H, Hilton Gillian R, Baker Lindsay A, Rubinstein John L, Kay Lewis E, *et al.* The polydispersity of αB-crystallin is rationalized by an interconverting polyhedral architecture. Structure. **2011**;19(12):1855-63.

377. Chandler SA, Benesch JLP. Mass spectrometry beyond the native state. Curr. Opin. Chem. Biol. **2018**;42:130-7.

378. Wright MA, Ruggeri FS, Saar KL, Challa PK, Benesch JL, Knowles TP. Analysis of αB-crystallin polydispersity in solution through native microfluidic electrophoresis. Analyst. **2019**;144:4413-24.

379. Sergeev YV, Hejtmancik JF, Wingfield PT. Energetics of domain−domain interactions and entropy driven association of β-crystallins. Biochemistry. **2004**;43(2):415-24.

380. de Jong S, Krylov SN. Pressure-based approach for the analysis of protein adsorption in capillary electrophoresis. Anal. Chem. **2012**;84(1):453-8.

381. Grossman PD, Colburn JC. Capillary electrophoresis: theory and practice. Elsevier Science; **1992**.

382. Ewing AG, Wallingford RA, Olefirowicz TM. Capillary electrophoresis. Anal. Chem. **1989**;61(4):292-303.

383. Popovici S-T, Kok WT, Schoenmakers PJ. Band broadening in size-exclusion chromatography of polydisperse samples. J. Chromatogr. **2004**;1060(1):237-52.

384. Karim MR, Shinagawa S, Takagi T. Electrophoretic mobilities of the complexes between sodium dodecyl sulfate and various peptides or proteins determined by free solution electrophoresis using coated capillaries. Electrophoresis. **1994**;15(1):1141-6.

385. Gudiksen KL, Gitlin I, Whitesides GM. Differentiation of proteins based on characteristic patterns of association and denaturation in solutions of SDS. Proc. Natl. Acad. Sci. U.S.A. **2006**;103(21):7968-72.

386. Thevarajah JJ, Sutton AT, Maniego AR, Whitty EG, Harrisson S, Cottet H, *et al.* Quantifying the heterogeneity of chemical structures in complex charged polymers through the dispersity of their distributions of electrophoretic mobilities or of compositions. Anal. Chem. **2016**;88(3):1674-81.

387. Stellwagen NC, Gelfi C, Righetti PG. The free solution mobility of DNA. Biopolymers. **1997**;42(6):687-703.

388. Cottet H, Gareil P, Theodoly O, Williams CE. A semi-empirical approach to the modeling of the electrophoretic mobility in free solution: application to polystyrenesulfonates of various sulfonation rates. Electrophoresis. **2000**;21(17):3529-40.

389. Beneito-Cambra M, Gareil P, Badet B, Badet-Denisot M-A, Delaunay N. First investigations for the characterization of glucosamine-6-phosphate synthase by capillary electrophoresis. J. Chromatogr. B. **2018**;1072:130-5.

390. Clouthier CM, Mironov GG, Okhonin V, Berezovski MV, Keillor JW. Real-time monitoring of protein conformational dynamics in solution using kinetic capillary electrophoresis. Angew. Chem. Int. Ed. **2012**;51(50):12464-8.

391. Shi Y, Rhodes NR, Abdolvahabi A, Kohn T, Cook NP, Marti AA, *et al.* Deamidation of asparagine to aspartate destabilizes Cu, Zn superoxide dismutase, accelerates fibrillization, and mirrors ALS-linked mutations. J. Am. Chem. Soc. **2013**;135(42):15897-908.

392. Chamieh J, Martin M, Cottet H. Quantitative analysis in capillary electrophoresis: transformation of raw electropherograms into continuous distributions. Anal. Chem. **2015**;87(2):1050-7.

393. Hilser VJ, Freire E. Quantitative analysis of conformational equilibrium using capillary electrophoresis: applications to protein folding. Anal. Biochem. **1995**;224(2):465-85.

394. Rochu D, Masson P. Multiple advantages of capillary zone electrophoresis for exploring protein conformational stability. Electrophoresis. **2002**;23(2):189-202.

CHAPTER 2:

DISULFIDE-LINKED γS-CRYSTALLIN DIMER

"No problem is too small or too trivial if we can really do something about it."

– Richard Feynman

DECLARATION

The following article is peer-reviewed and published in the *Journal of Molecular Biology*.

*Thorn DC, *Grosas AB, Mabbitt PD, Ray NJ, Jackson CJ, Carver JA. The Structure and Stability of the Disulfide-Linked γS-Crystallin Dimer Provide Insight into Oxidation Products Associated with Lens Cataract Formation. Journal of Molecular Biology. 2019;431(3):483-97. doi: https://doi.org/10.1016/j.jmb.2018.12.005

*These authors contributed equally.

All experimental work was carried out by the author, except where otherwise stated below, under the supervision of Professor John Carver.

The contributions of other authors are as follows: The author, Dr David Thorn, and Professor John Carver conceived the study. Dr David Thorn and Dr Nicholas Ray purified the protein. Dr David Thorn and Dr Peter Mabbitt crystallised the protein and Dr Peter Mabbitt and Professor Colin Jackson solved the X-ray crystal structure. Dr David Thorn assisted in the acquisition and analysis of small angle X-ray scattering data and *de novo* envelope reconstruction, size exclusion chromatography, transmission electron microscopy, dynamic light scattering and spectroscopy. The author wrote the publication with input from the co-authors.

The Structure and Stability of the Disulfide-Linked γS-Crystallin Dimer Provide Insight into Oxidation Products Associated with Lens Cataract Formation

David C. Thorn†, Aidan B. Grosas†, Peter D. Mabbitt, Nicholas J. Ray, Colin J. Jackson and John A. Carver

Research School of Chemistry, *The Australian National University, Acton, ACT 2601, Australia*

Correspondence to John A. Carver: *john.carver@anu.edu.au*
https://doi.org/10.1016/j.jmb.2018.12.005
Edited by Sheena Radford

Abstract

The reducing environment in the eye lens diminishes with age, leading to significant oxidative stress. Oxidation of lens crystallin proteins is the major contributor to their destabilization and deleterious aggregation that scatters visible light, obscures vision, and ultimately leads to cataract. However, the molecular basis for oxidation-induced aggregation is unknown. Using X-ray crystallography and small-angle X-ray scattering, we describe the structure of a disulfide-linked dimer of human γS-crystallin that was obtained via oxidation of C24. The γS-crystallin dimer is stable at glutathione concentrations comparable to those in aged and cataractous lenses. Moreover, dimerization of γS-crystallin significantly increases the protein's propensity to form large insoluble aggregates owing to non-cooperative domain unfolding, as is observed in crystallin variants associated with early-onset cataract. These findings provide insight into how oxidative modification of crystallins contributes to cataract and imply that early-onset and age-related forms of the disease share comparable development pathways.

© 2018 Elsevier Ltd. All rights reserved.

Introduction

The eye lens contains a high concentration of crystallin proteins arranged in a well-ordered, short-range array that allows for lens transparency and the refraction of light onto the retina, thus ensuring proper vision [1–3]. In mammals, the crystallins comprise three types (α, β, and γ), of which there are several isoforms. The α-crystallins are members of the small heat-shock protein family, whereas the β- and γ-crystallins are a structurally homologous superfamily of proteins that are not related to small heat-shock proteins [4–6]. β- and γ-crystallins have a monomeric mass of approximately 20 kDa and consist of two domains, each containing two Greek-key β-sheet motifs, which are adjoined by a short linking peptide [6–9].

There is little protein turnover in the eye lens, and thus, the crystallins are long-lived proteins that must maintain their structural integrity throughout life to preserve lens transparency [10–12]. Cataract occurs due to a loss of crystallin protein stability and the subsequent propensity of crystallins to partially unfold, leading to aggregation and precipitation [6,13]. While cataract acquired during early life commonly stems from destabilizing, inheritable mutations in crystallin proteins, age-related cataract is thought to originate from cumulative post-translational modifications (PTMs) [14,15]. Oxidation is a prevalent crystallin PTM in both aged and cataractous lenses [16,17] that increases the aggregation propensity of some crystallins *in vitro* [18–20]. Cysteine residues are the principal site of protein oxidation [16,21], and disulfide-linked crystallins are a major component of the insoluble fraction of cataractous lenses [22,23]. The key factor in preventing crystallin oxidation is the cellular reductant glutathione, the levels of which diminish with age, to the extent that it is severely depleted in cataractous lenses [24–26].

γS-crystallin (γS) is one of the major crystallins in the human lens [27], and its abundance increases with age due to postnatal expression [28]. Human γS in cataractous lenses is oxidized at specific cysteine residues [17], including *S*-methylation [29], *S*-glutathionylation [30], and intermolecular disulfide bond formation [23]. Indeed, γS forms disulfide-

© 2018 Elsevier Ltd. All rights reserved.

linked dimers *in vitro* [31]. Disulfide-linked dimerization similarly occurs for the R14C mutant of γD-crystallin, leading to increased aggregation propensity and hereditary juvenile-onset cataract [32]. In light of this and the enhanced oxidative conditions in the aging lens, a detailed understanding of the structural and physiological implications of disulfide-linked dimerization of wild-type γS is needed. Herein, we isolated disulfide-linked, dimeric human γS and determined its structure by X-ray crystallography and small-angle X-ray scattering (SAXS). The significance of the disulfide bonding arrangement of the three clustered cysteine residues at positions 22, 24, and 26 is discussed. Furthermore, we provide biophysical and biochemical evidence for the role of the γS dimer in age-related cataract and the potential molecular mechanisms underlying this role.

Results

Structure of human γS disulfide-linked dimer

Human γS monomer was expressed heterologously in *Escherichia coli* and purified using anion-exchange and size-exclusion chromatography (SEC). Previously, γS was reported to undergo time-dependent dimer formation under ambient, oxidative conditions at slightly elevated pH, for example, pH 8 [31]. In the present study, γS monomer was readily converted to dimer at physiological pH (i.e., pH 7) by concentrating the monomer to a minimum of 20 mg mL^{-1} and leaving the protein at 4 °C for 1 week. Monomeric and dimeric forms of γS were isolated and checked for homogeneity and correct mass using SEC coupled to multi-angle light scattering (Fig. S1a).

The γS dimer was crystallized and its structure determined by X-ray crystallography to a resolution of 2.1 Å (Table 1). The crystal structure shows the canonical two-domain Greek-key motif β-strand arrangement for each monomeric subunit (Fig. 1a, top). The asymmetric unit consists of a dimer of γS arranged in the same orientation as is observed for the crystal structure of truncated βB1-crystallin [34] and the solution structure of full-length βB2-crystallin [35], which has been previously described as a QR configuration [36]. There are three cysteine residues in a loop close to the N-terminus of γS: C22, C24, C26. In each monomer in the asymmetric unit, the loop is cyclized by one intramolecular disulfide bond between C22 and C26 in each subunit (Fig. 1a, bottom, c and d). No other disulfide bonds were present in the crystallographic dimer observed in the asymmetric unit. However, electrospray ionization mass spectrometry (MS) of the γS dimer indicated the presence of three disulfide bonds, that is, a loss of 6 Da from the expected mass of two monomers (Fig. S1b). Inspection of symmetry mates within the crystal lattice revealed an alternative dimer interface having a C24:C24' disulfide bond as its principal intermolecular contact. Analysis of the dimer interface using PISA [37] indicated that this alternative interface buries 648 $Å^2$ of protein surface area. In addition to the intermolecular disulfide bond, there are several non-covalent intermolecular interactions at the dimer interface, most notably two hydrogen bonds between the δ-oxygen of D23 and the amide hydrogen of A27 on either subunit, as well as an aromatic ring interaction between Y32 of each subunit.

Table 1. Data processing and refinement statistics for γS disulfide-linked dimer

Data collection	
Wavelength (Å)	0.9763
Space group	*P*1
Cell dimensions	
a, *b*, *c* (Å)	49.63, 52.37, 53.97
α, *β*, *γ* (°)	108.72, 111.46, 105.50
Total reflections	61,796 (5062)
Unique reflections	23,548 (2372)
Multiplicity	2.6 (2.6)
Completeness (%)	93.6 (94.4)
Mean *I*/sigma(*I*)	7.8 (1.3)
Wilson *B*-factor ($Å^2$)	33.3
R-merge	0.032 (0.559)
R-pim	0.022 (0.390)
CC1/2	0.997 (0.757)
Refinement	
R-work/*R*-free	0.2248/0.2716 (0.3067/0.3783)
Number of non-hydrogen atoms	3029
Macromolecules	2892
Water	137
Protein residues	348
RMS (bonds, Å)	0.008
RMS (angles, °)	0.95
Ramachandran preferred (%)	94.45
Ramachandran allowed (%)	4.36
Ramachandran outliers (%)	1.16
Clashscore	6
Average *B*-factor ($Å^2$)	32.55
Macromolecules	32.53
Solvent	32.88
PDB ID	6FD8

Highest-resolution shell is shown in parentheses.

Proteins can adopt quaternary arrangements in solution that are distinct from their crystal structure, for example, the structurally and functionally related βB2-crystallin dimer [35]. Thus, we used SAXS to characterize the structure of the monomer and disulfide-linked γS dimer in solution. The pairwise distance distribution function (*P*(r)) was examined to ensure shape and size differences between γS monomer and dimer were evident in the SAXS data. The *P*(r) of the monomer shows a unimodal distribution consistent with both domains being closely associated with a maximum diameter (D_{max}) of 59 Å and a radius of gyration (R_g) of 18.5 Å. The *P*(r) of the dimer shows a bimodal distribution consistent with

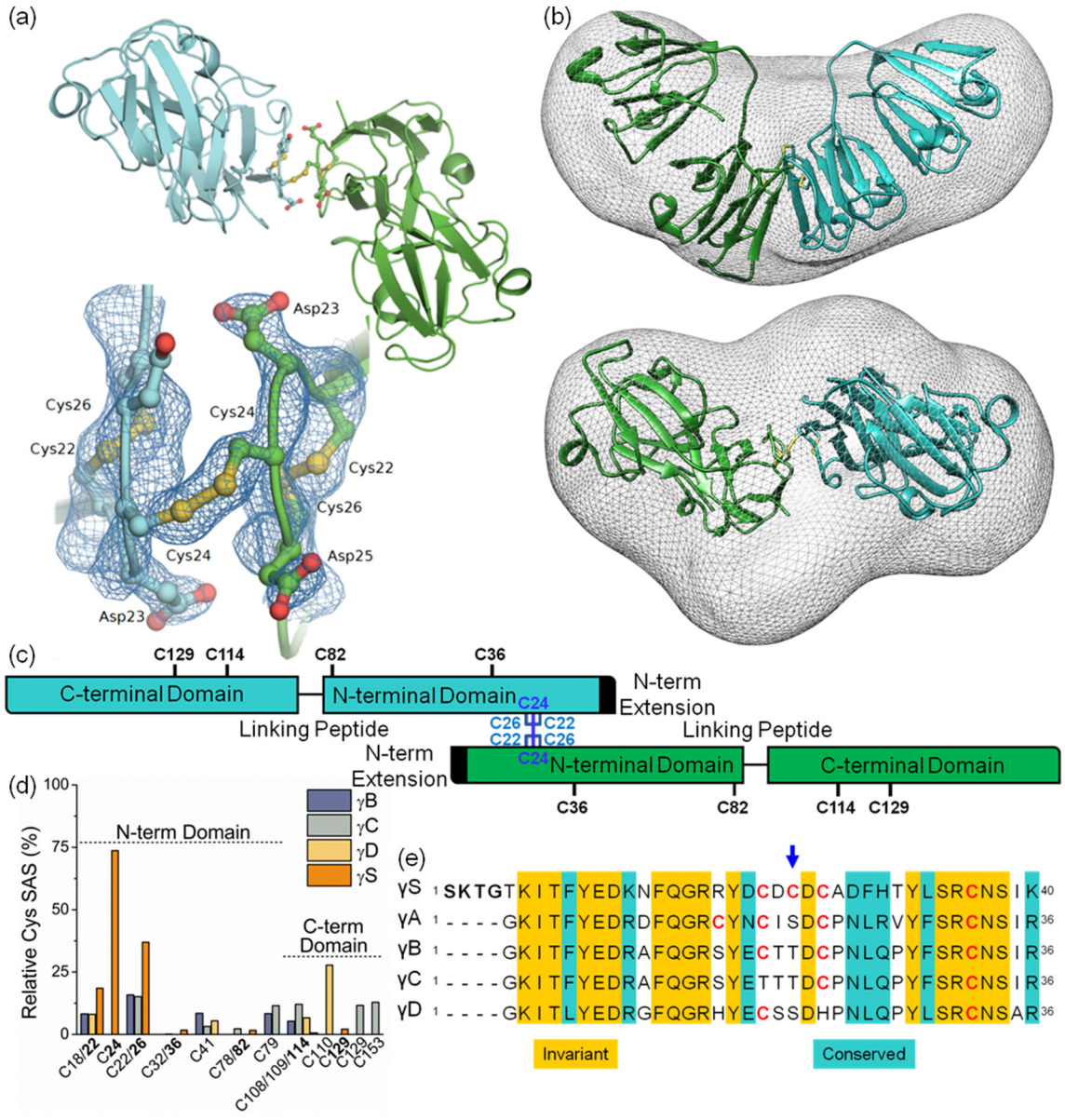


Fig. 1. Structure of the γS dimer. (a) Crystal structure of γS dimer with the inter- (C24) and intramolecular (C22 and C26) disulfide bond arrangement shown as a whole view of the dimer (top) and as a magnified view within the $2mF_o - dF_c$ map (blue lines, contoured at 1.5σ) centered on the intermolecular disulfide bond (bottom). (b) Crystal structure of γS dimer fitted into the *ab initio* SAXS shape envelope (grey line, contoured at 2σ) as shown from a side view (top) and from above (bottom). (c) Linearized schematic of the γS dimer structure with prominent structural features and residues indicated and colored concordantly with the crystal structure above. (d) Comparison of the percentage SAS of all cysteine residues in monomers of γB (PDB: 2JDF), γC (PDB: 2NBR), γD (PDB: 1HK0), and γS (PDB: 2M3T). The structure for γA is not available. SAS is relative to a Gly-Cys-Gly tripeptide in an extended conformation [33]. The positions of cysteine residues in the amino acid sequences of γS and other human γ-crystallins are in bold and normal typeface, respectively. C108 and C109 denote cysteine residues in γC/γD and γB, respectively. (e) Sequence alignment of human γS (residues 1–40) with four other human γ-crystallins highlighting: (i) invariant and chemically conserved amino acid residues as shown in the key; (ii) cysteine residues, C, in bold typeface and red; and (iii) C24 with a blue arrow.

the "V" shape of the disulfide-linked dimer, which exhibits greater distance between the C-terminal domains of each subunit than the alternative non-disulfide linked dimer, giving a D_{max} of 74 Å and a R_g of 24.3 Å (Fig. S1d). A direct comparison of the predicted one-dimensional scattering profiles from the structures of the monomer (PDB: 2M3T) and disulifde-linked dimer with that of the experimentally

obtained scattering data shows that the monomer structure fits well with a χ value of 1.19 (Fig. S1e, light blue). However, the dimer's fit is not as good with a χ value of 2.19 due to a discrepancy in the mid-q range (Fig. S1e, orange). To explore the basis of this structural discrepancy in the dimer, an *ab initio* shape envelope was built using the γS dimer SAXS data within which the crystal structure was fitted (Fig. 1b, top and bottom). The general domain arrangement along the longitudinal axis is consistent with that of the disulfide-linked dimer (Fig. 1b, top), thus verifying the conformation of the biological assembly. There are, however, two protuberances in the transverse plane of the SAXS envelope that correlate with the three-dimensional space that would be sampled by an unstructured and flexible N-terminal extension (Fig. 1b, bottom), encompassing the first four N-terminal residues (Fig. 1c, black, and e, black boldface text) [38]. A similar protuberance is noted in the *ab initio* SAXS envelope of βB2-crystallin dimer due to flexible terminal extensions [35,39]. Given the absence of the first three residues of the N-terminal extension in the crystal structure (due to insufficient electron density) and the high *B*-factor for the fourth residue (G4) (Fig. S1f), it is clear the N-terminal extension maintains flexibility in the dimer. Taken together, these results highlight the source of the discrepancy between the crystal structure and the solution SAXS data as being the N-terminal extension which protrudes from the domain core structure in the dimer (Fig. S1g), as it does in the monomer [40].

C24 is unique to γS and is the most solvent-exposed cysteine in human γ-crystallins

To explore whether this disulfide bonding arrangement observed in human γS is likely to be conserved in γ-crystallins, we compared the sequence of γS to that of human isoforms γA-crystallin (γA), γB-crystallin (γB), γC-crystallin (γC), and γD-crystallin (γD) (Fig. 1e). C22 and C26 (C18 and C22 in other γ-crystallins) are both conserved in γA and γB (Fig. 1e), implying that an intramolecular disulfide bond could form between C18 and C22 in these two γ-crystallins. However, C24 (Fig. 1e, blue arrow) is unique to γS with either serine or threonine substituting for cysteine at the analogous site (position 20) in other isoforms. To assess the relative probability of other human isoforms forming disulfide-linked dimers, we also examined the solvent accessible surface area (SAS) of all cysteines in the monomeric structures of γS, γB, γC, and γD relative to the side-chain SAS of Cys in a Gly-Cys-Gly tripeptide in an extended conformation [33] (Fig. 1d). C24 is significantly more solvent-exposed than any other cysteine in human γ-crystallins, followed by C26, although the latter is unlikely to be involved in the formation of an intermolecular disulfide bond owing to its participation in an intramolecular disulfide bond to C22 (Fig. 1a, bottom, c and d). With the exception of C110 in γD, all other cysteine residues in human γ-crystallins have a SAS of 20% or less (Fig. 1c and d), implying that significant formation of a native disulfide-linked dimer is confined to γS. An interspecies comparison of γS revealed that C24 (Fig. S1c, blue arrow) is highly conserved in eutherians (placental mammals), whereas it is not present in marsupials, birds, reptiles and fish, implying that disulfide-linked dimerization of a variety of eutherian orthologs is possible, as has been observed for bovine γS [31]. Moreover, the absence of C24 in other vertebrates suggests that it is a recent mutation on the evolutionary timescale, arising in a late common ancestor to eutherians (Fig. S1c).

The γS dimer persists in a reducing environment relevant to aging and cataractous lenses

The young human lens is a reducing environment due to the abundance of reduced glutathione (GSH) at a concentration of approximately 6 mM, which diminishes to approximately 2 mM upon aging and to less than 1 mM upon cataract formation [24–26]. We investigated the susceptibility of the γS dimer to dissociation over a range of physiologically relevant reducing environments using analytical SEC. Ratios of GSH and oxidized glutathione (GSSG) to a total concentration of 6 mM were used to modulate the reduction potential of the solution. Analytical SEC profiles showed significant changes over a 72-h period for those samples containing GSH (Fig. S2a). Interestingly, there was a minor linear shift, most prominent at 72 h, in the retention volume of the monomer peak maximum over the GSH:GSSG series (Fig. S2b), which most likely reflects an increase in hydrodynamic size of the monomer due to glutathionylation of the protein (see the following paragraph). From the peak integration values, the percentage of dimer was calculated for each time point and reduction potential, and then plotted against time (Fig. 2a). The results show that over 72 h, the dimer dissociates at an exponential rate that slows with decreasing reducing potential. For example, the dimer was reduced twice as fast by 3 mM GSH, and five times as fast by 6 mM GSH, when compared with 1 mM GSH (Fig. 2b). By 72 h, the percentage of the dimer remained relatively constant over time for all conditions; at 1, 3, and 6 mM GSH, there was 22%, 9%, and 3% dimer remaining, respectively (Fig. 2a). The fully oxidized environment (6 mM GSSG) showed no significant changes in the quantity of disulfide-linked dimer over the experimental time course. Thus, γS is monomeric at GSH:GSSG ratios associated with healthy young lenses, but the dimer becomes increasingly stable as GSH:GSSG

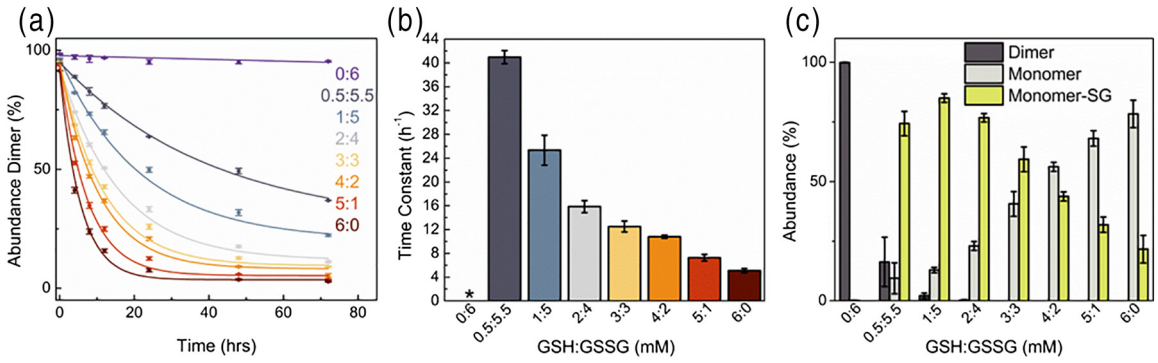


Fig. 2. The effect of glutathione on the abundance, reduction rate and reduction products of the γS dimer. (a) Percentage abundance of the γS dimer with different GSH:GSSG ratios over time as monitored by SEC. The data were fitted to a single exponential decay curve ($R^2 > 0.99$) except for 0:6 GSH:GSSG, which was fitted to a straight line function. (b) Time constant from the single exponential decay curve fitted against the relevant GSH:GSSG ratio. The errors are the standard deviation of two independent repeats. The asterisk denotes data for 0:6 GSH:GSSG that were not fitted to a single exponential decay curve. (c) Relative percentage abundance of the γS dimer, monomer, and glutathionylated monomer (monomer-SG) determined by electrospray ionization MS. The errors are the standard deviation of three independent repeats.

ratios approach those found in aged and cataractous lenses.

MS was used to confirm the analytical SEC results and to detect any chemical modifications to γS upon reduction of the dimer. Masses for the disulfide-linked dimer and the monomer were detected as well as a third species corresponding to *S*-glutathionylated γS monomer (Fig. S2c). Semi-quantitative analysis of all detectable MS species was achieved with a standard curve using purified monomer and dimer (Fig. S2d). It is evident that the dimer ionizes poorly compared to the monomer, thus affecting quantification of the dimer by MS, particularly at lower concentrations, and thereby accounting for discrepancies in its abundance when compared to the concentration sensitive UV-detection in SEC. Nonetheless, the data acquired following 72 h of reduction showed that the abundance of the dimer decreases with increasing reduction potential, alongside a concomitant increase in the abundance of monomeric species (Fig. 2c), concordant with the trend observed by analytical SEC. The abundance of unmodified and *S*-glutathionylated monomer changes across the GSH:GSSG series. At 1:5 GSH:GSSG, for example, 85% of monomer is *S*-glutathionylated; however, this decreases linearly across the series, falling to 22% at 6:0 GSH:GSSG (Fig. 2c). The unmodified monomer consequently shows the inverse trend, consistent with the linear decrease in the monomer's retention volume at higher GSSG concentrations, as noted in SEC (Fig. S2b), implying that glutathionylation occurs via a disulfide exchange mechanism whereby the protein thiol reacts with GSSG. Furthermore, given that no glutathionylation was detected for the dimer (Fig. S2c), glutathionylation occurs at either C22, C24, or C26, which are otherwise disulfide bonded in the dimer. Examination of the unmodified monomer mass peak after reduction of the dimer reveals the highest intensity mass peak shifts with increased reduction potential from 20872 Da to 20874 Da at 1:5 and 6:0 GSH:GSSG, respectively (Fig. S2e). A similar trend can be observed for the *S*-glutathionylated monomer mass peak, largely shifting from 21177 to 21179 Da at 1:5 and 6:0 GSH:GSSG, respectively (Fig. S2f). This 2-Da difference implies that the C22–C26 intramolecular disulfide remains intact after initial reduction of the dimer, but the bond is ultimately reduced at higher reduction potentials. It also implies that γS glutathionylation occurs initially at C24, consistent with it being the most solvent-exposed cysteine (Fig. 1d).

Aggregation propensity is significantly increased when γS forms a disulfide-linked dimer

In view of the link between the oxidation of crystallins and cataract [41] and the increased aggregation propensity of the R14C γD disulfide-linked dimer [32], we investigated the aggregation propensity of the disulfide-linked γS dimer compared to its monomeric counterpart. A light scattering assay at 60 °C showed that the disulfide-linked γS dimer was far more aggregation prone than the monomer from 0.5 to 3.0 mg mL^{-1} (Fig. 3a). On average, the dimer's light scattering maximum was approximately 3.6 times greater than that of the monomer (Fig. 3b). Similarly, the rate of aggregation of the dimer was approximately 2.8 times faster (Fig. 3c). However, the lag time, reflecting the time required for the protein to form a critical nucleus (from which to seed aggregation), was approximately 1.2 times longer for the dimer than for the monomer (Fig. 3d). Thus, while the dimer aggregates faster and to a greater extent,

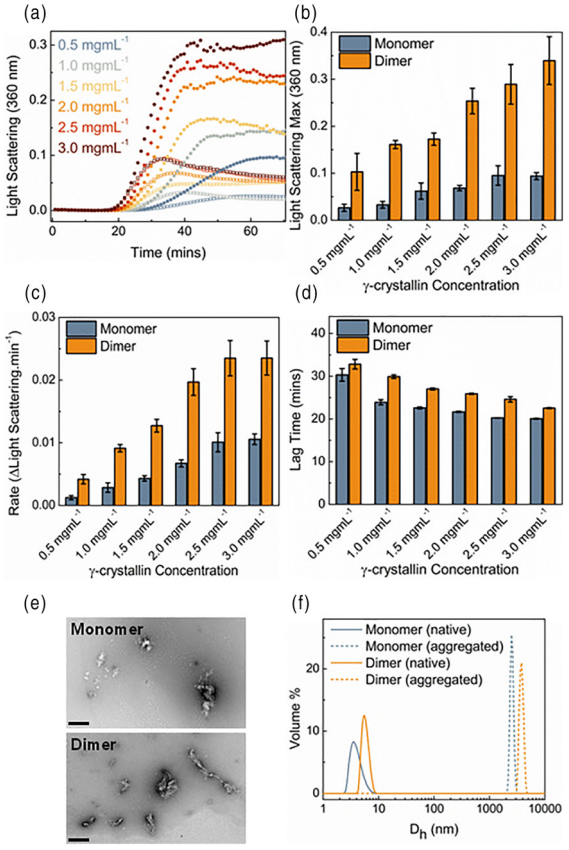


Fig. 3. Aggregation propensity of the γS monomer and dimer. (a) Light scattering (360 nm) assay of γS monomer (hollow squares) and dimer (filled circles) from 0.5 to 3.0 mg mL^{-1} with 0.5 mg mL^{-1} increments. Light scattering kinetic parameters including (b) maximum light scattering, (c) aggregation rate, and (d) lag time of aggregation for γS monomer (light blue) and dimer (orange). Errors are given as the standard deviation of three independent repeats. (e) Negative stain TEM of 1 mg mL^{-1} monomer (top) and dimer (bottom) aggregates formed after 2 h at 60 °C. The scale bar (black, bottom left) represents 500 nm. (f) Dynamic light scattering of 2 mg mL^{-1} native monomer (light blue, solid line) and dimer (orange, solid line) at 25 °C, and aggregated monomer (light blue, dashed line) and dimer (orange, dashed line) after 35 min at 60 °C.

intermolecular interactions in the early stages of aggregation are hindered in comparison to the monomer. The reduced accessibility of specific interfacial residues in the dimer, along with steric considerations, may be factors underlying the extended lag time for the dimer. Transmission electron microscopy (TEM) showed that the aggregates formed by the monomer (Fig. 3e, top) and dimer (Fig. 3e, bottom) appeared essentially amorphous in form with little ordered structure. Dynamic light scattering showed that the aggregates formed by the dimer were, on average, larger than those of the monomer. Initially, the native monomer and dimer measured at 25 °C had an average hydrodynamic diameter (D_h) of 4.0 ± 0.9 and 5.8 ± 0.8 nm, respectively (Fig. 3f, solid line). Upon incubation at 60 °C, there was a rapid, time-dependent increase in D_h for both the monomer and dimer, as anticipated. After 35 min of incubation, the approximate time at which a steady state was observed in the light scattering assay (Fig. 3a), the monomer and dimer exhibited a D_h of 2,541 ± 173 and 3,787 ± 301 nm, respectively (Fig. 3f, dashed line). Thus, the increased light scattering in solution is due, in part, to the increased hydrodynamic size of the dimer aggregates, as well as a potentially higher yield of aggregated protein.

γS dimer undergoes non-cooperative thermal unfolding despite having a similar structure to the monomer

Far-UV circular dichroism (CD) spectra and 8-anilinonaphthalene-1-sulfonic acid (ANS) binding experiments, respectively, showed that the overall secondary structure (Fig. 4a) and surface hydrophobicity (Fig. 4b) of γS are not significantly altered upon dimerization. Indeed, structural alignment of the γS dimer with three other γS structures, that is, the human γS C-terminal domain crystal structure [42] (PDB: 1HA4), the human γS NMR structure [40] (PDB: 2M3T), and the chicken γS crystal structure [43] (PDB: 5VH1) (Fig. 4c–e, respectively), gave C_α RMSD values no greater than 1.9 Å, indicating that the overall fold of the dimer is similar to those previously determined for γS. However, the dimer shows increased intrinsic tryptophan fluorescence relative to that of the monomer (Fig. 4c), suggesting that the fluorescence of one or more of the protein's four tryptophan residues is less efficiently quenched in the dimeric form. Close examination of the environment of the tryptophan residues within the dimer structure revealed that the otherwise highly quenched W162 is not coordinated to a water molecule, in contrast to the truncated γS C-terminal domain structure (PBD: 1HA4), thereby having an impact on the efficient quenching of W162 and potentially accounting for the higher quantum yield [44].

In the absence of evidence of any significant perturbations in the native structure of the disulfide-linked dimer, we examined its thermal stability in an effort to account for its increased aggregation propensity. Four spectroscopic methods, intrinsic tryptophan fluorescence, light scattering, ANS fluorescence, and CD (Fig. 5a–d, respectively), were used to compare the thermal stability of the monomer and dimer. It is apparent in all the unfolding curves that the dimer exhibits biphasic behavior while the monomer is monophasic in its unfolding. It is concluded that the dimer unfolds non-cooperatively, that is, one domain before the other, whereas unfolding of the monomer is concerted (Fig. 5e). Similar thermally induced biphasic unfolding has been noted in the

cataract-associated G18V variant in monomeric γS [45]. From the data in Fig. 5, a mid-point of thermal unfolding (T_m) or aggregation (T_{agg}) was calculated for each transition (Table S1), which showed the mid-point of the dimer's first transition to be lower than that for the single value of the monomer but the second transition to be at a higher temperature. Examination of the difference in T_m values (ΔT_m) between the monomer and those of the dimer and comparison to ΔT_m values reported in the literature for γS and its independent domains (Fig. S3a) implied that the first and second transitions of the dimer represent the unfolding of the N- and C-terminal domains, respectively. For example, while Mills *et al.* [46] reported a ΔT_m for the isolated N- and C-terminal domains of γS of 6.0 °C, the ΔT_m between the first and second transition of the dimer is 6.7 °C (Fig. S3b). Similarly, the ΔT_m between the full-length monomer and each domain [46] is comparable to the ΔT_m between the monomer and each transition of the dimer, implying that domain stability differences are largely conserved and, moreover, that unfolding of the N-terminal domain precedes that of the C-terminal domain in the γS dimer (Fig. 5f). Interestingly, the G18V variant, which also exhibits biphasic thermal unfolding, gave a ΔT_m between its first and second transition of 8.0 °C [45], comparable with the ΔT_m of 7.9 °C between the first and second transition of the dimer (Fig. S3c). Potentially, therefore, the biphasic unfolding induced in γS by the cataract-associated G18V mutation parallels the behavior induced by disulfide-linked dimerization.

Discussion

The evidence for cysteine oxidation, including disulfide crosslinks, between crystallin proteins in aging and cataractous lenses is extensive [16,17, 21–23,47]. Previous studies have shown that disulfide bonds can form *in vitro* in γ-crystallins [18,31, 32,48], but most studies do not demonstrate the ability for these disulfides to be viable in a reducing environment commensurate with that of the lens. Human lens GSH concentrations decrease with age

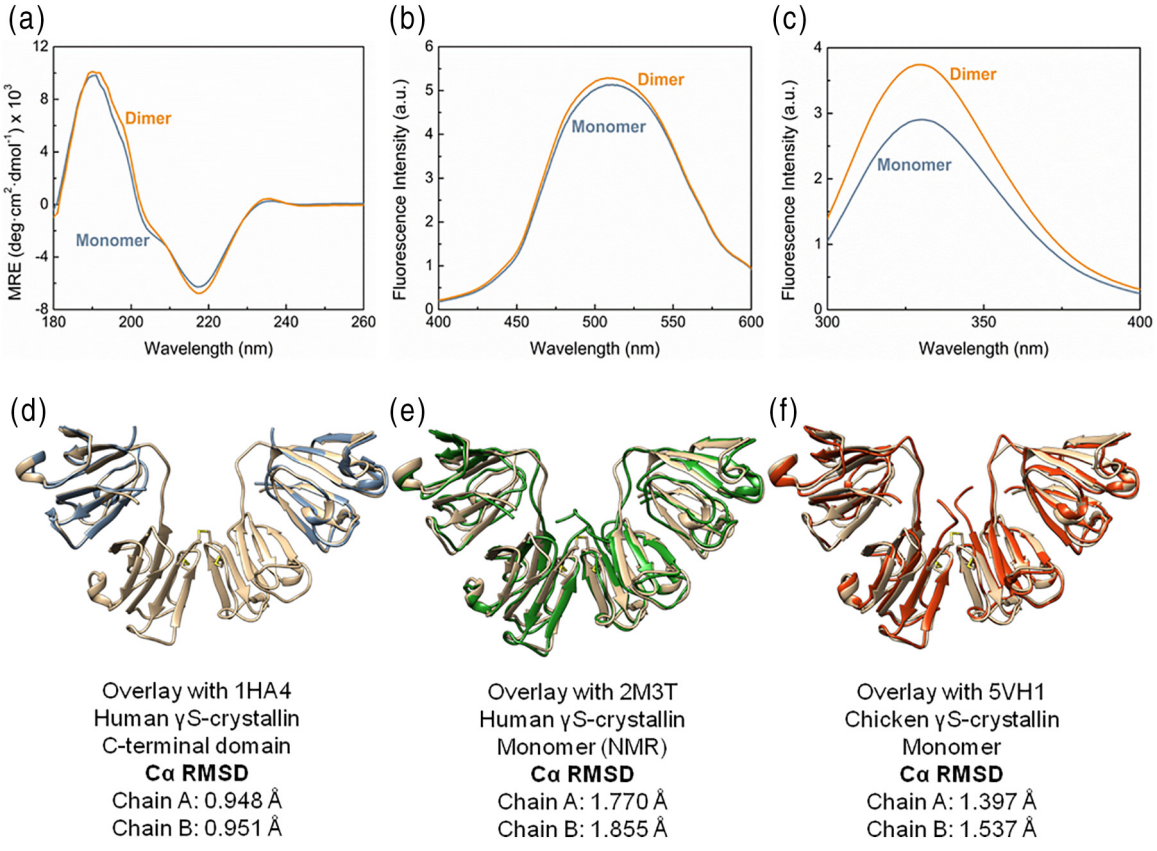


Fig. 4. Comparison of the secondary structure, surface hydrophobicity and overall conformation of the γS monomer, γS dimer and related structures. (a) Far-UV CD, (b) ANS fluorescence, and (c) tryptophan fluorescence spectra of γS monomer (light blue) and dimer (orange). Structural overlay between human γS dimer (beige) and (d) human γS C-terminal domain crystal structure (blue; PDB: 1HA4), (e) human γS NMR structure (green; PDB: 2M3T) and (f) chicken γS crystal structure (red; PDB: 5VH1). The RMSD is given as a quantitative measure of structural similarity.

from approximately 6 to 2 mM and can fall below 1 mM in cataract [24–26]. We have shown that the γS disulfide-linked dimer becomes increasingly stable at GSH concentrations found in aging and cataractous lenses. While these studies were completed at physiological GSH concentrations, it is worth noting that the concentration of γS in the lens is likely to be in the order of 100 times more than that studied here [6,49]. Furthermore, the aging, GSH-depleted lens would need to contend with the stress associated with the oxidation of other lenticular components [50]. Thus, we consider the experimental values of dimer abundance determined herein to be conservative estimates in relation to the prevalence of the disulfide-linked γS dimer *in vivo*.

Glutathionylation of crystallin proteins in normal and cataractous human lenses has been noted numerous times with the prevalence of this modification being positively correlated with aging [24,51–53]. γS extracted from the human lens is *S*-glutathionylated at two sites, with one modification at either C22, C24, or C26 and another at C82 [30]. We observed a single glutathionylation of γS by MS and inferred that the modification occurred at C24, consistent with its high solvent exposure (Fig. 1d). In natively folded γS, C82 is unlikely to be glutathionylated due to its very low solvent exposure, and as such, glutathionylation at C82 *in vivo* is likely an indicator of partial unfolding of the N-terminal domain. The abundance of glutathionylated γS monomer is directly proportional to the concentration of GSSG (Fig. 2c and S2b), indicative of mixed disulfide formation via an exchange mechanism which has been previously noted as the preferred mechanism for the addition of glutathione to crystallins [52,54]. Furthermore, the amount of glutathionylation at 6:0 GSH:GSSG is likely due to the reduction of dimer consequently forming GSSG. This mechanism is consistent with C24 being present largely in the reduced thiol (protonated) form at physiological p*H*, due to an elevated pK_a value as a result of flanking negatively charged aspartyl residues in the "DCDCDC" motif from residues 21–26 [31]. A disulfide exchange mechanism also has implications for the formation of the γS dimer, which could foreseeably arise via a reaction between *S*-glutathionylated and unmodified γS monomers, resulting in the formation of a protein-protein disulfide link and GSH (Fig. 6). Indeed, γS has been proposed to function as a "redox sink" in the aging lens, that is, acting as a reducing agent to consume GSSG and replenish GSH levels in the process of forming disulfide-linked dimers [6,31]. However, given the enhanced aggregation propensity of the γS dimer, such replacement would be a trade-off in order to maintain a reducing environment in the lens.

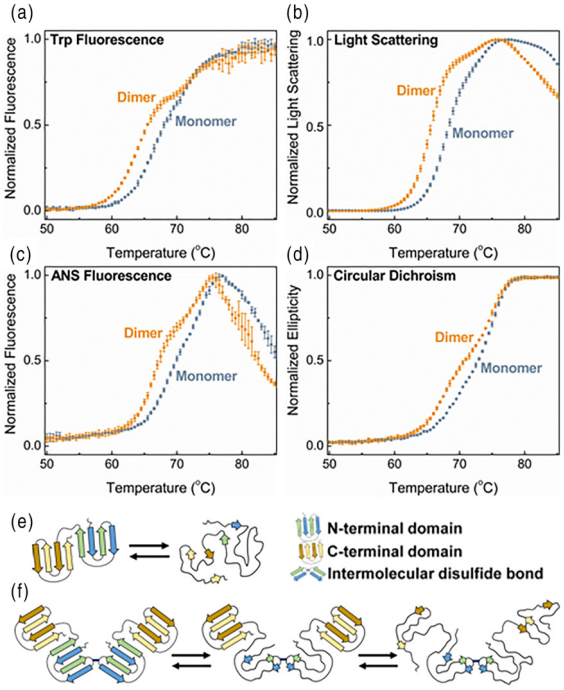


Fig. 5. Thermal stability of the γS monomer and dimer. The γS monomer (light blue) and dimer (orange) were monitored using (a) tryptophan fluorescence (Trp) (ratio 345/329 nm), (b) light scattering at 360 nm, (c) ANS fluorescence at 480 nm, and (d) CD at 218 nm. Curves are the average of three independent repeats, and the errors are given as standard deviation. (e) Schematic diagram depicting how the γS monomer unfolds cooperatively, while the γS dimer (f) unfolds non-cooperatively, initially via its N-terminal domain. Structural features of the monomer and dimer are denoted in the key adjacent to panel e.

The link between PTMs, particularly oxidation, disulfide bond formation, and cataract is prevalent, with a prevailing sentiment being that the formation of age-related cataract is a function of PTMs, heat, and time [14–16]. Single-residue mutations such as G18V in γS [45] and R14C (which can form disulfide-linked dimers) in γD [32] have a heightened aggregation propensity and are associated with cataract. However, both variants are structurally comparable to their wild-type counterparts. Similarly, the γS disulfide-linked dimer is also far more aggregation prone, but its structure is not altered when compared to the monomer. The thermal denaturation of the γS dimer also resembles that of the cataract-associated G18V variant [45]. Overall, thermal denaturation indicates that disulfide-linked dimerization of γS decouples the unfolding cooperativity of the N- and C-terminal domains. However, while a loss of cooperativity is evident, the relative stability difference between the N- and C-terminal domains does not appear to be affected upon comparison with the isolated N- and C-terminal domains (Fig. S3a). Possible causes of this are likely to be related to subtle changes in domain flexibility and altered interactions with water upon dimerization. Indeed, molecular dynamics studies on

homology structures of wild-type and G18V γS reveal that changes in inter-strand salt-bridge interactions accompany changes in dynamics that allow water permeable openings in the N-terminal domain of G18V [55]. This could further disrupt bonding networks that lead to a loss of the cooperative unfolding mechanism between the two domains as has been noted for this cataract-associated variant [45]. Furthermore, NMR studies on murine cataract-associated F9S, which also exhibits non-cooperative domain unfolding, show changes to the tryptophan (a "Trp corner") involved in hydrogen bonding and stabilization of a β-turn in the N-terminal domain. As Trp or Tyr corners have been intimately linked to the folding/unfolding mechanism of Greek-key folds [56,57], this could highlight an additional role in the maintenance of folding/unfolding cooperativity.

The biphasic unfolding of the dimer indicates a thermodynamically stabilized intermediate, which is likely characterized by an unfolded N-terminal domain and a folded C-terminal domain, as has been observed in other destabilized γ-crystallins [18,45,58–60]. Consistent with this, molecular dynamics simulations have implied that γD can aggregate through domain swapping via a partially folded intermediate characterized by an unfolded N-terminal domain but an otherwise folded C-terminal domain [61]. Indeed, the γS disulfide-linked dimer is far more aggregation prone compared to the monomer, displaying ostensibly amorphous looking aggregates via TEM (Fig. 3e, bottom). However, it has been demonstrated that partially ordered P23T γD aggregates are indistinguishable from amorphous aggregates when viewed by TEM [62], indicating that further investigation into the substructure of the aggregates observed here is needed to obtain a more complete understanding of their molecular underpinnings. The present study suggests that the disulfide-linked γS dimer is highly analogous to the G18V and dimer-forming R14C variants, which are both associated with progressive juvenile-onset cataract [32,45]. Such cataract-associated variants therefore appear to exhibit structural markers similar to those that may emerge in wild-type γ-crystallins later in life due to aging, for example, the formation of non-native disulfides [18]. We propose that oxidation and subsequent dimerization of γS preferentially destabilizes its N-terminal domain, as is evident in G18V, resulting in a thermodynamically stable intermediate that exhibits increased aggregation and thereby exacerbates light scattering associated with age-related cataract (Fig. 6).

In summary, we have characterized an oxidative PTM of human γS and provided a molecular basis for its role in age-related cataract. Determination of the crystal structure of the oxidation product showed γS in the form of a dimer linked via an intermolecular disulfide bond at C24 [31] and also revealed an intramolecular disulfide between C22 and C26. No gross conformational changes were observed upon oxidation. We have demonstrated that this PTM is stable at GSH concentrations akin to those in aged and cataractous lenses and the PTM increases the propensity of γS to aggregate. Along with other conformational features shared by the γS dimer and cataract-associated variants, these findings strongly suggest that the occurrence of this PTM in the aging lens leads to or worsens the severity of cataract. The mechanism outlined here may apply more broadly to other crystallin proteins that undergo oxidative PTMs as part of the oxidation-driven aggregation cascade that underlies age-related cataract.

Materials and Methods

Production of γS monomer and dimer

A pET43.1 plasmid encoding recombinant human γS (178 amino acids; UniProt P22914) was purchased from Genscript and expressed in BL21(DE3) *E. coli* cells. Cells were cultured initially at 37 °C for 4–5 h. Expression was then induced with 500 μM IPTG, and the cell culture was incubated overnight at 30 °C. Cells were pelleted, resuspended in DEAE column buffer (20 mM Tris–HCl, pH 8.0), and lysed using sonication. The cell lysate was loaded onto an anion-exchange column (HiPrep DEAE FF 16/10; GE Healthcare) and a peak containing γS eluted in the flow-through. Fractions containing predominantly γS were concentrated to ~1 mL, loaded onto a preparative SEC column (HiLoad 16/600 Superdex 75 pg; GE Healthcare) and separated in PBS (50 mM sodium phosphate, 100 mM NaCl, pH 7.0. The peak corresponding to monomeric γS was collected. This initial fraction of monomeric γS was concentrated to a minimum of 20 mg mL^{-1} using an Amicon Ultra-15 Centrifugal Filter Unit (Merck Millipore) and left at 4 °C for 1 week to allow for large-scale dimerization of γS (up to 60% of the total protein). Monomeric and dimeric forms of γS were then purified by two sequential separations via preparative SEC, eluting with 20 mM sodium phosphate buffer (pH 7.0). Within 30 min of elution, respective fractions of monomer and dimer were frozen with dry ice in 0.2–0.5 mL aliquots at protein concentrations of 4-8 mg mL^{-1}. Aliquots were thawed immediately prior to use.

Crystallization, data collection, and refinement

Purified γS dimer was buffer exchanged to 10 mM HEPES (pH 7.2) and concentrated to 10 mg mL^{-1}. Crystals were obtained by vapor diffusion in sitting drops containing 0.2 M sodium tartrate dibasic dihydrate (pH 7.3) and 20% w/v polyethylene glycol

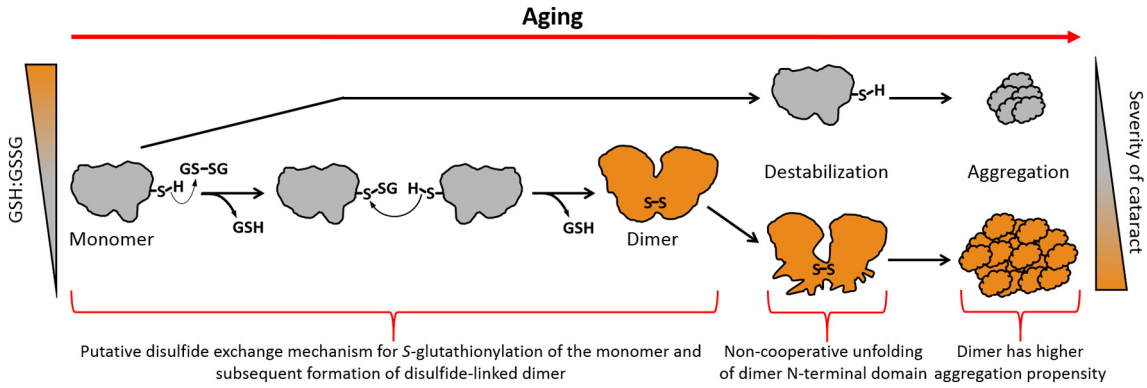


Fig. 6. Putative schematic representation of the process of oxidation/dimerization, destabilization, and aggregation of human γS.

3350 (Hampton Research). For cryoprotection, drops were supplemented with 30% w/v polyethylene glycol 3350 and crystals were subsequently vitrified by submersion in liquid nitrogen. Diffraction data were collected at a wavelength of 0.9763 Å on the MX2 beam line of the Australian Synchrotron. Diffraction data were indexed, scaled, and integrated using XDS [63]. Phases were obtained, via molecular replacement, using PhaserMR [64] with the crystal structure of chicken γS [43] (PDB 5VH1) as the search model. Cycles of refinement and manual rebuilding were completed using PHENIX and Coot, respectively [65,66].

SAXS measurements and *ab initio* envelope reconstruction

Experiments were conducted on a NanoSTAR II SAXS (Bruker) with a rotating anode Cu Kα radiation source at a wavelength of 1.541 Å and using a q range of 0.013–0.39 $Å^{-1}$. Samples of γS monomer (3 mg mL^{-1}) and dimer (6 mg mL^{-1}) were prepared in 20 mM sodium phosphate (pH 7.0). Samples were loaded into a sealed 2 mm quartz capillary cell and placed under vacuum. Scattering was collected in 1-h increments for a total of 4 h in a temperature controlled sample environment at 25 °C. No change in the scattering of the monomer or dimer was observed over the 4-h acquisition period. The data were averaged and subtracted from the scattering of the buffer using PRIMUS [67]. Low q points near the beam stop were excluded. The P(r) and the R_g were calculated using GNOM [68]. Using CRYSOL, the scattering profiles of the γS monomer (PBD: 2M3T) and dimer structures (PDB: 6FD8) were modeled and fitted to their respective experimental scattering profiles [69]. Using GASBOR [70], 15 *ab initio* shape envelopes were initially generated which were then aligned, filtered, and averaged using the DAMAVER suite [71]. This averaged structure underwent refinement using DAMMIN [72] to produce the final *ab initio* envelope. Using UCSF Chimera [73], the final envelope was converted to a volumetric map, contoured to 2σ and aligned with the γS dimer crystal structure, giving a correlation coefficient for the fit of 0.89.

Structural and sequence analysis

The SAS of all cysteine residues in the monomeric structures of γB (PDB: 2JDF), γC (PDB: 2NBR), γD (PDB: 1HK0), and γS (PDB: 2M3T) was investigated using UCSF Chimera [73]. In the case of NMR-derived structures (γC and γS), the SAS was averaged across all conformers. The SAS was converted to a percentage relative to the side-chain SAS of a Cys in a Gly-Cys-Gly tripeptide in an extended conformation [33]. Sequence alignments were generated using the Clustal Omega server [74] and colored according to the consensus symbols generated on the basis of their Gonnet PAM 250 matrix score [≤0.5 (.) weakly conserved (green); >0.5 (:) strongly conserved (blue); = 1.0 (*) invariant (orange)]. The phylogenetic tree was constructed using the National Center for Biotechnology Information taxonomy database. Structural overlay of the γS dimer was completed using UCSF Chimera and the RMSD of the C_α atoms was reported as the measure of structural similarity between other γS structures with the PDB accession codes 1HA4, 2M3T, and 5VH1.

Reduced and oxidized glutathione treatments of γS dimer

The γS dimer at a concentration of 0.3 mg mL^{-1} was incubated in filtered (0.44 μm) 20 mM sodium phosphate (pH 7.0) at room temperature with different ratios of GSH (G4251, Sigma) and GSSG (G4376, Sigma). GSH:GSSG molar ratios were 0:6, 0.5:5.5, 1:5, 2:4, 3:3, 4:2, 5:1, and 6:0, maintaining a 6 mM final concentration of glutathione.

Analytical SEC

γS dimer was prepared with GSH:GSSG ratios as previously detailed. Following treatment for 0, 4, 8, 12, 24, 48, and 72 h, γS dimer was separated into its monomeric and dimeric components using a Superdex 75 10/300 GL SEC (GE Healthcare) column attached to an ÄKTA Pure FPLC (GE Healthcare) with UV absorbance (280 nm) detection. Monomer and dimer peaks were baseline corrected and integrated using Unicorn 6.3 software (GE Healthcare). All experiments were performed at room temperature using filtered (0.44 μm) 20 mM sodium phosphate (pH 7.0) at a flow rate of 0.8 mL min^{-1}.

Mass spectrometry

Experiments were performed using an Orbitrap Elite mass spectrometer equipped with a HESI-II electrospray ionization source coupled to an UltiMate 3000 UHPLC (Thermo Scientific). γS dimer was prepared with GSH:GSSG ratios as previously detailed. Samples (7 μL) were injected into the mass analyzer following treatment for 24 and 72 h. A mass range of 200 to 4000 at a resolution of 240,000 was acquired. The data were extracted and analyzed using Thermo Xcalibur Qual software. Peaks corresponding to the monomer, monomer–glutathione adduct, and dimer were plotted as intensity *versus* mass (Fig. S2c) and quantified (Fig. 2c) using standard curves. Standard curves were generated using purified γS monomer and dimer at approximately 0.02, 0.05, 0.1, 0.2, 0.3, 0.35, and 0.4 mg mL^{-1} and mixed to achieve a total concentration of approximately 0.4 mg mL^{-1} (Fig. S2d). The standard curve samples were mixed to diminish any effects on the final intensity count due to ionization transfer between monomer and dimer species during experiments where the monomer was being generated from the dimer. The ionization intensity of the monomer–glutathione adduct was assumed to be the same as that of the monomer.

Transmission electron microscopy

Samples for TEM were prepared by adding 2 μL of 1 mg mL^{-1} γS monomer or dimer incubated for 2 h at 60 °C to Formvar and carbon-coated copper grids (ProSciTech, Australia). The grids were then washed three times with 10 μL of Milli-Q water and negatively stained with 10 μL of uranyl acetate (2% w/v). Samples were viewed using a Hitachi H7100FA transmission electron microscope (Tokyo, Japan).

Thermal stability

Experiments were performed using an Applied Photophysics Chirascan spectrophotometer attached to a Quantum Northwest TC 125 PELTIER temperature controller. CD measurements were acquired in a 0.1 cm pathlength quartz cuvette, while light scattering and Trp/ANS fluorescence experiments were acquired in a 1 cm pathlength quartz cuvette. All cuvettes were fitted with stoppers to prevent evaporation. Proteins were prepared in 20 mM sodium phosphate (pH 7.0) at a protein concentration of 0.3 mg mL^{-1}. Thermal stability was assessed by ramping the temperature in 0.5 °C increments from 25 to 90 °C. At each increment, measurements at a single wavelength were acquired for 1.5 s for 5 repeats with a 30 s equilibration between each temperature increment. Trp and ANS fluorescence were measured at excitation wavelengths of 295 and 350 nm, respectively. The concentration of ANS used was 300 μM. The acquisition wavelengths were 218 nm (CD), 345/329 nm (Trp fluorescence), 360 nm (light scattering), and 480 nm (ANS fluorescence). Experiments were performed in triplicate. A single or double Boltzmann sigmoidal function was fitted to the thermal unfolding curves to obtain the T_m or T_{agg} values.

Aggregation assay

Kinetic aggregation assays were performed using a Biotek Synergy 2 microplate reader. Protein aggregation was induced at 60 °C and monitored using light scattering at 360 nm. A "slow shaking" setting was engaged for the duration of the assay. The aggregation of γS monomer and dimer was studied at 0.5, 1.0, 1.5, 2.0, 2.5, and 3.0 mg mL^{-1}. Aggregation kinetics were fitted to a single Boltzmann function using Origin (OriginLab Corporation) and the fitting parameters were used to calculate the lag time and rate of aggregation as previously described [75].

Dynamic light scattering

Experiments were performed using a Zetasizer Nano ZS (Malvern Instruments) with a built-in PELTIER temperature control system. A standard He–Ne laser operating at a wavelength of 633 nm and a scattering detection angle of 173° (backscattering) was used for data collection. γS monomer and dimer solutions were filtered through Millex Durapore (0.22 μm) filters and subsequently loaded into a semi-micro quartz cuvette fitted with a stopper. The final concentration of protein was 2 mg mL^{-1} in 20 mM sodium phosphate (pH 7.0). Size measurements of the native monomer and dimer were performed at 25 °C. The sample was then placed at 60 °C, and aggregation was monitored continuously until the protein was fully aggregated. Size measurements collected within each 2-min increment were averaged. Figure 3e shows data for the fully aggregated sample acquired after 35 min, consistent with the time at which a steady state is reached in aggregation assays monitored by light scattering.

SEC with multi-angle light scattering detection

γS monomer and dimer were loaded onto a Superdex 75 10/300 GL SEC (GE Healthcare) column in filtered (0.22 μm) 20 mM sodium phosphate (pH 7.0) with multi-angle light scattering (DAWN HELEOS 8; Wyatt Technologies) and refractive index detection (Optilab rEX; Wyatt Technologies). The multi-angle detectors were normalized using monomeric bovine serum albumin (Sigma, A1900). A *dn*/*dc* value of 0.1983 mL g^{-1} for human γS [3] was used. The data were processed using ASTRA (Wyatt Technologies).

Spectroscopy

CD and ANS fluorescence spectra were acquired in 2 mM sodium phosphate (pH 7.0), while Trp fluorescence was acquired in 20 mM sodium phosphate (pH 7.0). All experiments were performed at a concentration of 0.3 mg mL^{-1} at 25 °C. CD spectra were acquired from 180 to 260 nm, ANS fluorescence from 400 to 600 nm, and Trp fluorescence from 300 to 400 nm. Fluorescence parameters were the same as those used for the thermal stability experiments, except that the excitation bandwidth was set to 5 nm for ANS fluorescence experiments. Three repeats were acquired with a step width of 1 nm for 4 s. These parameters were used for three independent experiments, which were averaged to produce the final spectrum.

Data deposition and accession numbers

The coordinates and structure factors have been deposited in the Protein Data Bank with accession number PDB ID 6FD8. SAXS data were deposited in the Small Angle Scattering Biological Data Bank (SASBDB) under the accession codes SASDEZ6 (gamma-crystallin S disulfide-linked dimer) and SASDE27 (gamma-crystallin S monomer).

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2018.12.005.

CRediT authorship contribution statement

David C. Thorn: Conceptualization, Formal analysis, Investigation, Writing - review & editing. **Aidan B. Grosas:** Conceptualization, Data curation, Formal analysis, Investigation, Visualization, Writing - original draft, Writing - review & editing. **Peter D. Mabbitt:** Formal analysis, Writing - review & editing, . **Nicholas J. Ray:** Investigation. **Colin J. Jackson:** Formal analysis, Resources, Supervision, Writing - review & editing. **John A. Carver:** Conceptualization, Funding acquisition, Project administration, Supervision, Writing - review & editing.

Acknowledgments

We thank Prof. Roger Truscott and Ms. Jen Xiang for insightful discussions. We acknowledge Dr. Paul Carr for helpful discussions related to protein crystallography. The Centre for Advanced Microscopy at ANU is acknowledged for their assistance with TEM. We are grateful to Dr. Adam Carroll and the team at the Joint Mass Spectrometry Facility, ANU, for help and guidance in performing and designing mass spectrometry experiments. We are grateful to Dr. Robert Knott at the Australian Nuclear Science and Technology Organisation, Lucas Heights, for his assistance with SAXS measurements. A.B.G. is supported by an Australian Postgraduate Award. D.C.T. is supported by a grant from the National Health and Medical Research Council of Australia (Grant No. 1068087) awarded to J.A.C.

Conflict of Interest: The authors declare no conflict of interest.

Received 31 October 2018;
Accepted 5 December 2018
Available online 13 December 2018

Keywords:
crystallin;
disulfide bond;
dimer;
oxidation;
cataract

†D.C.T. and A.B.G. contributed equally to this work.

Current address: P.D. Mabbitt, MRC Protein Phosphorylation and Ubiquitination Unit, University of Dundee, Dundee, UK.

Abbreviations used:
γS, γS-crystallin; PTM, post-translational modification; MS, mass spectrometry; SAS, solvent accessible surface area; GSH, reduced glutathione; GSSG, oxidized glutathione; ANS, 8-anilino-1-naphthalenesulfonic acid; CD, circular dichroism; TEM, transmission electron microscopy.

References

[1] M. Delaye, A. Tardieu, Short-range order of crystallin proteins accounts for eye lens transparency, Nature 302 (1983) 415.

[2] A. Ponce, C. Sorensen, L. Takemoto, Role of short-range protein interactions in lens opacifications, Mol. Vis. 12 (2006) 879–884.

[3] H. Zhao, P.H. Brown, M.T. Magone, P. Schuck, The molecular refractive function of lens γ-crystallins, J. Mol. Biol. 411 (2011) 680–699.
[4] H.P. Driessen, P. Herbrink, H. Bloemendal, W.W. de Jong, Primary structure of the bovine β-crystallin Bp chain: internal duplication and homology with γ-crystallin, Eur. J. Biochem. 121 (1981) 83–91.
[5] D. Hogg, L. Tsui, M. Gorin, M. Breitman, Characterization of the human β-crystallin gene Hu β A3/A1 reveals ancestral relationships among the beta gamma-crystallin superfamily, J. Biol. Chem. 261 (1986) 12420–12427.
[6] H. Bloemendal, W. de Jong, R. Jaenicke, N.H. Lubsen, C. Slingsby, A. Tardieu, Ageing and vision: structure, stability and function of lens crystallins, Prog. Biophys. Mol. Biol. 86 (2004) 407–485.
[7] T. Blundell, P. Lindley, L. Miller, D. Moss, C. Slingsby, I. Tickle, B. Turnell, G. Wistow, The molecular structure and stability of the eye lens: X-ray analysis of γ-crystallin II, Nature 289 (1981) 771.
[8] G. Inana, J. Piatigorsky, B. Norman, C. Slingsby, T. Blundell, Gene and protein structure of a β-crystallin polypeptide in murine lens: relationship of exons and structural motifs, Nature 302 (1983) 310.
[9] C. Slingsby, N.J. Clout, Structure of the crystallins, Eye 13 (1999) 395.
[10] S. Bassnett, Lens organelle degradation, Exp. Eye Res. 74 (2002) 1–6.
[11] N. Lynnerup, H. Kjeldsen, S. Heegaard, C. Jacobsen, J. Heinemeier, Radiocarbon dating of the human eye lens crystallines reveal proteins without carbon turnover throughout life, PLoS One 3 (2008), e1529.
[12] B.H. Toyama, J.N. Savas, S.K. Park, M.S. Harris, N.T. Ingolia, J.R. Yates III, M.W. Hetzer, Identification of long-lived proteins reveals exceptional stability of essential cellular structures, Cell 154 (2013) 971–982.
[13] M.G. Friedrich, R.J. Truscott, Membrane association of proteins in the aging human lens: profound changes take place in the fifth decade of life, Invest. Ophthalmol. Vis. Sci. 50 (2009) 4786–4793.
[14] K.L. Moreau, J.A. King, Protein misfolding and aggregation in cataract disease and prospects for prevention, Trends Mol. Med. 18 (2012) 273–282.
[15] R.J. Truscott, M.G. Friedrich, The etiology of human age-related cataract. Proteins don't last forever, Biochim. Biophys. Acta Gen. Subj. 1860 (2016) 192–198.
[16] R.J. Truscott, Age-related nuclear cataract—oxidation is the key, Exp. Eye Res. 80 (2005) 709–725.
[17] P.G. Hains, R.J. Truscott, Proteomic analysis of the oxidation of cysteine residues in human age-related nuclear cataract lenses, Biochim. Biophys. Acta Proteins Proteom. 1784 (2008) 1959–1964.
[18] E. Serebryany, J.C. Woodard, B.V. Adkar, M. Shabab, J.A. King, E.I. Shakhnovich, An internal disulfide locks a misfolded aggregation-prone intermediate in cataract-linked mutants of human γD-crystallin, J. Biol. Chem. 291 (2016) 19172–19183.
[19] E. Serebryany, T. Takata, E. Erickson, N. Schafheimer, Y. Wang, J.A. King, Aggregation of Trp > Glu point mutants of human gamma-D crystallin provides a model for hereditary or UV-induced cataract, Protein Sci. 25 (2016) 1115–1128.
[20] S. Ramkumar, X. Fan, B. Wang, S. Yang, V.M. Monnier, Reactive cysteine residues in the oxidative dimerization and Cu^{2+} induced aggregation of human γD-crystallin: implications for age-related cataract, Biochim. Biophys. Acta Mol. Basis Dis. 1864 (2018) 3595–3604.
[21] R.J. Truscott, R.C. Augusteyn, Oxidative changes in human lens proteins during senile nuclear cataract formation, Biochim. Biophys. Acta 492 (1977) 43–52.
[22] A. Spector, D. Roy, Disulfide-linked high molecular weight protein associated with human cataract, Proc. Natl. Acad. Sci. U. S. A. 75 (1978) 3244–3248.
[23] S.R.A. Hanson, A. Hasan, D.L. Smith, J.B. Smith, 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 (2000) 195–207.
[24] J.J. Harding, Free and protein-bound glutathione in normal and cataractous human lenses, Biochem. J. 117 (1970) 957–960.
[25] A. Kamei, Glutathione levels of the human crystalline lens in aging and its antioxidant effect against the oxidation of lens proteins, Biol. Pharm. Bull. 16 (1993) 870–875.
[26] L.M. Bova, M.H.J. Sweeney, J.F. Jamie, R.J.W. Truscott, Major changes in human ocular UV protection with age, Invest. Ophthalmol. Vis. Sci. 42 (2001) 200–205.
[27] K.J. Lampi, Z. Ma, M. Shih, T.R. Shearer, J.B. Smith, D.L. Smith, L.L. David, Sequence analysis of βA3, βB3, and βA4 crystallins completes the identification of the major proteins in young human lens, J. Biol. Chem. 272 (1997) 2268–2275.
[28] J.A. Thomson, R.C. Augusteyn, Ontogeny of human lens crystallins, Exp. Eye Res. 40 (1985) 393–410.
[29] V.N. Lapko, D.L. Smith, J.B. Smith, *S*-methylated cysteines in human lens γS-crystallins, Biochemistry 41 (2002) 14645–14651.
[30] J. Craghill, A.D. Cronshaw, J.J. Harding, The identification of a reaction site of glutathione mixed-disulphide formation on γS-crystallin in human lens, Biochem. J. 379 (2004) 595–600.
[31] F. Skouri-Panet, F. Bonneté, K. Prat, O.A. Bateman, N.H. Lubsen, A. Tardieu, Lens crystallins and oxidation: the special case of γS, Biophys. J. 89 (2001) 65–76.
[32] A. Pande, J. Pande, N. Asherie, A. Lomakin, O. Ogun, J. A. King, N.H. Lubsen, D. Walton, G.B. Benedek, Molecular basis of a progressive juvenile-onset hereditary cataract, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 1993–1998.
[33] S. Miller, J. Janin, A.M. Lesk, C. Chothia, Interior and surface of monomeric proteins, J. Mol. Biol. 196 (1987) 641–656.
[34] M.A. Smith, O.A. Bateman, R. Jaenicke, C. Slingsby, Mutation of interfaces in domain-swapped human βB2-crystallin, Protein Sci. 16 (2007) 615–625.
[35] Z. Xi, M.J. Whitley, A.M. Gronenborn, Human βB2-crystallin forms a face-en-face dimer in solution: an integrated NMR and SAXS study, Structure 25 (2017) 496–505.
[36] B. Bax, R. Lapatto, V. Nalini, H. Driessen, P. Lindley, D. Mahadevan, T. Blundell, C. Slingsby, X-ray analysis of βB2-crystallin and evolution of oligomeric lens proteins, Nature 347 (1990) 776.
[37] E. Krissinel, K. Henrick, Inference of macromolecular assemblies from crystalline state, J. Mol. Biol. 372 (2007) 774–797.
[38] P.G. Cooper, J.A. Carver, J.A. Aquilina, G.B. Ralston, R.J. Truscott, A ^{1}H NMR spectroscopic comparison of γS- and γB-crystallins, Exp. Eye Res. 59 (1994) 211–220.

[39] J.A. Carver, P.G. Cooper, R.J.W. Truscott, 1H-NMR spectroscopy of βB2-crystallin from bovine eye lens, Eur. J. Biochem. 213 (1993) 313–320.
[40] C.N. Kingsley, W.D. Brubaker, S. Markovic, A. Diehl, A.J. Brindley, H. Oschkinat, R.W. Martin, Preferential and specific binding of human αB-crystallin to a cataract-related variant of γS-crystallin, Structure 21 (2013) 2221–2227.
[41] A. Taylor, K.J.A. Davies, Protein oxidation and loss of protease activity may lead to cataract formation in the aged lens, Free Radic. Biol. Med. 3 (1987) 371–377.
[42] A.G. Purkiss, O.A. Bateman, J.M. Goodfellow, N.H. Lubsen, C. Slingsby, The X-ray crystal structure of human γS-crystallin C-terminal domain, J. Biol. Chem. 277 (2002) 4199–4205.
[43] V. Sagar, S.K. Chaturvedi, P. Schuck, G. Wistow, Crystal structure of chicken γS-crystallin reveals lattice contacts with implications for function in the lens and the evolution of the βγ-crystallins, Structure 25 (2017) 1068-78.e2.
[44] J. Chen, P.R. Callis, J. King, Mechanism of the very efficient quenching of tryptophan fluorescence in human γD- and γS-crystallins: the γ-crystallin fold may have evolved to protect tryptophan residues from ultraviolet photodamage, Biochemistry 48 (2009) 3708–3716.
[45] Z. Ma, G. Piszczek, P.T. Wingfield, Y.V. Sergeev, J.F. Hejtmancik, The G18V CRYGS mutation associated with human cataracts increases γS-crystallin sensitivity to thermal and chemical stress, Biochemistry 48 (2009) 7334–7341.
[46] I.A. Mills, S.L. Flaugh, M.S. Kosinski-Collins, J.A. King, Folding and stability of the isolated Greek key domains of the long-lived human lens proteins γD-crystallin and γS-crystallin, Protein Sci. 16 (2007) 2427–2444.
[47] J.J. Harding, Disulphide cross-linked protein of high molecular weight in human cataractous lens, Exp. Eye Res. 17 (1973) 377–383.
[48] S. Najmudin, V. Nalini, H.P.C. Driessen, C. Slingsby, T.L. Blundell, D.S. Moss, P.F. Lindley, Structure of the bovine eye lens protein γB(γII)-crystallin at 1.47 Å, Acta Crystallogr. D 49 (1993) 223–233.
[49] R.J. Siezen, J.A. Thomson, E.D. Kaplan, G.B. Benedek, Human lens gamma-crystallins: isolation, identification, and characterization of the expressed gene products, Proc. Natl. Acad. Sci. U. S. A. 84 (1987) 6088–6092.
[50] V.M. Berthoud, E.C. Beyer, Oxidative stress, lens gap junctions, and cataracts, Antioxid. Redox Signal. 11 (2009) 339–353.
[51] M.H. Garner, A. Spector, Selective oxidation of cysteine and methionine in normal and senile cataractous lenses, Proc. Natl. Acad. Sci. U. S. A. 77 (1980) 1274–1277.
[52] J.N. Liang, M.R. Pelletier, Spectroscopic studies on the mixed disulfide formation of lens crystallin with glutathione, Exp. Eye Res. 45 (1987) 197–206.
[53] M.F. Lou, J.E. Dickerson, Protein-thiol mixed disulfides in human lens, Exp. Eye Res. 55 (1992) 889–896.
[54] C. Slingsby, L. Miller, The reaction of glutathione with the eye-lens protein γ-crystallin, Biochem. J. 230 (1985) 143–150.
[55] W.D. Brubaker, J.A. Freites, K.J. Golchert, R.A. Shapiro, V. Morikis, Douglas J. Tobias, R.W. Martin, Separating instability from aggregation propensity in γS-crystallin variants, Biophys. J. 100 (2011) 498–506.
[56] J.M. Hemmingsen, K.M. Gernert, J.S. Richardson, D.C. Richardson, The tyrosine corner: a feature of most Greek key β-barrel proteins, Protein Sci. 3 (1994) 1927–1937.
[57] S. Bagby, S. Go, S. Inouye, M. Ikura, A. Chakrabartty, Equilibrium folding intermediates of a Greek key β-barrel protein, J. Mol. Biol. 276 (1998) 669–681.
[58] S. Lee, B. Mahler, J. Toward, B. Jones, K. Wyatt, L. Dong, G. Wistow, Z. Wu, A single destabilizing mutation (F9S) promotes concerted unfolding of an entire globular domain in γS-crystallin, J. Mol. Biol. 399 (2010) 320–330.
[59] E. Serebryany, J.A. King, The βγ-crystallins: native state stability and pathways to aggregation, Prog. Biophys. Mol. Biol. 115 (2014) 32–41.
[60] M.J. Whitley, Z. Xi, J.C. Bartko, M.R. Jensen, M. Blackledge, A.M. Gronenborn, A combined NMR and SAXS analysis of the partially folded cataract-associated V75D γD-crystallin, Biophys. J. 112 (2017) 1135–1146.
[61] P. Das, J.A. King, R. Zhou, Aggregation of γ-crystallins associated with human cataracts via domain swapping at the C-terminal β-strands, Proc. Natl. Acad. Sci. U. S. A. 108 (2011) 10514–10519.
[62] J.C. Boatz, M.J. Whitley, M. Li, A.M. Gronenborn, P.C.A. van der Wel, Cataract-associated P23T γD-crystallin retains a native-like fold in amorphous-looking aggregates formed at physiological pH, Nat. Commun. 8 (2017), 15137.
[63] W. Kabsch, XDS, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 125–132.
[64] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J. Read, Phaser crystallographic software, J. Appl. Crystallogr. 40 (2007) 658–674.
[65] P.D. Adams, P.V. Afonine, G. Bunkóczi, V.B. Chen, I.W. Davis, N. Echols, J.J. Headd, L.-W. Hung, G.J. Kapral, R. W. Grosse-Kunstleve, PHENIX: a comprehensive Python-based system for macromolecular structure solution, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 213–221.
[66] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, Acta Crystallogr. D Biol. Crystallogr. 60 (2004) 2126–2132.
[67] P.V. Konarev, V.V. Volkov, A.V. Sokolova, M.H. Koch, D.I. Svergun, PRIMUS: a Windows PC-based system for small-angle scattering data analysis, J. Appl. Crystallogr. 36 (2003) 1277–1282.
[68] D. Svergun, Determination of the regularization parameter in indirect-transform methods using perceptual criteria, J. Appl. Crystallogr. 25 (1992) 495–503.
[69] D. Svergun, C. Barberato, M.H. Koch, CRYSOL—a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates, J. Appl. Crystallogr. 28 (1995) 768–773.
[70] D.I. Svergun, M.V. Petoukhov, M.H. Koch, Determination of domain structure of proteins from X-ray solution scattering, Biophys. J. 80 (2001) 2946–2953.
[71] V.V. Volkov, D.I. Svergun, Uniqueness of ab initio shape determination in small-angle scattering, J. Appl. Crystallogr. 36 (2003) 860–864.
[72] D.I. Svergun, Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing, Biophys. J. 76 (1999) 2879–2886.
[73] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera—a visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612.
[74] F. Sievers, A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, Fast,

scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, Mol. Syst. Biol. 7 (2011) 539.
[75] D. Cox, E. Selig, M.D. Griffin, J.A. Carver, H. Ecroyd, Small heat shock proteins prevent α-synuclein aggregation via transient interactions and their efficacy is affected by the rate of aggregation, J. Biol. Chem. 291 (2016) 22618–22629.

Supplementary Data

The structure and stability of the disulfide-linked γS-crystallin dimer provide insight into oxidation products linked with cataract

David C. Thorn, Aidan B. Grosas, Peter D. Mabbitt, Nicholas J. Ray, Colin J. Jackson, and John A. Carver

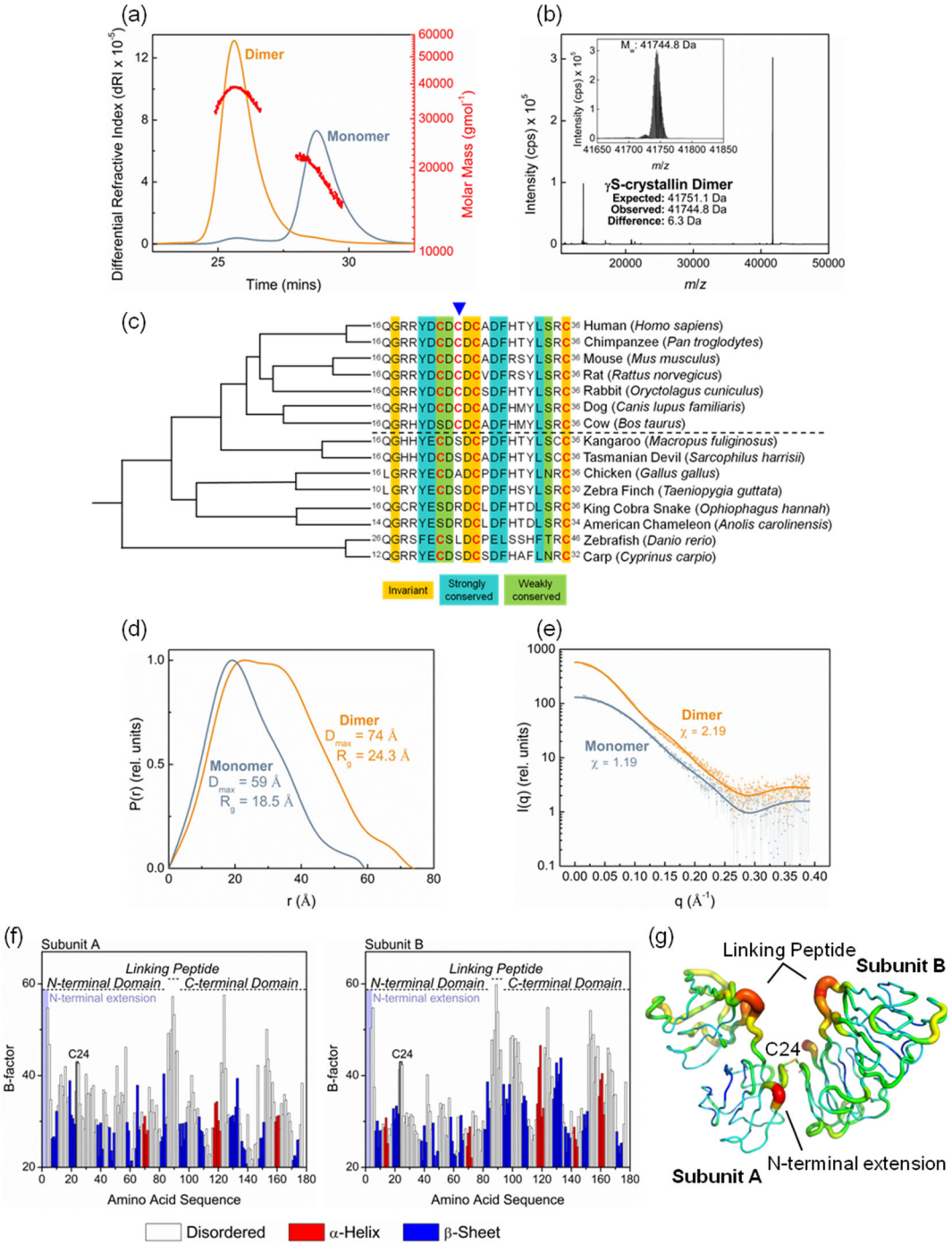


Fig. S1. Further structural characterization of the γS dimer. (a) SEC-MALS of purified monomeric (M_w: 19.5 kDa) and dimeric (M_w: 38.5 kDa) fractions of γS. (b) Intact mass spectrometry of the γS dimer with the *inset* showing the relevant mass peak and the weight average mass (M_w) for the isotopic distribution of the dimer. (c) Interspecies γS sequence alignment of the region centered around C24 (indicated with a blue triangle) and phylogenetic analysis using whole genome sequences. Species are indicated with both common and taxonomic, binomial names. Cysteines are colored red and in bold typeface. Sequence conservation is highlighted according to the key at the bottom of the figure. The horizontal dotted line indicates a differentiation between eutherians and other vertebrates. (d) P(r) function from SAXS data for γS monomer and dimer with the D_{max} and real-space R_g values generated by GNOM. (e) CRYSOL fits of the crystal structure (solid line) to the experimental SAXS data (scatter) for the γS monomer (light blue) (PDB: 2M3T) and dimer (orange) (6FD8). The χ value indicates the goodness of the fit. (f) B-factor plot for subunits A (*left*) and B (*right*) of γS dimer colored by secondary structure as noted in the key below. The N- and C-terminal domains, as well as the N-terminal extension, are indicated. C24 residue is highlighted. (g) γS dimer structure colored by B-factors. For residues with higher B-factors, the polypeptide backbone is represented as broader and colored nearer to red.

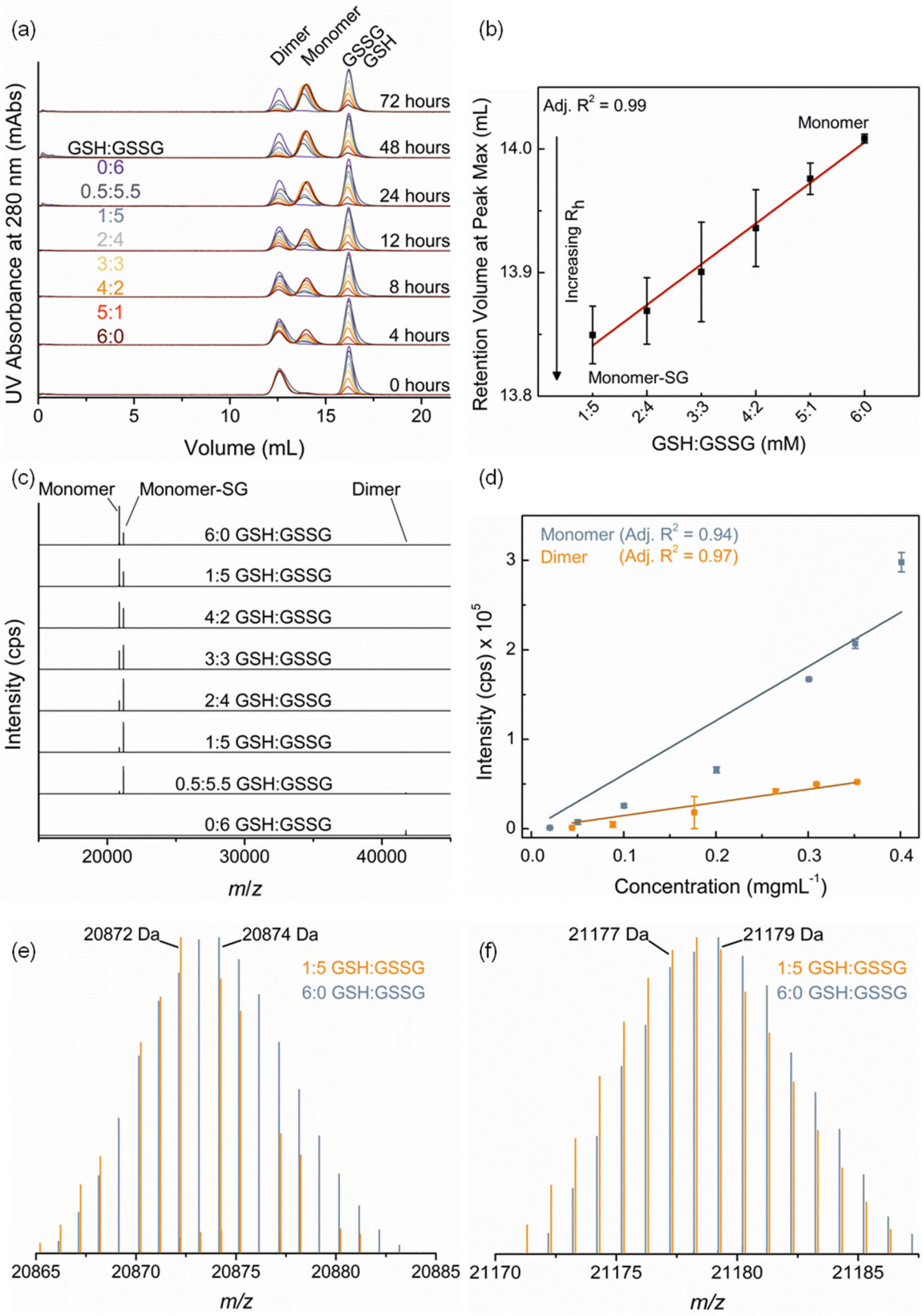

(a)
Dimer
Monomer
GSSG
GSH
72 hours
48 hours
24 hours
12 hours
8 hours
4 hours
0 hours
GSH:GSSG
0:6
0.5:5.5
1:5
2:4
3:3
4:2
5:1
6:0
UV Absorbance at 280 nm (mAbs)
Volume (mL)
(b)
Adj. R2 = 0.99
Monomer
Monomer-SG
Increasing Rh
Retention Volume at Peak Max (mL)
GSH:GSSG (mM)
(c)
Monomer
Monomer-SG
Dimer
6:0 GSH:GSSG
1:5 GSH:GSSG
4:2 GSH:GSSG
3:3 GSH:GSSG
2:4 GSH:GSSG
1:5 GSH:GSSG
0.5:5.5 GSH:GSSG
0:6 GSH:GSSG
Intensity (cps)
m/z
(d)
Monomer (Adj. R2 = 0.94)
Dimer (Adj. R2 = 0.97)
Intensity (cps) x 10^5
Concentration (mgmL^-1)
(e)
20872 Da
20874 Da
1:5 GSH:GSSG
6:0 GSH:GSSG
m/z
(f)
21177 Da
21179 Da
1:5 GSH:GSSG
6:0 GSH:GSSG
m/z

Fig. S2. Monitoring the resistance of the γS dimer to reduction using SEC and MS. (a) Chromatograms showing the separation of γS monomer and dimer as well as oxidized glutathione (GSSG) and reduced glutathione (GSH) in samples of γS dimer treated with increasing ratios of GSH to GSSG for various lengths of time. GSH:GSSG treatments of the dimer are colored as per the figure legend. (b) The retention volumes at peak maxima for the γS monomer from SEC experiments after 72 hours. A shift in retention volume indicates an increase in the hydrodynamic size of the monomer with increasing GSSG, indicative of protein glutathionylation. Error bars represent the standard deviation from two independent experiments. (c) Mass spectra of γS dimer as a function of increasing GSH:GSSG. Spectra are stacked along the *y*-axis for ease of comparison. (d) Standard curves for the intensity of intact mass spectra against either γS monomer or dimer concentration. To account for possible ionization exchange between the monomer and dimer, the standard curve samples were formulated using mixtures of different concentrations of both monomer and dimer. No exchange between these species was evident on the experimental timescale. Error bars represent the standard deviation from two independent experiments. (e) Isotopic peak distributions of the γS monomer produced when dimer is treated with the low (1:5) or high (6:0) GSH:GSSG. Maximum intensity of the distribution shows a shift from 20,872 Da to 20,874 Da with increased GSH concentration, implying the loss of the intramolecular disulfide bond in the monomer at higher reduction potential. The respective sample intensities were normalized and the masses offset by 0.1 Da for comparison. (f) As in *E*, except here the *S*-glutathionylated monomer species. The mass of 21,179 Da for 6:0 GSH:GSSG is indicative of *S*-glutathionylation only while a mass of 21,177 Da for 1:5 GSH:GSSG is indicative of *S*-glutathionylation with an additional disulfide bond.

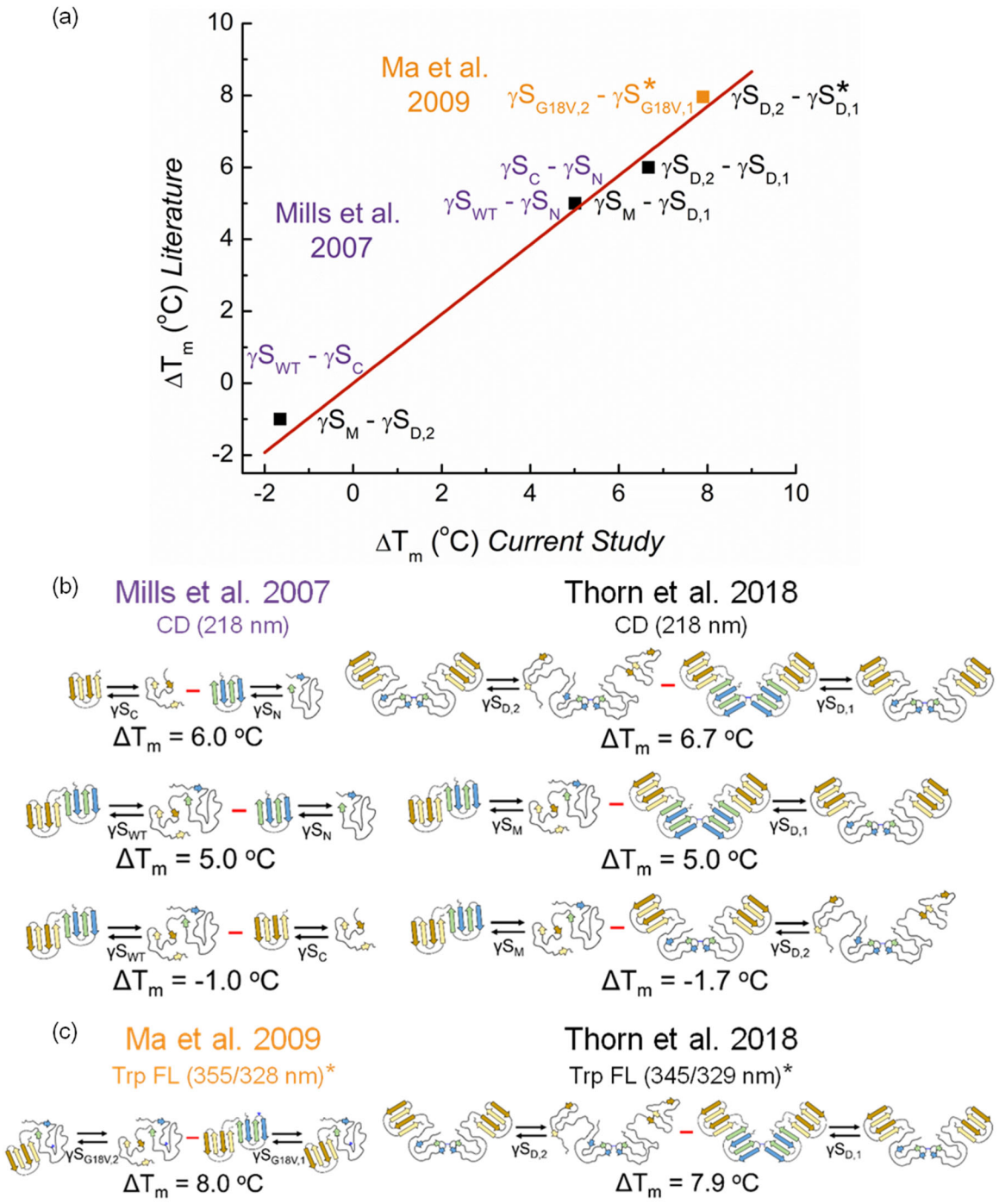


Fig. S3. Correlation of the thermal unfolding transitions of the γS monomer and dimer with those of the N- and C-terminal domains. (a) A correlation plot between ΔT_m determined in this study on wildtype γS (Thorn et al. 2018, black, *x*-axis) and those reported in previous studies on the isolated N- and C-terminal domains (Mills et al. 2007, purple, *y*-axis) (17) as well as the biphasic G18V variant (Ma et al. 2009, orange, *y*-axis) (18). Abbreviations: wildtype γS (γS_{WT}), γS N-terminal domain only (γS_N), γS C-terminal domain only (γS), monomeric γS (γS_M), dimeric γS first transition ($\gamma S_{D,1}$), dimeric γS second transition ($\gamma S_{D,2}$), G18V γS first

transition ($\gamma S_{G18V,1}$), and G18V γS second transition ($\gamma S_{G18V,2}$). Asterisk indicates a ΔT_m derived from Trp fluorescence. (b) Schematic representation of γS unfolding and their respective ΔT_m values derived from thermal unfolding monitored via CD. The red line indicates the subtraction of one T_m from another to calculate ΔT_m. (c) As in (b) except here comparing ΔT_m values from Trp fluorescence. The valine side-chain in G18V γS is represented as a blue Y-shaped protrusion from the polypeptide backbone.

Table S1. Thermal unfolding mid-point (T_m) and difference (ΔT_m) for γS monomer and dimer.

		T_m (°C)			
Oligomeric State	Transition	Trp FL (345/329 nm)	LS (360 nm)	ANS FL (480 nm)	CD (218 nm)
Monomer	Single (M)	67.9 ± 0.1	68.5 ± 0.2	71.0 ± 0.2	72.9 ± 0.1
Dimer	First (D,1)	63.9 ± 0.0	65.4 ± 0.1	66.5 ± 0.1	67.9 ± 0.2
	Second (D,2)	71.8 ± 0.2	72.4 ± 0.1	72.7 ± 0.4	74.6 ± 0.1
		ΔT_m			
	D,2 – D,1	7.9	7.0	6.2	6.7
	M – D,1	3.9	3.1	4.5	5.0
	M – D,2	-4.0	-3.8	-1.7	-1.7

Errors for T_m values are given as the standard deviation of three independent repeats.

CHAPTER 3:

MACROMOLECULAR CROWDING AND αB-CRYSTALLIN

"Above all, don't fear difficult moments. The best comes from them."

– Rita Levi-Montalcini

DECLARATION

The following manuscript is formatted in the style of the *Journal of Biological Chemistry*.

All experimental work was carried out by the author, except where otherwise stated below, under the supervision of Professor John Carver.

The contributions of other authors are as follows: The author and Professor John Carver conceived the study. Dr Agata Rekas purified the deuterated form of the protein and assisted with sample preparation for small angle neutron scattering. Dr Jitendra Mata acquired small angle neutron scattering data and completed data reduction. Dr David Thorn prepared imaging grids and acquired transmission electron microscopy data. The author wrote the manuscript with input from the co-authors.

The aggregation of αB-crystallin under crowding conditions is prevented by αA-crystallin: Implications for α-crystallin stability and lens transparency

Aidan B. Grosas[‡], Agata Rekas[§], Jitendra P. Mata[¶], David C. Thorn[‡] and John A. Carver[‡,1]

[‡]Research School of Chemistry, The Australian National University, Acton, ACT 2601, Australia, [§]National Deuteration Facility, Australian Nuclear Science and Technology Organisation (ANSTO), Lucas Heights, NSW 2234, Australia, and [¶]ACNS, Australian Nuclear Science and Technology Organisation (ANSTO), Lucas Heights, NSW 2234, Australia

[1]To whom correspondence may be addressed: Research School of Chemistry, Building 137, Sullivans Creek Rd, Australian National University, Acton ACT 2601, Australia.

Email: john.carver@anu.edu.au.

The authors declare no conflict of interest.

Abbreviations: αAc, αA-crystallin; αBc, αB-crystallin; ACD, alpha-crystallin domain; CD, circular dichroism; CTE C-terminal extension; CTR, C-terminal region; D_{max}, maximum dimension; DTT, dithiothreitol; NTR, N-terminal region; P(r), pair-wise distance distribution function; PTMs, post-translational modifications; R_g, radius of gyration; SANS, small angle neutron scattering; sHsp, small heat-shock protein; T_{agg}, mid-point of aggregation; TEM, transmission electron microscopy; ThT, thioflavin T; T_m, mid-point of unfolding; Trp, tryptophan; λ_{bcm}, barycentric mean fluorescence

Keywords: αB-crystallin, αA-crystallin, aggregation, lens, cataract, macromolecular crowding, small-angle scattering, fibril formation, molecular chaperone, structure

Abstract

One of the most crowded biological environments is the eye lens which contains a high concentration of crystallin proteins. The molecular chaperones αB-crystallin (αBc) with its lens partner αA-crystallin (αAc) prevent deleterious protein aggregation and cataract formation. However, some forms of cataract are associated with structural alteration and dysfunction of αBc. While many studies have investigated the structure and function of αBc under dilute *in vitro* conditions, the effect of crowding on these aspects is not well understood despite its *in vivo* relevance. The structure and chaperone ability of αBc under conditions that mimic the crowded lens environment were investigated using Ficoll 400 and bovine γ-crystallin as crowding agents and a variety of biophysical methods, principally contrast variation small-angle neutron scattering. Under crowding conditions, αBc unfolds, increases in size/oligomeric state, decreases in thermal stability and chaperone ability, and forms kinetically distinct amorphous and fibrillar aggregates. However, the presence of αAc stabilizes αBc against aggregation. These results provide a molecular rationale for the aggregation of αBc in the crowded lens, which exhibits marked structural and functional similarities to cataract-associated mutants R120G and D109A αBc under dilute conditions. Strategies that maintain or restore αBc stability, as αAc natively does, might therefore provide an avenue for the therapeutic treatment of cataract.

Introduction

There are three types of mammalian lens crystallins: α-crystallins, which are members of the small heat-shock protein (sHsp) family and the β- and γ-crystallins, which together are part of a structurally homologous superfamily containing Greek key β-sheet motifs (1-3). Within the lens, crystallins have a structural role in maintaining a short-range array that facilitates a smooth refractive index gradient for correct light refraction and lifelong lens transparency (4-6). Crystallins are long-lived, as necessitated by the lack of protein turnover in the lens fibre cells (7,8). However, they are subject to structural perturbations that can cause their partial unfolding and aggregation leading to lens opacification, i.e. the scattering of incident light, and, ultimately, cataract (3,9). Deleterious structural alterations can manifest either with age, via environmental insults and/or the accumulation of post-translational modifications (PTMs) (10), or are apparent from birth/early age via mutations (11). The α-crystallins have an additional important function as molecular chaperone proteins to curtail lens protein unfolding and aggregation (12).

Of the crystallin proteins in the human lens, ~40 % are α-crystallins (13,14) which comprise two subunits, αB-crystallin (αBc) and αA-crystallin (αAc), that share ~60 % sequence identity (3). On average, the αBc:αAc ratio in the young mammalian lens is 1:3 (14) and this changes to a 2:3 ratio by ~55 years of age in the human lens (15), i.e. the proportion of αBc increases. While αAc is predominantly a lenticular protein (16), αBc is ubiquitous, where it acts as a molecular chaperone to minimise the aggregation of extra-lenticular proteins, for example, those associated with neurodengerative diseases (17) and degenerative muscular diseases (18,19).

The eye lens is a highly crowded environment containing ~90 % crystallin protein at 300 - 400 mg/mL (3,9). Such *in vivo* conditions contrast starkly with the dilute *in vitro* conditions under which most crystallin studies have been conducted. The crowded environment *in vivo* means that the crystallins are susceptible to entropically driven excluded volume effects (20). Under these conditions, the folding, conformation, oligomerization, and chaperone action of the major lens proteins αBc and αAc can be significantly affected (21). Between room and physiological temperature, both αBc and αAc are highly soluble and stable under dilute *in vitro* conditions (3,22). However, *in vivo* studies highlight a critical role for αAc in stabilizing αBc in the crowded environment of the lens. An αAc knockout

mouse model developed early-onset cataract, the aggregates of which were rich in αBc (23). Similarly, a homozygous recessive W9X nonsense mutation of the αAc gene in two human families, which essentially abolished functional αAc in the lens, also led to early-onset cataract (24). Conversely, αBc knockout mice were relatively normal in their lens development (25).

Under dilute *in vitro* conditions, αBc is a polydisperse oligomer (26) with a 24-mer being one of its most populated forms (Fig. 1, *A, top*) (27). Dynamic subunit exchange between oligomers is a key feature of αBc which allows for the highly responsive shifts in oligomeric distribution needed for molecular chaperone function and hetero-oligomerization with related sHsps, e.g. αAc (28-30). There are three distinct regions in αBc of which the N-terminal region (NTR) and C-terminal region (CTR) (incorporating the C-terminal extension (CTE)) are largely disordered and flank the central conserved α-crystallin domain (ACD) which adopts an immunoglobulin-like, β-sheet-rich fold (Fig. 1, *A, bottom*) (31,32). Compared to αBc, αAc exhibits many similar structural and functional features with some notable differences: a lower isoelectric point, diminished hydrophobic surface area, reduced NTR solvent exposure, and generally poorer *in vitro* chaperone activity (33). The small number of *in vitro* based crowding experiments on the α-crystallins using non-proteinaceous crowding agents show that αBc size and molar mass increase (34) and αAc subunit exchange rate is slowed compared to the dilute environment (35). Additionally, the chaperone activities of human αBc, αAc, and bovine α-crystallin are all diminished (35,36).

Previous *in vitro* crowding studies on αBc have yet to account for the lens opacification and cataract formation observed in αAc knockout mice (23) and in humans carrying the W9X αAc mutation (24). A detailed examination of αBc structure, stability and function in crowded conditions representative of the lens could offer mechanistic insight into how αBc can form aggregates that contribute toward lens opacification and, further, how interaction with its lens partner, αAc, functions to maintain a normal lens phenotype (37). Herein, we use a polysaccharide, Ficoll 400, and a protein, bovine γ-crystallin, as crowding agents to mimic the high concentrations inherent to the lens in order to study the physical, structural and functional changes of αBc relative to a dilute environment. Under crowded conditions, we show that αBc is destabilized, unfolds, increases in size, loses chaperone activity, and forms kinetically distinct amorphous and fibrillar aggregates. *In vivo*, such behaviour

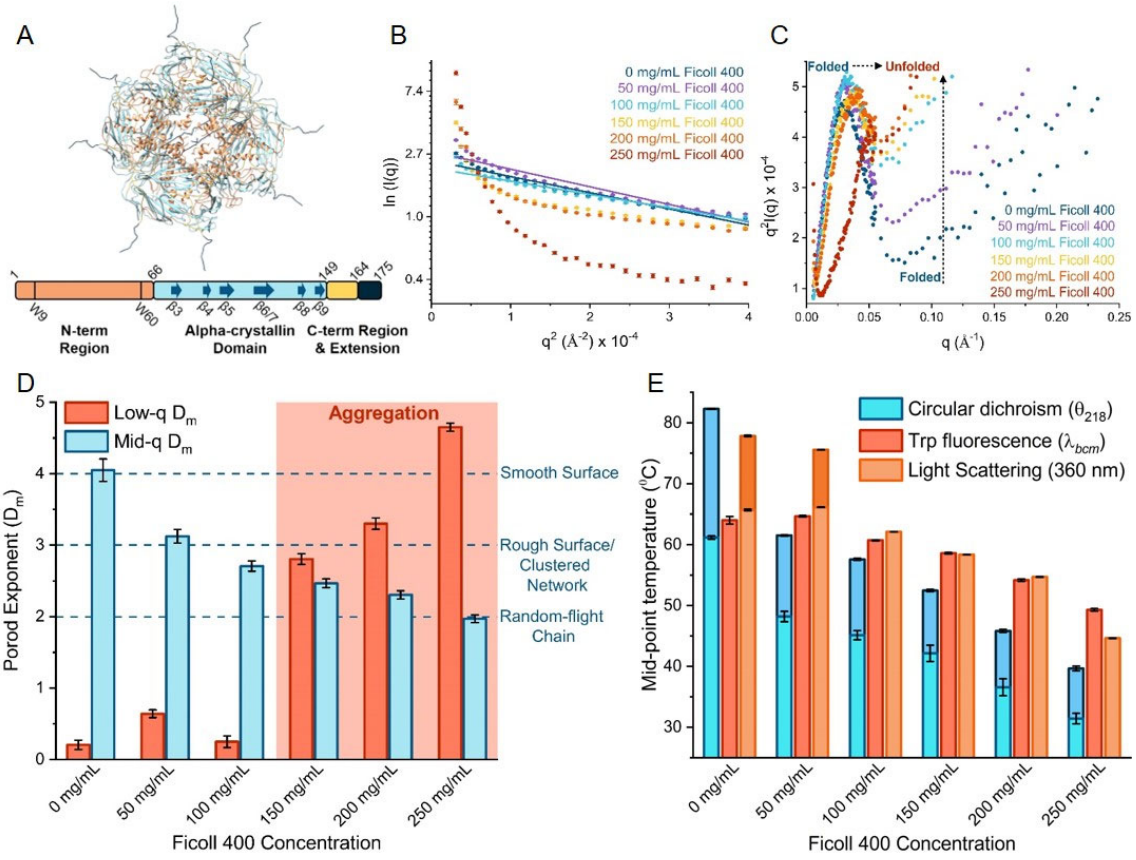


Figure 1. Structural changes and the thermal stability of αBc under conditions of increasing macromolecular crowding. *A*, atomic-level model of αBc 24mer (PDB: 3J07) (27) (*top*) and a linearized schematic of the αBc monomer (38) (*bottom*) both coloured according to specific regions i.e. NTR (1-65, *red*), the ACD (66-148, *blue*), the CTR (149-163, *yellow*) and the CTE (164-175, *navy blue*). The two tryptophan residues (W9 and W60, *dark red line*) in the NTR and β-strands (β3-β9, *dark blue arrows*) (39) within the ACD are noted in the linearized schematic (*bottom*). *B*, Guinier and *C*, Kratky plot of 2.5 mg/mL deuterated αBc in Ficoll 400 concentrations from 0-250 mg/mL with 50 mg/mL increments at 37 °C in 28% D_2O. *D*, correlation of Porod exponent (mid-q D_m, blue (0.05 $Å^{-1} < q < 0.40$ $Å^{-1}$)) with the gradient of the Guinier region (low-q D_m, red (0.006 $Å^{-1} < q < 0.018$ $Å^{-1}$)) against the aforementioned crowding conditions in *B* and *C*. Errors are given as the standard error of the fit. *E*, thermal stability of 0.3 mg/mL αBc in Ficoll 400 concentrations from 0-250 mg/mL with 50 mg/mL increments as measured using 218 nm CD (*blue*), barycentric mean fluorescence (λ_{bcm}) Trp fluorescence (*red*) and 360 nm light scattering (*orange*). Both 218 nm CD (0 – 250 mg/mL Ficoll 400) and 360 nm light scattering (0 and 50 mg/mL Ficoll 400) have two transitions associated with thermal denaturation of αBc. A light colour shade is applied to the first transition while a dark colour shade is applied to the second. Errors are given as the standard error of a sigmoidal fit to the thermal denaturation data.

would contribute directly to lens opacification, if not for the concomitant chaperone action of αAc that significantly increases the stability of αBc under crowded conditions and prevents aggregation. We

discuss how age-related modifications or mutations *in vivo* might contribute to cataract formation through either destabilizing αBc and/or inactivating αAc, and how therapeutic approaches to cataract prevention could be efficacious in such instances.

Results

Macromolecular crowding induces unfolding and thermal instability in αBc that leads to aggregation

Contrast variation small-angle neutron scattering (SANS) provides a proven platform to investigate the effects of macromolecular crowding on protein structure (40,41). The neutron scattering of a highly concentrated hydrogenated crowding agent is attenuated using a molecule-dependent volume percentage of D_2O, enabling the scattering of only the partner deuterated molecule (in this case, a protein) to be observed. In this study, hydrogenated Ficoll 400 was used as a model-crowding agent as its average mass is similar to that of the αBc oligomer. Deuterated αBc was mixed with increasing concentrations of hydrogenated Ficoll 400 whose scattering was matched out with 28 % D_2O enabling the neutron scattering of only deuterated αBc to be observed.

Quantitative estimation of a protein's radius of gyration (R_g) from a Guinier plot (ln (I(q)) vs q^2 where q the scattering vector) is contingent on the scattering from the low-q region maintaining linearity. Qualitative examination of the low-q region of a Guinier plot from αBc scattering in 0 – 100 mg/mL of Ficoll 400 remained linear indicating a stable size (Fig. 1*B*). However, Ficoll 400 concentrations from 150 – 250 mg/mL caused a significant upturn indicating protein aggregation. Interestingly, we noted the appearance of a Bragg peak in the SANS data with increasing temperature and time, particularly at 250 mg/mL Ficoll 400 (Fig. S1, *A*). Given Bragg peaks usually arise due to scattering from a defined repeating structure, this could indicate the presence of an ordered αBc aggregate (42). While protein aggregation from αBc precluded the determination of a R_g value across the entire Ficoll 400 concentration series, the influence of aggregation on the scattering intensity outside of this low-q region is significantly diminished (40). As such, we analyzed these data using a Kratky plot ($q^2I(q)$ vs. q) which provides a qualitative measure of the protein's compactness and globularity i.e. its degree of foldedness (43). The Kratky plot for αBc at 0 mg/mL Ficoll 400 shows a trace with a prominent peak that trends downward towards the q-axis at mid-q values but stops just short of

convergence with the x-axis (Fig. 1, *C*). This indicates a largely folded globular protein with some unfolded regions, i.e. the central ACD and terminal regions, respectively (Fig. 1 *A*, *top*), and is consistent with a previous SAXS finding on related sHsps (44). As macromolecular crowding increases with Ficoll 400 concentrations from 50 to 200 mg/mL, the peak shifts to a higher-q value and there is an upturn in the mid-q region indicating an increasingly unfolded polypeptide. Finally, the peak at 250 mg/mL Ficoll 400 is lost and a major upturn in the mid-q region is evident indicating that αBc is largely unfolded (Fig. 1, *C*).

The foldedness of αBc can also be measured using the Porod exponent (D_m) which is determined over a mid-q range ($0.05\ Å^{-1} < q < 0.40\ Å^{-1}$) using a power-law relationship (Eqn. 1, see *Experimental Procedures*) (Fig. S1, *B*). The D_m value reflects the internal scaling of interatomic distances over different length scales and can be interpreted as a fractal dimension, sensitive to the degree of compaction of the protein chain (45,46). A D_m value of 4 indicates a surface fractal which is smooth and globular-like, 3 is either a rough surface or a mass fractal equivalent to a clustered network of chains, and 2 is indicative of a random flight chain (47,48). The D_m value of αBc at 0 mg/mL Ficoll 400 begins at 4.0 ± 0.2, indicating an unperturbed globular structure, and decreases sequentially with increasing Ficoll 400 concentration, indicating a loss of structure with increased crowding. At 250 mg/mL Ficoll 400, the D_m value is 2.0 ± 0.1 indicating a random flight chain, i.e. a largely unfolded protein (Fig. 1, *D*, *blue*), which is consistent with the Kratky plot. We also applied Eqn. 1 over the low-q range ($0.006\ Å^{-1} < q < 0.018\ Å^{-1}$), where determination of the R_g value from the Guinier plot over the experimental series of 0 – 250 mg/mL Ficoll 400 no longer satisfied the condition of $qR_g \leq 1.3$ due to protein aggregation, giving rise to a low-q upturn (Fig. S1, *B*). This analytical method provided a unitless measure of the gradient of that upturn with a higher gradient value indicating greater aggregation. Similar to the Guinier analysis (Fig. 1, *B*), this value showed increasing aggregation from 150 mg/mL upwards in concentration of Ficoll 400 (Fig. 1, *D*, *red*). Together, these values provide a representation of the interplay between the effects of macromolecular crowding on αBc foldedness and general structure using the mid-q D_m (Porod exponent) and on aggregation based on the low-q D_m. In summary, aggregation of αBc occurs between 100 and 150 mg/mL Ficoll 400 where αBc undergoes

significant but not complete loss of structure. Above 150 mg/mL, αBc continues to unfold and aggregate as crowding increases (Fig. 1, *D*).

To investigate the impact of macromolecular crowding on the stability of the αBc native state, the thermal stability was measured using either the mid-point of unfolding (T_m) or aggregation (T_{agg}) at increasing concentrations of Ficoll 400. As αBc is heterogeneous in structure, different spectroscopic methods provided information on its various structural regions. Circular dichroism (CD) ellipticity at 218 nm, representative of β–sheet secondary structure, provided information on structural changes associated with the β–sheet-rich ACD (Fig. 1, *A, bottom (blue arrows*)). Tryptophan (Trp) fluorescence provided information on the local environment of the NTR via W9 and W60 (Fig. 1, *A, bottom (dark red lines*)) and was analysed via λ_{bcm} (Eqn. 2) (Fig. S1, *C*). Finally, light scattering at 360 nm provided information on the overall size increase of αBc at the onset of aggregation. Thermal denaturation of αBc, monitored by CD, showed two transitions: the early decrease in ellipticity at 218 nm (lower T_m) was interpreted as a conformational change in the ACD (33), while the late increase (higher T_m) is indicative of denaturation (Fig. S1, *D*). Light scattering displayed a biphasic curve for Ficoll 400 concentrations of 0 and 50 mg/mL, consistent with a previous study under similar dilute conditions (49), while all other aggregation curves were monophasic (Fig. S1, *E*). In the absence of Ficoll 400, the first transition T_m values for CD (61.1 ± 0.3 °C), Trp fluorescence (63.0 ± 0.6 °C) and T_{agg} for light scattering (65.7 ± 0.1 °C) (Fig. 1, *E*), were consistent with literature values (3,49,50). The second CD transition at 82.3 ± 0.1 °C has not been previously reported in the literature to our knowledge. The effects of macromolecular crowding on αBc thermal stability are most evident when comparing results in the presence of 0 and 250 mg/mL Ficoll 400, with differences in T_m (ΔT_m) of 29.7 °C (β–sheet secondary structure change) and 42.6 °C (denaturation) for CD, 14.7 °C for Trp fluorescence and a difference in T_{agg} (ΔT_{agg}) of 21.1 °C for light scattering. The CD T_m values are the most drastically affected indicating that the ACD is the most susceptible structural motif to crowding, in line with the observation above that unfolding precedes αBc aggregation under crowded conditions. Overall, for all methods of analysis, the mid-point temperature values decrease with increasing Ficoll 400 concentration, indicating that macromolecular crowding significantly destabilizes αBc.

αBc undergoes kinetically distinct and concentration-dependent amorphous and fibrillar aggregation in a highly crowded environment

To investigate further the nature of the crowding-induced aggregation of αBc and to decouple kinetic and temperature effects inherent to the SANS studies, a crowded solution of 250 mg/mL Ficoll 400 at 37 °C and pH 7.3 was used to approximate the environment of the eye lens. The aggregation kinetics of 2.5 mg/mL αBc was monitored using light scattering at 360 nm (which is sensitive to both amorphous and fibrillar aggregation) and thioflavin T (ThT) fluorescence (which is sensitive to fibrillar aggregation); the latter was measured to inquire into the origin of the putative Bragg peak in the SANS data (51). Light scattering exhibited a strong sigmoidal response with a mid-point of ~0.5 hours and a maximum intensity at ~5 hours after which time a steady decrease in light scattering occurred due to protein precipitation (Fig. 2, *A*, *blue*). ThT fluorescence was initially high and decreased until ~10 hours (*see later*) after which the fluorescence increased in an exponential manner with a mid-point of ~16 hours and a final fluorescence maximum at 33 hours (Fig. 2, *A*, *red*).

Negative stain transmission electron microscopy (TEM) was employed to understand better the aggregation processes that were monitored by light scattering and ThT fluorescence. As an initial control, the micrograph of αBc at 0 hours under non-crowded conditions displays individually dispersed, spherical oligomers of ~15 nm in diameter consistent with previous studies (52,53) (Fig. 2, *B*). In contrast, at 0 hours under crowded conditions (in the presence of 250 mg/mL Ficoll 400) individual αBc oligomers were not easily discernible, implying that volume exclusion effects forced the oligomers closer together (Fig. 2, *C*). TEM at specific time-points of the kinetic experiment (Fig. 2, *A, black dashed lines* labelled *D – G*) provided a snapshot of the size and morphology of the αBc aggregates in crowded conditions. At 0.5 hours of incubation, spheroidal aggregates of approximately 40-50 nm were observed (Fig. 2, *D*, *black triangles*). After 5 hours, amorphous aggregates ranging from approximately 40-160 nm in diameter were observed (Fig. 2, *E*, *black* and *white triangles*, respectively), which grew to be very large over a further 11 hours of incubation (i.e. by 16 hours), however no fibrillar species were discernible, potentially due to their small size (Fig. 2, *F*). Finally, samples at 33 hours contained fibrillar species of various lengths and morphologies (Fig. 2, *G*, *black triangles*). In summary, these experiments confirmed that the light scattering kinetics represents fast-forming amorphous

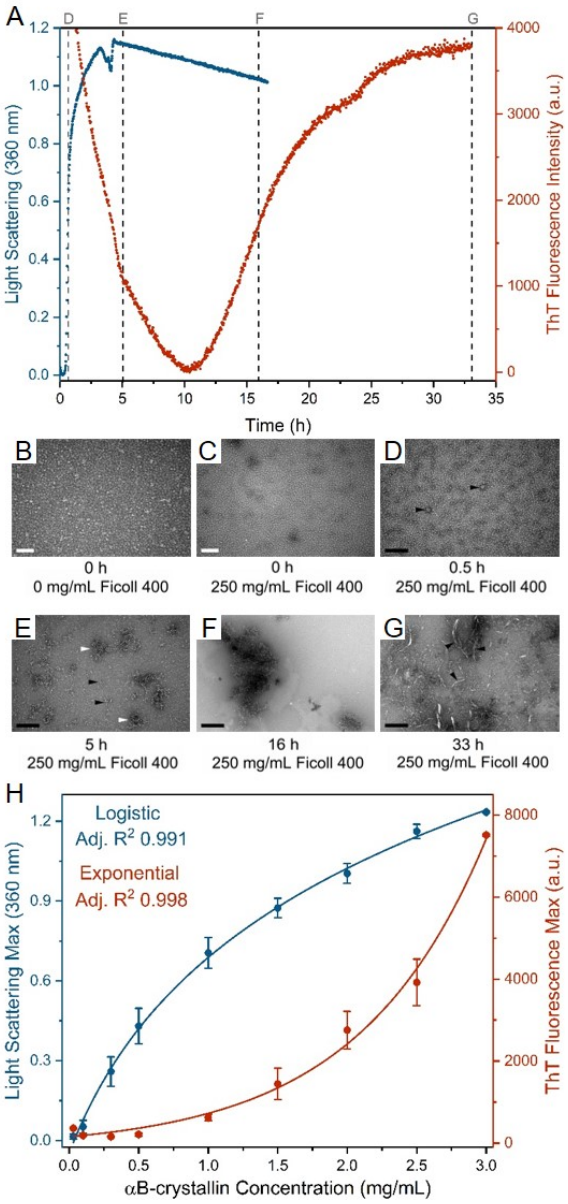


Figure 2. Characterizing the aggregation of αBc under conditions of macromolecular crowding. *A*, representative aggregation assay assessing the kinetics of light scattering at 360 nm (*left axis*, *blue*) and ThT fluorescence (*right axis*, *red*) for 2.5 mg/mL αBc in 250 mg/mL Ficoll 400 at 37 °C. Black dashed lines at 0.5, 5, 16, and 33 h, labelled D, E, F, and G, respectively, relate to the corresponding TEM micrographs. *B-G*, representative TEM micrographs of 2.5 mg/mL αBc in a dilute environment compared with the crowded environment sampled at specific time points during the kinetic experiment. Below each micrograph is listed the corresponding sampling time and Ficoll 400 concentration. White and black scale bars represent 100 nm and 200 nm, respectively. Black and/or white arrows are used to highlight noteworthy aggregates. *H*, light scattering maximum at 360 nm (*left axis*, *blue*) and ThT fluorescence maximum (*right axis*, *red*) for αBc from 0.03-3.00 mg/mL in 250 mg/mL Ficoll 400 at 37 °C. Errors are given as the standard deviation of three independent repeats. Light scattering maxima are fitted to a logistic function with an adjusted R^2 of 0.991 (*solid line*, *blue*) while the ThT fluorescence maxima are fitted to an exponential function with an adjusted R^2 of 0.998 (*solid line*, *red*).

aggregation of αBc while the ThT fluorescence kinetics represents the formation of slow-forming fibrillar species (Fig. 2, *A*).

A plot of the maximum light scattering and ThT fluorescence against αBc concentration revealed differences in the thermodynamics between the formation of the amorphous and fibrillar αBc species in crowded conditions of 250 mg/mL Ficoll 400 (Fig. 2, *H* and Fig. S2, *A* and *B*). Light scattering exhibited a logistic trend while final ThT fluorescence values showed an exponential trend. Notably, initial ThT fluorescence values over the αBc concentration range display a Gaussian trend with maximum at 1.5 mg/mL αBc (Fig. S2, *C*). This suggests that the source of the initial high ThT fluorescence and its subsequent decrease over approximately eight hours (Fig. S2, *B*) reflect a ThT-binding structure of αBc that becomes increasingly exposed, and is then lost, under highly crowded conditions. Overall, these data indicate that the formation of amorphous aggregations in highly crowded conditions is favoured at low αBc concentrations but is limited at higher concentrations (≥ 1.5 mg/mL). Conversely, ThT fluorescence data show that the formation of fibrillar species is disfavoured at low αBc concentrations but increasingly favoured at higher concentrations. In conjunction with the kinetic data, a theoretical reaction coordinate diagram depicting amorphous aggregation as kinetically stable and fibrillar aggregation as thermodynamically stable, either as separate pathways or as a single pathway with amorphous aggregation as the reaction intermediate, can be used to rationalize these results (Fig. S2, *D*).

The effects of macromolecular crowding on αBc structure, stability and aggregation propensity are accentuated by bovine γ-crystallin relative to Ficoll 400

The use of polysaccharides such as Ficoll 400 is a mainstay of *in vitro* crowding studies, however their effects may not adequately mimic the biological environment (54). To verify the relevance of Ficoll 400 as a model for the crowded environment of the eye lens, we used γ-crystallin extracted from bovine lens homogenate as a biologically relevant crowding agent (Fig. S3, *A*). Given the potential of protein crowding agents to obfuscate biophysical studies of protein targets, contrast matched SANS was again employed. The scattering of 125 mg/mL bovine γ-crystallin was matched out by 40 % D_2O and its effect

on deuterated αBc structure, stability and aggregation propensity was compared with 125 mg/mL Ficoll 400 and no crowding (dilute environment) in buffer containing 28 % D_2O.

Guinier analysis and the pair-wise distance distribution function (P(r)) were used to determine the initial Guinier and real-space R_g values, respectively, as well as the P(r) derived maximum dimension (D_{max}), for αBc at 30 °C (Fig. 3, *A* and *B*, respectively). The Guinier (52.1 ± 0.5 Å and 53.5 ± 0.8 Å) and real-space (52.2 Å and 53.3 Å) R_g values were similar for conditions of no crowding and crowding using 125 mg/mL Ficoll 400, respectively, and were consistent with the previously reported R_g value for αBc (39,55). The D_{max} of 190 Å for αBc under conditions of no crowding precisely matched that previously reported (39), while crowding using 125 mg/mL Ficoll 400 gave a D_{max} of 178 Å. As the R_g values are similar and the shape of the P(r) function matches well until the highest r values (Fig. 3, *B*), it is interesting that the P(r) function for αBc in 125 mg/mL Ficoll 400 abruptly drops, thereby producing a smaller D_{max}. The likely explanation for this behaviour is mild compaction of the CTE (39,56), while the general oligomeric size and shape are maintained. In contrast, crowding with 125 mg/mL γ-crystallin showed a ~28 Å increase in R_g value accompanied by an 86 Å increase in D_{max}, relative to no crowding (Fig. 3, *B*), indicating a significant alteration in the overall structure of αBc (perhaps due to unfolding) or a shift in the size distribution of the oligomeric array. The general shape of the P(r) plot indicates that under conditions of no crowding and 125 mg/mL Ficoll 400, αBc is largely globular with extended regions as discerned by the predominant bell-shaped peak and a tailing shoulder at higher r values (Fig. 3, *B*, *black* and *orange*). However, crowding using 125 mg/mL γ-crystallin leads to an increase in the bell-shaped peak maximum r value consistent with an increase in overall size and an extensive multi-modal shoulder suggesting a more extended structure and/or new extended populations (Fig. 3, *B*, *blue*).

A normalized Kratky plot ($qRg^2 \cdot I(q)/I(0)$ vs. qR_g) (57) was used to determine relative changes in compactness and flexibility (i.e. 'foldedness') for αBc under no crowding and crowding conditions at 30 °C. In this plot, a globular folded protein exhibits a bell-shaped peak maximum value of 1.104 at a qR_g of √3 (Fig. 3, *C*, *dashed lines*) while any departure from this intersection point indicates increased flexibility and hence unfolding (58). Under conditions of no crowding and crowding using 125 mg/mL of Ficoll 400, αBc has a normalized Kratky plot indicative of a partially folded protein with the latter

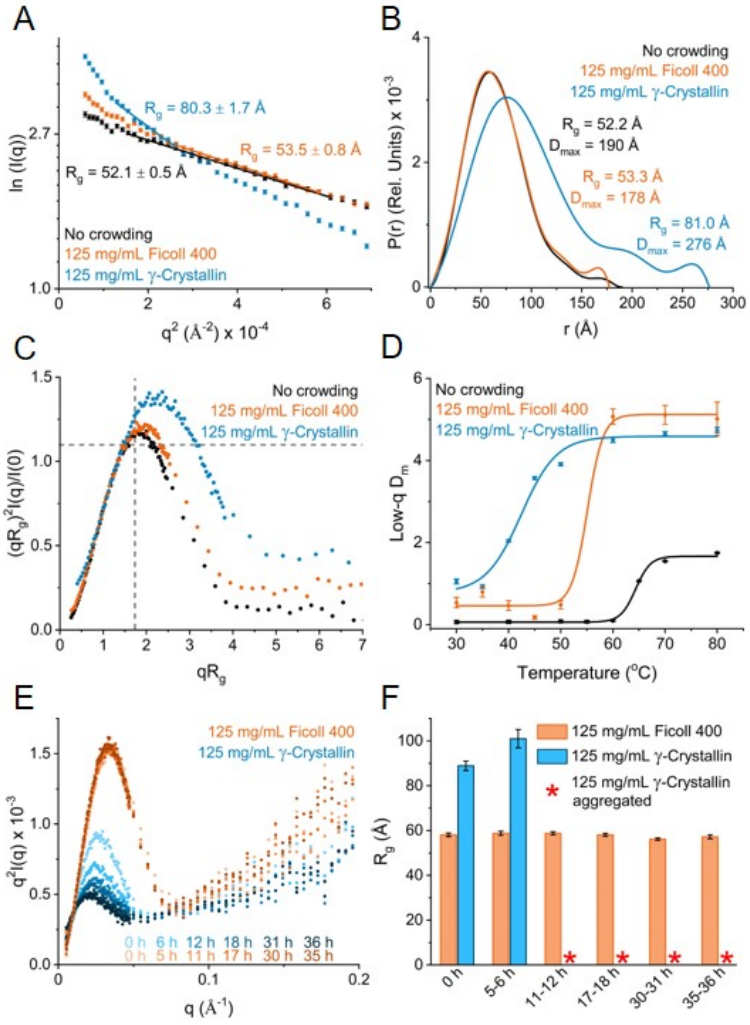


Figure 3. Effects of macromolecular crowding on αBc using bovine γ-crystallin compared to Ficoll 400 and the dilute environment via contrast variation SANS. *A*, Guinier plot, *B*, P(r) function and, *C*, dimensionless Kratky plot for deuterated αBc in conditions of 'No crowding' (*black*), '125 mg/mL Ficoll 400' (*orange*), and '125 mg/mL γ-crystallin' (*blue*) at 30 °C. Parameters derived from these analyses are given on the plot in the colour corresponding to the crowding condition. *D*, thermally induced aggregation curves in the aforementioned crowding conditions followed as a function of the low-q D_m (0.006 $Å^{-1} < q < 0.018$ $Å^{-1}$) or otherwise the gradient of the Guinier region which is sensitive to sample aggregation (*squares*). The data are fitted with a Boltzmann function (*solid line*). *E*, Kratky plot of αBc in 125 mg/mL Ficoll 400 (*orange, light to dark with time*), and 125 mg/mL γ-crystallin (*blue*, *light to dark with time*) at 37 °C over a period of 36 hours. *F*, R_g value of αBc in 125 mg/mL Ficoll 400 (*orange*), and 125 mg/mL γ-crystallin (*blue*) at 37 °C over a period of 36 hours. A *red asterisk* for αBc in 125 mg/mL γ-crystallin indicates the sample was aggregated, i.e. an upturn in the scattering data of the Guinier region no longer allowed the accurate determination of the Guinier R_g value within the limits of $q_{max}R_g < 1.3$. No crowding and 125 mg/mL Ficoll 400 experiments were conducted in 28% D_2O while the 125 mg/mL γ-crystallin experiments were conducted in 40% D_2O. All experiments used 2.5 mg/mL deuterated αBc.

condition indicating slightly more unfolding (Fig. 3, *C*, *black* and *orange*), commensurate with the D_m values of 3.7 ± 0.3 and 2.7 ± 0.2 respectively (Fig. S3, *B*). Crowding with 125 mg/mL γ-crystallin induces further unfolding relative to both of the other conditions as implied by the Kratky plot and a D_m value of 2.2 ± 0.1 (Figs. 3, *C*, *blue* and S3, *B*, *blue*).

Given the decrease in thermal stability with increasing Ficoll 400 concentration (Fig. 1, *E*), αBc was examined to determine whether it was impacted in a similar way by 125 mg/mL γ-crystallin. Data were acquired from 30 – 80 °C by measuring the low-q D_m value for αBc under each condition via fitting of the power-law function (Fig. S3, *C* – *E*) as used in Fig. 1, *D*. The resultant data could be fitted to a sigmoidal Boltzmann function to extract the T_{agg} value, similar to that obtained from light scattering profiles (Fig. S1, *E*). The T_{agg} values for αBc under conditions of no crowding and 125 mg/mL Ficoll 400 were 64.3 ± 0.3 °C and 55.1 ± 1.8 °C respectively (Fig. 3, *D*, *black* and *orange*), largely in agreement with the T_{agg} from light scattering data (Fig. 1, *E*, *orange*). However, 125 mg/mL γ-crystallin induced a T_{agg} value for αBc of 42.4 ± 1.1 °C (Fig. 3, *D*, *blue*). Thus, the thermal stability of αBc is mildly affected in 125 mg/mL Ficoll 400 but markedly so in 125 mg/mL γ-crystallin relative to the absence of crowding agents.

Owing to the markedly reduced thermal stability of αBc, differences in the effects of macromolecular crowding between Ficoll 400 and bovine γ-crystallin on the kinetics of structural and size changes for αBc at 37 °C were evident from the low- and mid-q region scattering profiles (Fig. S3, *F* and *G*). Moreover, the Kratky plot showed significant structural change for αBc in 125 mg/mL γ-crystallin over 36 hours from partially unfolded to increasingly unfolded as indicated by the steady loss of the bell-shaped peak (Fig. 3, *E*, *blue*) while the structure of αBc in 125 mg/mL Ficoll 400 αBc remained unchanged (Fig. 3, *E*, *orange*). Similarly, the Guinier-derived R_g value for αBc in 125 mg/mL γ-crystallin was elevated to 88.9 ± 2.1 Å at 0 hour, then increased further to 101 ± 4.1 Å between 5 – 6 hours and was highly aggregated at time points after that (Fig. 3, *F*, *blue*). In contrast, αBc in 125 mg/mL Ficoll 400 had no significant change in both Kratky plot and R_g value giving an average R_g value of 57.9 Å at 37 °C with a standard deviation of 1.0 Å over 36 hours (Fig. 3, *F*, *orange*).

These thermodynamic and kinetic results imply that changes to the structure, stability and aggregation propensity of αBc are more pronounced under conditions of macromolecular crowding

induced by biologically relevant bovine γ-crystallin than by the polysaccharide Ficoll 400. A variety of reasons could account for this: (i) the large molar concentration difference (approximately 20 times higher) of γ-crystallin compared to Ficoll 400, (ii) the greater excluded volume effects for the smaller crowder, γ-crystallin (59), or (iii) 'soft' interactions between the protein crowder and αBc that preferentially destabilize αBc relative to that induced by Ficoll 400 (60). However, importantly, the underlying destabilizing effects on αBc structure and aggregation are the same as those noted at high concentrations of Ficoll 400. It is concluded therefore that the effects on αBc structure using Ficoll 400 as a crowding agent observed herein are physiologically relevant.

Macromolecular crowding diminishes the molecular chaperone ability of αBc

The molecular chaperone action of αBc, along with that of αAc, is the primary defence against the unfolding and aggregation of the structural lens proteins, β- and γ-crystallin (61) and is of importance extralenticularly in cellular proteostasis (62). Thus, it is important to understand how macromolecular crowding affects the chaperone action of αBc. Accordingly, this was assessed *in vitro* against the amorphous aggregation of insulin induced via reduction of its disulfide bonds in the presence of increasing Ficoll 400 concentration. Reduced insulin and αBc were examined individually (Fig. S4, *A* and *C*) and also together in a 1:2 w:w ratio, respectively (Fig. S4, *B*). From these experiments, aggregation metrics including the lag time, rate of aggregation and maximum light scattering (Fig. 4, *A* – *C*, respectively), were calculated in the absence of Ficoll 400 and in the presence of increasing Ficoll 400 concentration.

The effect of macromolecular crowding on insulin aggregation at 37 °C results in a decrease in lag time and increase in both the rate of aggregation and maximum light scattering with increasing Ficoll 400 concentration (Fig. 4, *A* – *C*, *blue*), consistent with macromolecular crowding theory describing the promotion of association and aggregation (63). The effect of macromolecular crowding on 0.3 mg/mL αBc aggregation (i.e. the control, without any insulin present) is not evident until 250 mg/mL Ficoll 400 (Fig. 4, *A* – *C*, *yellow* and Fig. S4, *C*) under which conditions the lag time is twelve times longer than that of insulin at the same Ficoll 400 concentration (Fig. 4, *A*). Therefore, αBc acts (in a chaperone manner) on reduced insulin (its B chain) while the former is in an oligomeric state i.e.

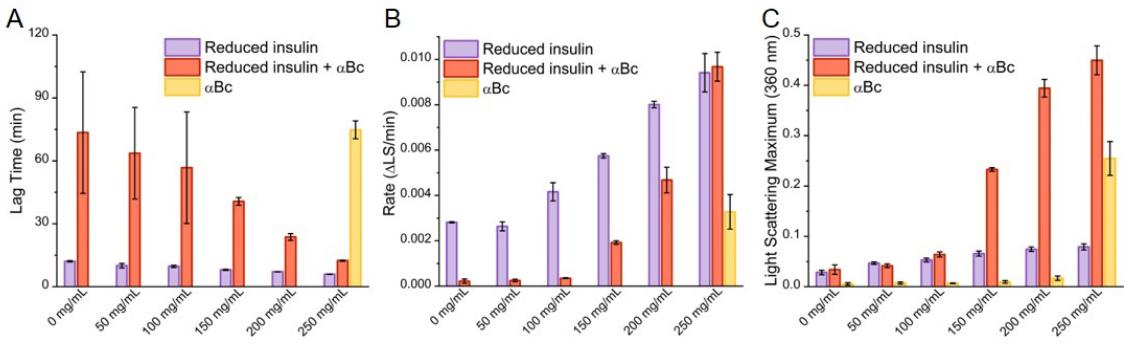


Figure 4. Chaperone ability of αBc against reduced insulin aggregation under conditions of increasing macromolecular crowding. *A*, lag time, *B*, rate of aggregation, and *C*, light scattering maximum for reduced insulin (*blue*), reduced insulin with αBc in at 1:2 w:w ratio (*red*) and αBc alone (*yellow*) at 37 °C. Errors are given as the standard deviation of three independent repeats.

before it has amorphously aggregated. Within the range of 0 – 100 mg/mL Ficoll 400, insulin aggregation is significantly inhibited by the chaperone action of αBc as shown by an extended lag time and slower rate of aggregation (Fig. 4, *A* and *B*, *red*). At Ficoll 400 concentrations equal to or greater than 150 mg/mL, the chaperone ability of αBc was impaired. Relative to insulin alone, the lag time for the insulin-αBc complex was approximately five, three, and two times longer at 150, 200 and 250 mg/mL Ficoll 400, respectively (Fig. 4, *A*, *red*). Similarly, the rate of aggregation was approximately three and two times slower for 150 and 200 mg/mL Ficoll 400, respectively, however there was no difference at 250 mg/mL Ficoll 400 (Fig. 4, *B*, *red*).

At Ficoll 400 concentrations of 150 mg/mL and above, aggregates of the insulin-αBc complex (Fig. 4, *C*, *red*) exhibit a light scattering maximum that is significantly greater than insulin alone (Fig. 4, *C*, *blue*). At 250 mg/mL Ficoll 400, the light scattering maximum for the insulin-αBc complex was 1.8 times greater than αBc alone. Thus, the destabilized target protein, i.e. reduced insulin, can induce co-aggregation of the chaperone αBc under highly crowded conditions.

Aggregation of αBc under conditions of macromolecular crowding can be curtailed by lens partner protein αAc via an increase in thermal stability

In addition to αBc, the eye lens also contains molecular chaperone protein αAc which forms hetero-oligomers with αBc, as described above at a ratio of approximately 1:3 αBc:αAc (14,15). To understand

the impact of α-crystallin hetero-oligomerization on kinetic and thermodynamic stability of αBc in a highly crowded environment, we examined both proteins separately and in 1:1, 1:2, and 1:3 αBc:αAc w:w ratios at 250 mg/mL Ficoll 400. Aggregation kinetics at 37 ºC showed that αBc aggregates in a concentration-dependent manner (Fig. 5, *A, light blue* and *purple*) (observed previously in Fig. 2, *H*), in contrast with αAc which does not aggregate under the same concentrations (Fig. 5, *A, grey* and *dark grey*). Mixing αBc:αAc at 1:1 and 1:2 w:w ratios significantly decreased the amount of aggregation relative to 0.1 mg/mL αBc alone and aggregation was no longer discernible at a ratio of 1:3 (Fig. 5, *A, yellow*, *orange*, and *red*, respectively). The percentage protection against aggregation for 1:1, 1:2 and 1:3 αBc:αAc was subsequently quantified as 59.3 ± 6.0 %, 78.6 ± 4.3 %, and 92.5 ± 1.7 % respectively (Fig. S5, *A*), relative to the light scattering maximum at 0.1 mg/mL αBc.

To understand the basis of the difference in aggregation propensity between αBc and αAc in a highly crowded environment, we examined the thermal stability of the homo- and hetero-oligomers at the same concentrations used in the kinetic aggregation assay (Fig. S5, *B*). The thermal stability of αBc in a highly crowded environment indicates that the protein is highly destabilized in a concentration-dependent manner as indicated by T_{agg} values of 44.9 ± 0.7 ºC and 49.7 ± 1.0 ºC for 0.3 and 0.1 mg/mL αBc, respectively (Fig. 5, *B, blue* and *light blue*). Interestingly, the thermal stability of αAc is not altered significantly by the same changes in protein concentration (Fig. 5, *B, grey* and *dark grey*). Mixing αBc with αAc led to enhanced thermal stability with the hetero-oligomeric complex having a ΔT_{agg} value of + 6.0 ºC on average relative to 0.1 mg/mL αBc (Fig. 5, *B, yellow*, *orange*, and *red*).

The relationship between the steady-state light scattering maximum from the 37 ºC kinetic aggregation assay and the T_{agg} value indicative of homo- and hetero-oligomer thermal stability was examined. Both values were plotted on a scatter plot and the points fitted well to a single exponential function (R^2 = 0.99) (Fig. 5, *C*). This correlation provides novel insight into how lens α-crystallin stability in a highly crowded environment can eventually lead to an exponential increase in light scattering under physiological conditions.

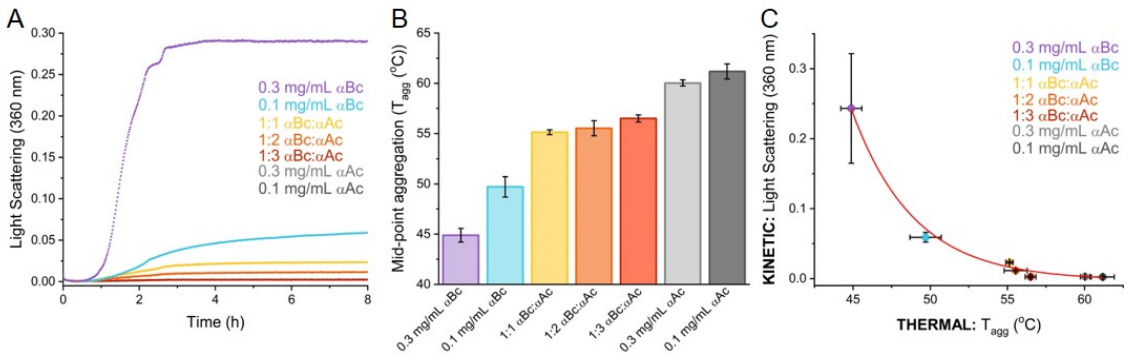


Figure 5. Chaperone action of αAc toward αBc and subsequent complex stability under conditions of macromolecular crowding. *A*, chaperone assay containing αBc and αAc in 250 mg/mL Ficoll 400 at 37 ºC. Both 0.3 and 0.1 mg/mL αAc exhibited no detectable light scattering over 8 hours (i.e. their scattering trace is below that for the 1:3 αBc:αAc trace). Data are the average of three independent repeats. *B*, T_{agg} of homo- and hetero-oligomers of αBc and αAc in 250 mg/mL Ficoll 400 with column colours corresponding to those indicated in *A*. Errors are the standard deviation of three independent repeats. *C*, light scattering maximum from the kinetic study (*A*) and the T_{agg} from the thermal stability study (*B*) reveal an exponential relationship between maximum light scattering in the kinetic study and complex thermal stability in 250 mg/mL Ficoll 400. The ratios of 1:1, 1:2, and 1:3 αBc:αAc w:w are based on an αBc concentration of 0.1 mg/mL. Each mixture was equilibrated for one hour at room temperature before being used.

Discussion

Macromolecular crowding has manifold consequences that can be difficult to predict for proteins with dynamic, supramolecular structures such as αBc and other sHsps. Previous studies have noted that macromolecular crowding has a largely disruptive effect on both physical and functional characteristics of αBc, αAc and whole α-crystallin (35,64-66). Given a number of studies implicate αBc aggregation in cataract, this led us to reason that the highly crowded *in vivo* environment of the eye lens could have an adverse impact on the structure and function of αBc, resulting in it contributing to numerous cataract phenotypes (18,23,24,67,68). Further, as macromolecular crowding reduces the chaperone function of both homo- and hetero-oligomeric forms, the biological purpose of the association of αBc with αAc to form α-crystallin under such conditions warrants examination. Our current study significantly builds on previous findings for αBc under conditions of macromolecular crowding, and places the results in the context of the highly crowded eye lens where the preservation of crystallin stability is vital for the maintenance of lens transparency.

We have identified additional consequences of macromolecular crowding on αBc, i.e. unfolding and significant thermal destabilization, which ultimately lead to αBc aggregation (Fig. 1). The unfolding likely involves the β–sheet rich ACD which is significantly destabilized under crowded conditions. These results were recapitulated using a biologically relevant crowding agent bovine γ-crystallin, i.e. there was a clear link between the loss of αBc structure and a decrease in stability, an increase in size and/or oligomeric state, and finally aggregation (Fig. 3). This was particularly striking given the concentration of γ-crystallin used here is only a conservative representation of the crowded environment of the eye lens, meaning these effects could be further magnified *in vivo*. The aggregation of αBc under crowded conditions occurred initially by relatively fast-forming amorphous and then slow-forming fibrillar species (Fig. 2). A loss of chaperone function and significant co-aggregation of αBc with destabilized insulin in highly crowded conditions was also noted (Fig. 4). Aggregation, and hence light scattering, of αBc in highly crowded conditions was curtailed by its lens partner protein, αAc, which maintained its stability in crowded conditions and stabilized αBc such that its aggregation was prevented (Fig. 5). These processes are depicted schematically leading to either the maintenance of lens transparency or the formation of cataract (Fig. 6).

The αBc ACD forms a dimer that provides the cornerstone for higher order assembly (27,52). Specific mutations that destabilize the dimeric interface, such as R120G and D109A (both residues are involved in a salt bridge across the ACD dimer interface in the APII register (69)), lead to similar structural and functional consequences to those observed in the present study on wild-type αBc under conditions of macromolecular crowding, i.e. both mutants exhibit unfolding, destabilization, an increase in size/oligomeric state and ultimately aggregation, with R120G also displaying a loss of chaperone activity and co-aggregation with a destabilized target protein (22,68,70,71). It follows that macromolecular crowding of αBc may have a similar deleterious effect on the ACD as does these cataract-associated mutations, thereby leading to structural alterations occasioning loss-of-function and aggregation. While excluded volume effects arising from macromolecular crowding are usually associated with the stabilization of the compact native form, they also encourage protein association (21,63,72). It would seem that further oligomerization of αBc relative to the dilute environment is favoured under crowded conditions as has been demonstrated for other sHsps, namely Hsp27 (HspB1)

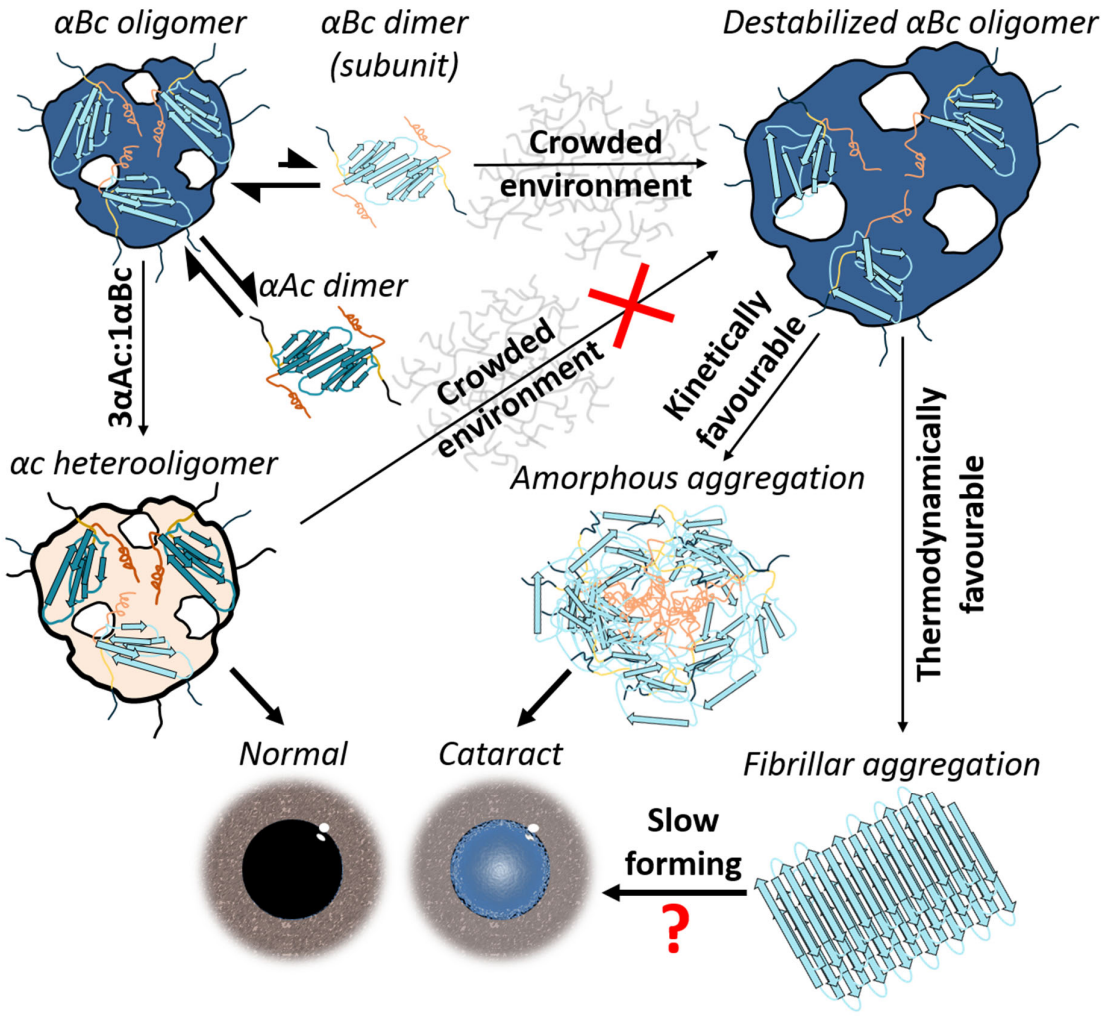


Figure 6. Schematic depicting the pathways associated with normal vision and cataract as a consequence of changes to αBc under conditions of macromolecular crowding. Under dilute conditions, oligomeric αBc (*blue*, *top left*) is folded and the molecules are coloured according to specific αBc regions as in Fig. 1, *A*, *bottom*. The oligomer is in equilibrium and undergoing subunit exchange with a dimeric species. Oligomeric αBc is destabilized under conditions of macromolecular crowding (*blue*, *top right*) and becomes unfolded, increases in size/oligomeric state, loses thermal stability, has reduced chaperone ability and is prone to co-aggregation with a destabilized and aggregating target protein. From destabilized oligomeric αBc, kinetically favourable amorphous aggregates form quickly (*middle right*). In the lens, this occurrence could potentially lead to opacification and eventually cataract formation. Thermodynamically favourable fibrillar aggregation (*bottom right*) forms slowly, likely via the amyloidogenic regions in the ACD (38) after unfolding and rearrangement. Fibrillar aggregates may be associated with late-stage cataract (*bottom left*). Hetero-oligomerization with lens partner protein αAc (*yellow*, *middle left*) prevents the aggregation of αBc via a chaperone mechanism which increases the thermal stability of αBc thus facilitating lens transparency and normal vision (*bottom left*). This protective mechanism under highly crowded conditions could be circumvented by mutations or PTMs that either further destabilize αBc or diminish αAc chaperone ability.

(73) and Hsp20 (HspB6) (74). In light of this, we suggest that the structural defects inherent to the αBc mutants like R120G and D109A are potentially exaggerated in the crowded lens, making it difficult for αAc to chaperone them effectively and leading to the large-scale aggregation that is associated with their congenital cataract phenotypes (18,68).

Partial unfolding at the ACD interface of the Hsp27 dissociated monomer under dilute conditions has been correlated with increased chaperone activity (75). Due to the conservation of the ACD, it was suggested this mechanism could be a ubiquitous feature across the sHsp superfamily. While αBc undergoes crowding-induced unfolding and significant destabilization of the ACD, the chaperone function was lost at higher crowding concentrations. The latter is consistent with previous studies using different crowding agents and target proteins (64,65). A loss of free subunits, i.e. chaperone-active dimeric species, along with an increased rate of target protein aggregation under crowded concentrations are potential reasons for the loss of chaperone function. However, as αBc's ACD alone has been shown to exhibit chaperone ability (69), it is curious that its unfolding does not lead to increased chaperone function in a similar mechanism to Hsp27, thereby offsetting the other unfavourable consequences imposed on αBc by crowding. Interestingly, the partially unfolded Hsp27 monomer species also has a heightened aggregation propensity at higher concentrations (75). Thus, there may be a delicate balance in the ACD between folding and unfolding for the optimal functioning and stability of αBc, a balance that seems to be tilted in favour of aberrant aggregation in a crowded environment. *In vivo*, PTMs that increase the chaperone activity of αBc (49,76-78) and interactions with endogenous compounds such as metals (65,79,80) could be pivotal in modulating the stability, structure and function of αBc under conditions of macromolecular crowding.

The aggregation of αBc under highly crowded conditions led to significant light scattering, as observed previously *in vivo* (23), and the appearance of kinetically distinct ThT positive fibrillar species (Fig. 2). Differences in the kinetic and concentration-dependent formation of amorphous and fibrillar aggregates suggest two possible mechanisms of formation. The first mechanism consists of two distinct formation pathways whereby both amorphous and fibrillar aggregation are in competition for the same pool of native αBc (Fig. S2, *D*, *red* and *blue*). The second mechanism is a single pathway where amorphous aggregation occurs as a metastable intermediate in the formation of fibrillar aggregates (81)

(Fig. S2, *D*, *purple*). This competitive process is depicted using a theoretical reaction coordinate diagram (Fig. S2, *D*) where, in both pathways, amorphous aggregation has a lower activation energy ($\Delta G^{\neq}$), indicating its kinetic favourability, while fibrillar aggregation has a significantly lower free energy (ΔG^{0}), indicating its thermodynamic favourability (Fig. S2, *D*). Fibrillar aggregates are generally the most thermodynamically favourable protein conformation (82) which is even more so the case in a crowded environment as they provide greater reduction in excluded volume (63). However, their formation in this instance, i.e. constant crowding concentration and temperature, is only viable at high αBc concentrations to reach the requisite higher activation energy for formation, relative to amorphous aggregate formation (Fig. 2, *H*). One potential reason for the slow kinetics of αBc fibrillar aggregate formation under highly crowded conditions, even at high αBc concentrations, is that significant structural rearrangement precedes this process. As much of the ACD of αBc is predicted to be amyloidogenic, and is ostensibly 'protected' from this high amyloidogenicity by the presence of the unstructured and dynamic flanking NTR and CTR (38), significant unfolding and destabilization of the ACD, which occurs as a consequence of crowding (Fig. 1), is likely required before the fibrillar aggregation pathway can be accessed. In agreement with this, under dilute conditions, αBc is converted into amyloid fibrils with mild denaturant and heat (83,84).

While amorphous aggregates are considered the hallmark of cataract (9,85), recent findings have presented evidence for amyloid structures in animal lenses (86) and post-mortem human lens tissues (87). Interestingly, while amyloid structures were not detected in juvenile human lenses, they were identified in mature non-cataract and cataract lenses (87). The authors suggested that amyloid structures may form before typical cataract symptoms manifest (87). The low abundance of amyloid structures observed in human cataract lenses (~7 % on average), can be explained by our observation *in vitro* that in the crowded environment of the lens the formation of mature fibres is kinetically slow. Thus, advanced, late-stage cataract lenses may exhibit fully formed fibrillar species, similar to those observed in this study. Potentially, this would be a useful *in vivo* finding as any therapeutic cataract treatment that staves off amorphous aggregation may still allow the formation of late-stage fibrillar aggregates, at least for αBc, which might also lead to a cataract-like phenotype.

Under dilute *in vitro* thermal stress conditions, it has been well demonstrated that αBc is the least stable α-crystallin subunit and that its thermally induced aggregation is prevented by the more stable αAc subunit (50,88-90). However, the effect of macromolecular crowding on this paradigm between αBc and αAc has yet to be explored despite its relevance to the eye lens. There was a clear contrast in the stability of the two subunits in highly crowded conditions with αBc being thermally and kinetically destabilized while αAc remained stable under simulated physiological conditions. Upon mixing αBc and αAc to form the α-crystallin hetero-oligomer, the thermal stability of the complex increased with increasing abundance of αAc. At a physiologically relevant ratio of 1:3 αBc:αAc, the kinetically favoured amorphous aggregation of αBc at 37 ºC was inhibited. These findings are consistent with the development of cataract in the mouse αAc and αBc knockout studies (23,25). The difference in stability could be related to the lower isoelectric point of αAc (3), thereby increasing electrostatic repulsion and discouraging aggregation, or features intrinsic to αAc's native structure, such as disulfide bonding (91), which affords it further protection.

The trend between αBc and αAc thermal stability and kinetic aggregation propensity in the crowded environment indicate that a loss of stability, below a threshold temperature, results in an exponential increase in kinetic light scattering (Fig. 5, *C*). While congenital cataract-associated αBc mutants, such as D109A (68) (which is also a mimic for αBc Asp isomerization (71)), R120G (18,70), D140N (92), and Q151X (93) largely exhibit different structural defects, the loss of stability relative to the wild-type is a common characteristic of all these mutants. Given αBc's susceptibility to destabilization and aggregation under crowded conditions, it can be expected that destabilizing mutations or PTMs could further impact αBc stability and hence increase kinetic aggregation. Likewise, mutations or PTMs that affect the stability and thereby the chaperone ability of αAc, such as the F71L (94), will allow the development of early-onset cataract due to ineffective stabilization of αBc. One of the therapeutic treatment options for cataract is the use of small molecules that aim to stabilize the native state of crystallins (9). One such promising treatment involves the specific thermal stabilization of α-crystallin *in vitro* via a sterol-based pharmacological chaperone that binds to the interface of the αBc ACD dimer (95). The sterol reversed cataract in an *in vivo* mouse model and improved the solubility of crystallins from human cataract lenses *ex vivo*. Our study affirms that approaches such as these are

viable to prevent the kinetic aggregation of αBc associated with cataract, and further imply that compound screening via stability assays in a simulated crowded environment might better inform current drug discovery efforts. Finally, we posit that these findings are also applicable to αBc more generally in the extra-lenticular environment where other sHsps such as the populous Hsp27 also stabilizes αBc *in vitro* (96) and *in vivo* (97).

Experimental procedures

Materials

All materials, including buffer components, Ficoll 400, ThT, dithiothreitol (DTT), and insulin from bovine pancreas (I5500) were purchased from Sigma. Prior to SANS experiments, 1 mL aliquots of 25 mM sodium phosphate buffer, 100 mM NaCl, pH 7.3 made up in H_2O, were lyophilized, and reconstituted with the desired D_2O:H_2O ratio. Preparation of 1 mL of 450 mg/mL Ficoll 400 stock solution involved weighing 450 mg of Ficoll 400 on an analytical balance and adding 625 µL of 25 mM sodium phosphate buffer, 100 mM NaCl, pH 7.3. The volume to be added was calculated based on the expected solution displacement of a 450 mg/mL Ficoll 400 solution with a density of 1.2 g/mL which was experimentally verified to produce a 1 mL solution. Upon buffer addition to Ficoll 400, the solution was left to solubilize at room temperature overnight or until clear. The solution was then centrifuged and diluted with the same buffer as needed. Ficoll 400 solutions used in SANS experiments were made the same way with the only change being the use of a buffer with the desired D_2O:H_2O ratio. Protein concentrations were determined using an ultra-low volume spectrophotometer (Nanodrop (Thermo Fisher Scientific Inc.)). An $E^{1\%}$ (280 nm) of 8.1 and 8.2 was used for αAc and αBc respectively (13), an average 21.0 was used for bovine γ-crystallin (98) and $E^{1\%}$ (278 nm) value of 10.6 was used for insulin as per the manufactures product information sheet.

Protein Expression, Deuteration and Purification

The expression vectors pET24a and pET43.1a encoding recombinant human αBc and αAc (Genscript) genes were transformed and expressed in BL21(DE3) *E. coli* cells. Both proteins were purified as previously described (99) with some modifications. Briefly, BL21(DE3) cells were induced

using IPTG and left overnight, with shaking, at room temperature. The resuspended cell pellet was lysed using sonication. The cell lysate was initially purified using anion-exchange chromatography (HiPrep™ DEAE FF 16/10 (GE Healthcare)) using a 20 mM Tris-HCl buffer, pH 8.5 and a 0 – 1 M NaCl gradient. Fractions containing the protein were concentrated and finally subjected to gel filtration chromatography ((HiPrep™ 16/60 Sephacryl® S-300 HR (GE Healthcare)), in 50 mM sodium phosphate buffer, 100 mM NaCl, pH 7.4). Purified fractions were dialysed against MilliQ water, lyophilized and stored at - 20 ºC.

Deuterated αBc was overexpressed in BL21(DE3) *E.coli* cells after being transformed with pET24a plasmid containing the human αBc gene sequence. The cells from single colonies were selected with kanamycin and gradually adapted to D_2O-based minimal "Mod C1" media (100) containing a step-wise increasing concentration of D_2O from 0 to 100 %, and stored as glycerol stocks at - 80 ºC until needed. The starter culture of 100 mL of 100 % D_2O minimal medium containing 40 g/L glycerol-d8 (Cambridge Isotope Labs) was used to inoculate 2 L RTF-5 fermentor bioreactors (Real Time Engineering) containing 900 mL of the same medium. Cells were cultured at 37 ºC with automatically regulated pH and dissolved oxygen, until the OD_{600} reached a value of 13 – 15, and were then induced to express αBc with 1 mM IPTG and the temperature lowered to 25 ºC. After 15 hours, the biomass was harvested and stored at - 80 ºC or on dry ice until purification. Deuterated αBc was purified as described above for the hydrogenated form. MALDI-TOF mass spectrometry was carried out by the Mass Spectrometry Core Facility, University of Sydney, to determine the deuteration level of the protein. By comparing the mass differences between H- and D-protein fragments relative to the calculated numbers of non-exchangeable H/D atoms (bound to carbon or sulfur), the level of deuterium substitution of αBc was determined to be 99.8%.

Bovine γ-crystallin was extracted and purified from whole lens homogeneate as previously described (101) and modified for a Tris buffer system (102). A HiPrep™ 16/60 Sephacryl® S-300 HR (GE Healthcare) gel filtration column was used to isolate the γ-crystallin fraction (Fig. S3, *A*) which was dialysed against MilliQ water, lyophilized and stored at - 20 ºC.

Small Angle Neutron Scattering

SANS experiments were conducted on the Quokka SANS instrument located at the Australian Centre for Neutron Scattering (ACNS), ANSTO, Sydney (103). Neutrons ($\lambda = 5$ Å) from the cold source OPAL reactor were detected utilizing 1.3 and 12 m detector lengths which provided a q range of 0.007 – 0.4 $Å^{-1}$. Samples were loaded into a 2 mm circular quartz cuvette which was sealed with a stopper and parafilm to avoid evaporative loss. The data were converted to absolute scale via normalization with the incident beam flux.

In order to exclusively observe neutron scattering from αBc under crowded conditions, we used deuterated αBc and eliminated the high background scattering of the hydrogenated crowding agents Ficoll 400 and bovine γ-crystallin through contrast matching. The D_2O match points of all solution components were predicted via the molecule's scattering length density and experimentally tested using contrast-series SANS experiments. The match points were 100, 28 and 40 % D_2O for deuterated αBc, hydrogenated Ficoll 400 and, hydrogenated bovine γ-crystallin, respectively. As such, experiments using hydrogenated Ficoll 400 were undertaken in 28 % D_2O while those using hydrogenated bovine γ-crystallin were undertaken in 40 % D_2O. Deuterated αBc in dilute conditions was studied with 28 % D_2O present. All SANS studies used 2.5 mg/mL deuterated αBc in a 25 mM sodium phosphate buffer, 100 mM NaCl, pH 7.3 along with relevant percentage of D_2O.

Initial experiments using deuterated αBc and hydrogenated Ficoll 400 at 0, 50, 100, 125, 150, 200, and 250 mg/mL were conducted at temperatures of 3, 25, 37, and 45 ºC which was sequentially ramped over approximately 48 hours. Experiments to compare and contrast the effect of a soluble polysaccharide, hydrogenated Ficoll 400, and a relevant protein, hydrogenated bovine γ-crystallin, as crowding agents on deuterated αBc were conducted thermodynamically and kinetically at a crowding concentration of 125 mg/mL. Thermodynamic experiments involved temperature ramping and scattering collection at eight temperatures within the range of 30 – 80 ºC. Kinetic experiments were completed at 37 ºC for a total of 36 hours. Scattering from all SANS experiments was collected for 1.5 hours per crowding concentration, temperature or time-point.

The data were reduced to a 1D scattering profile using in-house macros within the IgorPro software suite (104). The data were subtracted from the scattering of an empty cell and buffer using

either IgorPro or PRIMUS (105). Low-q points near the beam stop were excluded. The data were analyzed by determining the Guinier R_g and I(0) value using PRIMUS and the real-space R_g and D_{max} via the P(r) function calculated using GNOM (106). Analysis using power-law fitting (Eqn. 1) was completed using SasView 4.2.0 (107).

$$I(q) = s \cdot |q|^{-D_m} + b \quad \text{(Eqn. 1)}$$

where *I(q)* is the scattering intensity in cm^{-1}, *s* is a scaling factor, D_m is the Porod exponent (otherwise known as the 'fractal dimension') and *b* is the background in cm^{-1}. Power-law analysis in the mid-q range ($0.05\ Å^{-1} < q < 0.40\ Å^{-1}$) provided the D_m value which is directly related to the overall structure of the scattering molecule based on internal scaling inherent to its inter-atomic distances. However, power-law analysis in the low-q range ($0.006\ Å^{-1} < q < 0.018\ Å^{-1}$) allowed a reproducible unitless measure of protein aggregation. For instance, if the determination of R_g via the Guinier plot no longer satisfied the condition of $qR_g \leq 1.3$ due to an upturn, this method provided a measure of the gradient of that upturn with a higher gradient value indicating greater aggregation. The data were also analyzed using Kratky or normalized Kratky plots. The latter was dependent on the acquisition of R_g and I(0) values.

Thermal Stability

CD, Trp fluorescence, and light scattering thermal stability experiments were conducted at 0.3 mg/mL αBc in 25 mM sodium phosphate buffer, 100 mM NaCl, pH 7.3 using 0, 50, 100, 150, 200, and 250 mg/mL Ficoll 400. Experiments were performed using an Applied Photophysics Chirascan spectrophotometer with scanning emission monochromator attachment. CD spectra were acquired using a 1 mm path length quartz cuvette while Trp fluorescence and light scattering measurements were undertaken using a 1 cm path length quartz cuvette. Both cuvettes were fitted with stoppers to minimize evaporative loss. Step-wise temperature ramping was used from 25 – 90 °C at 0.5 °C increments per step using a Quantum Northwest TC 125 Peltier temperature controller. There was a 30 second equilibration time between each incremental temperature increase. CD spectra were acquired from 200 – 260 nm and Trp fluorescence spectra from 300 – 400 nm both with a wavelength step width of 1 nm

acquired for 1.5 seconds per step. Light scattering data were acquired and analyzed against temperature at the single wavelength of 360 nm to determine T_{agg}. The CD ellipticity at 218 nm, sensitive to β–sheet, was analyzed against temperature to obtain a T_m value largely related to the β-sheet rich ACD. A plot of the Trp fluorescence spectra versus temperature provided a T_m value related to the local environment of the NTR of αBc and were best analyzed using the whole spectrum for each temperature via the λ_{bcm} (Eqn. 2).

$$\lambda_{bcm} = \frac{\sum I(\lambda)\cdot\lambda}{\sum I(\lambda)} \quad \text{(Eqn. 2)}$$

where λ is a wavelength within the emission range and $I(\lambda)$ is the fluorescence intensity at that wavelength.

Both T_m and T_{agg} values were elucidated via fitting of either a single Boltzmann function (Eqn. 3) for distinct monophasic transitions or double Boltzmann function (Eqn. 4) for clear biphasic transitions using the mid-point(s) of the sigmoidal transition (x_0, *x01*, and *x02*).

$$y = \frac{A_1 - A_2}{1+e^{(x-x_0)/dx}} + A_2 \quad \text{(Eqn. 3)}$$

where A_1 and A_2 are the lower and upper asymptotic values, respectively, *dx* is the rate of unfolding/aggregation against temperature, and x_0 is the mid-point (T_m or T_{agg}) of the monophasic transition.

$$y = y_0 + A\left[\frac{p}{1+e^{\frac{x-x01}{k1}}} + \frac{1-p}{1+e^{\frac{x-x02}{k2}}}\right] \quad \text{(Eqn. 4)}$$

where y_0 is the y-axis offset, *A* is the sum of the first and second asymptotes, *p* is the fraction of *A* that is equal to the value of the first asymptote (i.e. if *A* = 10 and the first asymptote is 3, then *p* = 0.3), *k1* and *k2* are the slope factors for the first and second transition, respectively, and *x01* and *x02* are the mid-points of the first and second transitions in the biphasic sigmoidal curve, respectively.

Thermal aggregation experiments on both αBc and αAc alone were undertaken at 0.1 and 0.3 mg/mL while mixtures of 1:1, 1:2, and 1:3 αBc:αAc were based on w:w ratios with αBc having a fixed concentration of 0.1 mg/mL. Mixtures of αBc and αAc were allowed to equilibrate under dilute conditions for 1 hour at room temperature before acquisition. The experiments were all conducted at

250 mg/mL Ficoll 400 with the same settings and parameters as previously described for light scattering T_{agg} measurements at 360 nm.

Aggregation and Chaperone Assays

Kinetic aggregation and chaperone assays were performed using a Biotek Synergy 2 microplate reader. Light scattering was monitored at 360 nm while ThT fluorescence used a 440 nm filter for excitation and a 485 nm filter for emission at a final dye concentration of 20 μM. Both clear (light scattering) and black (ThT fluorescence) 96 well plates (Greiner Bio-One) were sealed to prevent evaporative loss and a 'slow shaking' setting was engaged for the duration of all assays. All experiments were conducted at 37 °C in 25 mM sodium phosphate buffer, 100 mM NaCl, pH 7.3.

The aggregation of αBc in highly crowded conditions was monitored with light scattering and ThT fluorescence using 250 mg/mL of spectroscopically silent Ficoll 400. Concentrations of αBc studied were 0.03, 0.1, 0.3, 0.5, 1.0, 1.5, 2.0, 2.5, and 3.0 mg/mL. The maximum light scattering and fluorescence from their respective assays were measured and plotted against αBc concentration.

The chaperone assay using reduced insulin was modified from a previous study (49) to allow for changes in lag time and rate of aggregation due to the presence of crowding agent. Modifications included a lower insulin (0.15 mg/mL) and DTT (10 mM) concentration which allowed a gradual change in aggregation across the experimental crowding series. The experimental crowding series consisted of 0, 50, 100, 150, 200, and 250 mg/mL of Ficoll 400 with a fixed concentration of 0.3 mg/mL αBc giving a 1:2 w:w ratio of insulin:αBc. Aggregation kinetics were fitted to a single Boltzmann function (Eqn. 3) using Origin (OriginLab Corporation) and the fitting parameters were used to calculate the lag time, rate of aggregation and maximum light scattering as previously described (108).

The kinetic aggregation of αBc, αAc and mixed hetero-oligomers was studied via light scattering. The experimental conditions related to the mixtures of αBc and αAc were the same as described in the final *Thermal Stability* section and the experiment was conducted over 8 hours. The percentage protection of the αBc:αAc hetero-oligomers versus 0.1 mg/mL αBc alone was calculated as previously described (108).

Transmission Electron Microscopy

TEM samples were prepared using a 1 in 5 dilution with Milli-Q water of 2.5 mg/mL αBc at 0 hour in 0 mg/mL Ficoll 400 and at 0 , 0.5 , 5 , 16 , and 33 hours in 250 mg/mL Ficoll 400 after incubation at 37 °C. The diluted sample (2 μL) was added to Formvar- and carbon-coated copper grids (ProSciTech, Australia). The grids were then washed three times with 10 μL of Milli-Q water and negatively stained with 10 μL of uranyl acetate (2 % w/v). Samples were viewed using a Hitachi H7100FA transmission electron microscope (Tokyo, Japan).

Author contributions– A.B.G., and J.A.C. conceptualization; A.B.G. data curation; A.B.G., and J.P.M. formal analysis; A.B.G., A.R., J.P.M., and J.A.C. funding acquisition; A.B.G., A.R., J.P.M., and D.C.T. investigation; A.B.G., and J.A.C. project administration; J.A.C. supervision; A.B.G. visualization; A.B.G. writing – original draft; A.B.G., A.R., J.P.M., D.C.T., and J.A.C. writing – review and editing.

Acknowledgements – We thank Prof. Heath Ecroyd (University of Wollongong) for insightful discussions. We thank Ms. Lily Wang (Australian National University) for the purification of bovine γ-crystallin. Access to SANS (Quokka, Lucas Heights, Sydney, Australia) at the Australian Centre for Neutron Scattering was supported through ANSTO beamtime awards P3858, P5094, and P6131. Access to large-scale protein deuteration facilities at the National Deuteration Facility was supported through ANSTO beamtime award P6132. The National Deuteration Facility at ANSTO is partly supported by the National Collaborative Research Infrastructure Strategy – an initiative of the Australian Government. The Centre for Advanced Microscopy at ANU for is acknowledged for their assistance with TEM. A.B.G. was supported by an Australian Postgraduate Award. D.C.T. was supported by a grant to J.A.C from the National Health and Medical Research Council of Australia (grant #1068087).

References

1. Driessen, H. P. C., Herbrink, P., Bloemendal, H., and de Jong, W. W. (1981) Primary structure of the bovine β-crystallin bp chain. *Eur. J. Biochem.* **121**, 83-91
2. Hogg, D., Tsui, L. C., Gorin, M., and Breitman, M. L. (1986) Characterization of the human beta-crystallin gene Hu beta A3/A1 reveals ancestral relationships among the beta gamma-crystallin superfamily. *J. Biol. Chem.* **261**, 12420-12427
3. Bloemendal, H., de Jong, W., Jaenicke, R., Lubsen, N. H., Slingsby, C., and Tardieu, A. (2004) Ageing and vision: structure, stability and function of lens crystallins. *Prog. Biophys. Mol. Bio.* **86**, 407-485
4. Delaye, M., and Tardieu, A. (1983) Short-range order of crystallin proteins accounts for eye lens transparency. *Nature* **302**, 415-417
5. Ponce, A., Sorensen, C., and Takemoto, L. (2006) Role of short-range protein interactions in lens opacifications. *Mol. Vis.* **12**, 879-884
6. Zhao, H., Brown, P. H., Magone, M. T., and Schuck, P. (2011) The molecular refractive function of lens γ-crystallins. *J. Mol. Biol.* **411**, 680-699
7. Bassnett, S. (2002) Lens organelle degradation. *Exp. Eye. Res.* **74**, 1-6
8. Lynnerup, N., Kjeldsen, H., Heegaard, S., Jacobsen, C., and Heinemeier, J. (2008) Radiocarbon dating of the human eye lens crystallines reveal proteins without carbon turnover throughout life. *PLOS ONE* **3**, e1529
9. Moreau, K. L., and King, J. A. (2012) Protein misfolding and aggregation in cataract disease and prospects for prevention. *Trends Mol. Med.* **18**, 273-282
10. Truscott, R. J. W., and Friedrich, M. G. (2016) The etiology of human age-related cataract. Proteins don't last forever. *Biochim. Biophys. Acta, Gen. Subj.* **1860**, 192-198
11. Clark, A. R., Lubsen, N. H., and Slingsby, C. (2012) sHSP in the eye lens: crystallin mutations, cataract and proteostasis. *Int. J. Biochem. Cell Biol.* **44**, 1687-1697
12. Horwitz, J. (1992) Alpha-crystallin can function as a molecular chaperone. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 10449-10453

13. Thomson, J. A., and Augusteyn, R. C. (1985) Ontogeny of human lens crystallins. *Exp. Eye. Res.* **40**, 393-410

14. Horwitz, J., Bova, M. P., Ding, L.-L., Haley, D. A., and Stewart, P. L. (1999) Lens α-crystallin: function and structure. *Eye* **13**, 403-408

15. Ma, Z., Hanson, S. R. A., 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

16. Piatigorsky, J. (1991) The recruitment of crystallins: new functions precede gene duplication. *Science* **252**, 1078-1079

17. Renkawek, K., Voorter, C. E. M., Bosman, G. J. C. G. M., van Workum, F. P. A., and de Jong, W. W. (1994) Expression of αB-crystallin in Alzheimer's disease. *Acta Neuropathol.* **87**, 155-160

18. Vicart, P., Caron, A., Guicheney, P., Li, Z., Prévost, M.-C., Faure, A., Chateau, D., Chapon, F., Tomé, F., Dupret, J.-M., Paulin, D., and Fardeau, M. (1998) A missense mutation in the αB-crystallin chaperone gene causes a desmin-related myopathy. *Nature Genet.* **20**, 92-95

19. Selcen, D., and Engel, A. G. (2003) Myofibrillar myopathy caused by novel dominant negative αB-crystallin mutations. *Ann. Neurol.* **54**, 804-810

20. Zimmerman, S. B., and Minton, A. P. (1993) Macromolecular crowding: biochemical, biophysical, and physiological consequences. *Annu. Rev. Biophys. Biomol. Struct.* **22**, 27-65

21. Kuznetsova, I., Turoverov, K., and Uversky, V. (2014) What macromolecular crowding can do to a protein. *Int. J. Mol. Sci.* **15**, 23090-23140

22. Treweek, T. M., Rekas, A., Lindner, R. A., Walker, M. J., Aquilina, J. A., Robinson, C. V., Horwitz, J., Der Perng, M., Quinlan, R. A., and Carver, J. A. (2005) R120G αB-crystallin promotes the unfolding of reduced α-lactalbumin and is inherently unstable. *FEBS J.* **272**, 711-724

23. Brady, J. P., Garland, D., Duglas-Tabor, Y., Robison, W. G., Groome, A., and Wawrousek, E. F. (1997) Targeted disruption of the mouse αA-crystallin gene induces cataract and cytoplasmic

inclusion bodies containing the small heat shock protein αB-crystallin. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 884-889

24. Pras, E., Frydman, M., Levy–Nissenbaum, E., Bakhan, T., Raz, J., Assia, E. I., Goldman, B., and Pras, E. (2000) A nonsense mutation (W9X) in CRYAA causes autosomal recessive cataract in an inbred Jewish Persian family. *Invest. Ophth. Vis. Sci.* **41**, 3511-3515

25. Brady, J. P., Garland, D. L., Green, D. E., Tamm, E. R., Giblin, F. J., and Wawrousek, E. F. (2001) αB-crystallin in lens development and muscle integrity: a gene knockout approach. *Invest. Ophth. Vis. Sci.* **42**, 2924-2934

26. Hochberg, G. K. A., and Benesch, J. L. P. (2014) Dynamical structure of αB-crystallin. *Prog. Biophys. Mol. Bio.* **115**, 11-20

27. Jehle, S., Vollmar, B. S., Bardiaux, B., Dove, K. K., Rajagopal, P., Gonen, T., Oschkinat, H., and Klevit, R. E. (2011) N-terminal domain of αB-crystallin provides a conformational switch for multimerization and structural heterogeneity. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 6409-6414

28. Benesch, J. L. P., Ayoub, M., Robinson, C. V., and Aquilina, J. A. (2008) Small heat shock protein activity is regulated by variable oligomeric substructure. *J. Biol. Chem.* **283**, 28513-28517

29. Delbecq, S. P., and Klevit, R. E. (2019) HSPB5 engages multiple states of a destabilized client to enhance chaperone activity in a stress-dependent manner. *J. Biol. Chem.* **294**, 3261-3270

30. Bova, M. P., Ding, L.-L., Horwitz, J., and Fung, B. K.-K. (1997) Subunit exchange of αA-crystallin. *J. Biol. Chem.* **272**, 29511-29517

31. Treweek, T. M., Meehan, S., Ecroyd, H., and Carver, J. A. (2015) Small heat-shock proteins: important players in regulating cellular proteostasis. *Cell. Mol. Life Sci.* **72**, 429-451

32. Haslbeck, M., Peschek, J., Buchner, J., and Weinkauf, S. (2016) Structure and function of α-crystallins: traversing from in vitro to in vivo. *Biochim. Biophys. Acta, Gen. Subj.* **1860**, 149-166

33. Datta, S. A., and Rao, C. M. (1999) Differential temperature-dependent chaperone-like activity of αA- and αB-crystallin homoaggregates. *J. Biol. Chem.* **274**, 34773-34778

34. Chebotareva, N. A., Eronina, T. B., Roman, S. G., Mikhaylova, V. V., Sluchanko, N. N., Gusev, N. B., and Kurganov, B. I. (2019) Oligomeric state of αB-crystallin under crowded conditions. *Biochem. Biophys. Res. Commun.* **508**, 1101-1105

35. Ghahghaei, A., Rekas, A., Price, W. E., and Carver, J. A. (2007) The effect of dextran on subunit exchange of the molecular chaperone αA-crystallin. *Biochim. Biophys. Acta, Proteins Proteom.* **1774**, 102-111

36. Roman, S. G., Chebotareva, N. A., Eronina, T. B., Kleymenov, S. Y., Makeeva, V. F., Poliansky, N. B., Muranov, K. O., and Kurganov, B. I. (2011) Does the crowded cell-like environment reduce the chaperone-like activity of α-crystallin? *Biochemistry* **50**, 10607-10623

37. Carra, S., Alberti, S., Arrigo, P. A., Benesch, J. L., Benjamin, I. J., Boelens, W., Bartelt-Kirbach, B., Brundel, B. J. J. M., Buchner, J., Bukau, B., Carver, J. A., Ecroyd, H., Emanuelsson, C., Finet, S., Golenhofen, N., Goloubinoff, P., *et al*. (2017) The growing world of small heat shock proteins: from structure to functions. *Cell Stress Chaperones* **22**, 601-611

38. Carver, J. A., Grosas, A. B., Ecroyd, H., and Quinlan, R. A. (2017) The functional roles of the unstructured N- and C-terminal regions in αB-crystallin and other mammalian small heat-shock proteins. *Cell Stress Chaperones* **22**, 627-638

39. Jehle, S., Rajagopal, P., Bardiaux, B., Markovic, S., Kühne, R., Stout, J. R., Higman, V. A., Klevit, R. E., van Rossum, B.-J., and Oschkinat, H. (2010) Solid-state NMR and SAXS studies provide a structural basis for the activation of αB-crystallin oligomers. *Nat. Struct. Mol. Biol.* **17**, 1037-1043

40. Johansen, D., Jeffries, C. M. J., Hammouda, B., Trewhella, J., and Goldenberg, D. P. (2011) Effects of macromolecular crowding on an intrinsically disordered protein characterized by small-angle neutron scattering with contrast matching. *Biophys. J.* **100**, 1120-1128

41. Goldenberg, David P., and Argyle, B. (2014) Minimal effects of macromolecular crowding on an intrinsically disordered protein: a small-angle neutron scattering study. *Biophys. J.* **106**, 905-914

42. Guilbaud, J.-B., and Saiani, A. (2011) Using small angle scattering (SAS) to structurally characterise peptide and protein self-assembled materials. *Chem. Soc. Rev.* **40**, 1200-1210

43. Kataoka, M., and Goto, Y. (1996) X-ray solution scattering studies of protein folding. *Folding Des.* **1**, 107-114

44. Weeks, S. D., Baranova, E. V., Heirbaut, M., Beelen, S., Shkumatov, A. V., Gusev, N. B., and Strelkov, S. V. (2014) Molecular structure and dynamics of the dimeric human small heat shock protein HSPB6. *J. Struct. Biol.* **185**, 342-354

45. Wang, Y., Trewhella, J., and Goldenberg, D. P. (2008) Small-angle X-ray scattering of reduced ribonuclease A: effects of solution conditions and comparisons with a computational model of unfolded proteins. *J. Mol. Biol.* **377**, 1576-1592

46. Johansen, D., Trewhella, J., and Goldenberg, D. P. (2011) Fractal dimension of an intrinsically disordered protein: small-angle X-ray scattering and computational study of the bacteriophage λ N protein. *Protein Sci.* **20**, 1955-1970

47. Beaucage, G. (1996) Small-angle scattering from polymeric mass fractals of arbitrary mass-fractal dimension. *J. Appl. Crystallogr.* **29**, 134-146

48. Hammouda, B. (2010) A new Guinier–Porod model. *J. Appl. Crystallogr.* **43**, 716-719

49. Ecroyd, H., Meehan, S., Horwitz, J., Aquilina, J. A., Benesch, Justin L. P., Robinson, Carol V., Macphee, Cait E., and Carver, John A. (2007) Mimicking phosphorylation of αB-crystallin affects its chaperone activity. *Biochem. J.* **401**, 129-141

50. Srinivas, P., Narahari, A., Petrash, J. M., Swamy, M. J., and Reddy, G. B. (2010) Importance of eye lens α-crystallin heteropolymer with 3:1 αA to αB ratio: stability, aggregation, and modifications. *IUBMB life* **62**, 693-702

51. Yoshimura, Y., Lin, Y., Yagi, H., Lee, Y.-H., Kitayama, H., Sakurai, K., So, M., Ogi, H., Naiki, H., and Goto, Y. (2012) Distinguishing crystal-like amyloid fibrils and glass-like amorphous aggregates from their kinetics of formation. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 14446-14451

52. Braun, N., Zacharias, M., Peschek, J., Kastenmüller, A., Zou, J., Hanzlik, M., Haslbeck, M., Rappsilber, J., Buchner, J., and Weinkauf, S. (2011) Multiple molecular architectures of the eye lens chaperone αB-crystallin elucidated by a triple hybrid approach. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 20491-20496

53. Garvey, M., Ecroyd, H., Ray, N. J., Gerrard, J. A., and Carver, J. A. (2017) Functional amyloid protection in the eye lens: retention of α-crystallin molecular chaperone activity after modification into amyloid fibrils. *Biomolecules* **7**, 1-20

54. Monteith, W. B., and Pielak, G. J. (2014) Residue level quantification of protein stability in living cells. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 11335-11340

55. Inoue, R., Takata, T., Fujii, N., Ishii, K., Uchiyama, S., Sato, N., Oba, Y., Wood, K., Kato, K., Fujii, N., and Sugiyama, M. (2016) New insight into the dynamical system of αB-crystallin oligomers. *Sci. Rep.* **6**, 1-7

56. Carver, J. A., Aquilina, J. A., Truscott, R. J. W., and Ralston, G. B. (1992) Identification by 1H NMR spectroscopy of flexible C-terminal extensions in bovine lens α-crystallin. *FEBS Lett.* **311**, 143-149

57. Durand, D., Vivès, C., Cannella, D., Pérez, J., Pebay-Peyroula, E., Vachette, P., and Fieschi, F. (2010) NADPH oxidase activator p67phox behaves in solution as a multidomain protein with semi-flexible linkers. *J. Struct. Biol.* **169**, 45-53

58. Receveur-Bréchot, V., and Durand, D. (2012) How random are intrinsically disordered proteins? A small angle scattering perspective. *Curr. Protein Pept. Sci.* **13**, 55-75

59. Shahid, S., Hassan, M. I., Islam, A., and Ahmad, F. (2017) Size-dependent studies of macromolecular crowding on the thermodynamic stability, structure and functional activity of proteins: in vitro and in silico approaches. *Biochim. Biophys. Acta, Gen. Subj.* **1861**, 178-197

60. Zhou, H.-X., Rivas, G., and Minton, A. P. (2008) Macromolecular crowding and confinement: biochemical, biophysical, and potential physiological consequences. *Annu. Rev. Biophys.* **37**, 375-397

61. Horwitz, J. (2000) The function of alpha-crystallin in vision. *Semin. Cell Dev. Biol.* **11**, 53-60

62. Carver, J. A., Ecroyd, H., Truscott, R. J. W., Thorn, D. C., and Holt, C. (2018) Proteostasis and the regulation of intra- and extracellular protein aggregation by ATP-independent molecular chaperones: lens α-crystallins and milk caseins. *Acc. Chem. Res.* **51**, 745-752

63. Minton, A. P. (2005) Influence of macromolecular crowding upon the stability and state of association of proteins: predictions and observations. *J. Pharm. Sci.* **94**, 1668-1675

64. Rekas, A., Adda, C. G., Andrew Aquilina, J., Barnham, K. J., Sunde, M., Galatis, D., Williamson, N. A., Masters, C. L., Anders, R. F., Robinson, C. V., Cappai, R., and Carver, J. A. (2004) Interaction of the molecular chaperone αB-crystallin with α-synuclein: effects on amyloid fibril formation and chaperone activity. *J. Mol. Biol.* **340**, 1167-1183

65. Chebotareva, N. A., Eronina, T. B., Sluchanko, N. N., and Kurganov, B. I. (2015) Effect of Ca2+ and Mg2+ ions on oligomeric state and chaperone-like activity of αB-crystallin in crowded media. *Int. J. Biol. Macromol.* **76**, 86-93

66. Roman, S. G., Chebotareva, N. A., and Kurganov, B. I. (2017) Anti-aggregation activity of small heat shock proteins under crowded conditions. *Int. J. Biol. Macromol.* **100**, 97-103

67. Truscott, R. J. W., Chen, Y. C., and Shaw, D. C. (1998) Evidence for the participation of αB-crystallin in human age-related nuclear cataract. *Int. J. Biol. Macromol.* **22**, 321-330

68. Fichna, J. P., Potulska-Chromik, A., Miszta, P., Redowicz, M. J., Kaminska, A. M., Zekanowski, C., and Filipek, S. (2017) A novel dominant D109A CRYAB mutation in a family with myofibrillar myopathy affects αB-crystallin structure. *Biochim. Biophys. Acta, Clin.* **7**, 1-7

69. Hochberg, G. K. A., Ecroyd, H., Liu, C., Cox, D., Cascio, D., Sawaya, M. R., Collier, M. P., Stroud, J., Carver, J. A., Baldwin, A. J., Robinson, C. V., Eisenberg, D. S., Benesch, J. L. P., and Laganowsky, A. (2014) The structured core domain of αB-crystallin can prevent amyloid fibrillation and associated toxicity. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 1562-1570

70. Bova, M. P., Yaron, O., Huang, Q., Ding, L., Haley, D. A., Stewart, P. L., and Horwitz, J. (1999) Mutation R120G in αB-crystallin, which is linked to a desmin-related myopathy, results in an irregular structure and defective chaperone-like function. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 6137-6142

71. Lyon, Y. A., Collier, M. P., Riggs, D. L., Degiacomi, M. T., Benesch, J. L. P., and Julian, R. R. (2019) Structural and functional consequences of age-related isomerization in α-crystallins. *J. Biol. Chem.* **294**, 7546-7555

72. Ellis, R. J. (2001) Macromolecular crowding: obvious but underappreciated. *Trends Biochem. Sci.* **26**, 597-604

73. Chebotareva, N. A., Makeeva, V. F., Bazhina, S. G., Eronina, T. B., Gusev, N. B., and Kurganov, B. I. (2010) Interaction of Hsp27 with native phosphorylase kinase under crowding conditions. *Macromol. Biosci.* **10**, 783-789

74. Sluchanko, N. N., Chebotareva, N. A., and Gusev, N. B. (2015) Quaternary structure of human small heat shock protein HSPB6 (Hsp20) in crowded media modeled by trimethylamine N-oxide (TMAO): effect of protein phosphorylation. *Biochimie* **108**, 68-75

75. Alderson, T. R., Roche, J., Gastall, H. Y., Dias, D. M., Pritišanac, I., Ying, J., Bax, A., Benesch, J. L. P., and Baldwin, A. J. (2019) Local unfolding of the HSP27 monomer regulates chaperone activity. *Nat. Commun.* **10**, 1068-1083

76. Nagaraj, R. H., Oya-Ito, T., Padayatti, P. S., Kumar, R., Mehta, S., West, K., Levison, B., Sun, J., Crabb, J. W., and Padival, A. K. (2003) Enhancement of chaperone function of α-crystallin by methylglyoxal modification. *Biochemistry* **42**, 10746-10755

77. Peschek, J., Braun, N., Rohrberg, J., Back, K. C., Kriehuber, T., Kastenmüller, A., Weinkauf, S., and Buchner, J. (2013) Regulated structural transitions unleash the chaperone activity of αB-crystallin. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 3780-3789

78. Nandi, S. K., Rakete, S., Nahomi, R. B., Michel, C., Dunbar, A., Fritz, K. S., and Nagaraj, R. H. (2019) Succinylation is a gain-of-function modification in human lens αB-crystallin. *Biochemistry* **58**, 1260-1274

79. Ganadu, M. L., Aru, M., Mura, G. M., Coi, A., Mlynarz, P., and Kozlowski, H. (2004) Effects of divalent metal ions on the αB-crystallin chaperone-like activity: spectroscopic evidence for a complex between copper(II) and protein. *J. Inorg. Biochem.* **98**, 1103-1109

80. Mainz, A., Bardiaux, B., Kuppler, F., Multhaup, G., Felli, I. C., Pierattelli, R., and Reif, B. (2012) Structural and mechanistic implications of metal binding in the small heat-shock protein αB-crystallin. *J. Biol. Chem.* **287**, 1128-1138

81. Adachi, M., So, M., Sakurai, K., Kardos, J., and Goto, Y. (2015) Supersaturation-limited and unlimited phase transitions compete to produce the pathway complexity in amyloid fibrillation. *J. Biol. Chem.* **290**, 18134-18145

82. Hartl, F. U., and Hayer-Hartl, M. (2009) Converging concepts of protein folding in vitro and in vivo. *Nat. Struct. Mol. Biol.* **16**, 574-581

83. Meehan, S., Berry, Y., Luisi, B., Dobson, C. M., Carver, J. A., and MacPhee, C. E. (2004) Amyloid fibril formation by lens crystallin proteins and its implications for cataract formation. *J. Biol. Chem.* **279**, 3413-3419

84. Meehan, S., Knowles, T. P. J., Baldwin, A. J., Smith, J. F., Squires, A. M., Clements, P., Treweek, T. M., Ecroyd, H., Tartaglia, G. G., Vendruscolo, M., MacPhee, C. E., Dobson, C. M., and Carver, J. A. (2007) Characterisation of amyloid fibril formation by small heat-shock chaperone proteins human αA-, αB- and R120G αB-crystallins. *J. Mol. Biol.* **372**, 470-484

85. Ecroyd, H., and Carver, J. A. (2008) Crystallin proteins and amyloid fibrils. *Cell. Mol. Life Sci.* **66**, 62-81

86. Zhang, T. O., Alperstein, A. M., and Zanni, M. T. (2017) Amyloid β-sheet secondary structure identified in UV-induced cataracts of porcine lenses using 2D IR spectroscopy. *J. Mol. Biol.* **429**, 1705-1721

87. Alperstein, A. M., Ostrander, J. S., Zhang, T. O., and Zanni, M. T. (2019) Amyloid found in human cataracts with two-dimensional infrared spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 6602-6607

88. Sun, T.-X., and Liang, J. J.-N. (1998) Intermolecular exchange and stabilization of recombinant Human αA- and αB-crystallin. *J. Biol. Chem.* **273**, 286-290

89. van Boekel, M. A. M., de Lange, F., de Grip, W. J., and de Jong, W. W. (1999) Eye lens αA- and αB-crystallin: complex stability versus chaperone-like activity. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **1434**, 114-123

90. Liang, J., Sun, T.-X., and Akhtar, N. J. (2000) Heat-induced conformational change of human lens recombinant alphaA-and alphaB-crystallins. *Mol. Vis.* **6**, 10-14

91. Ray, N. J., Hall, D., and Carver, J. A. (2017) A structural and functional study of Gln147 deamidation in αA-crystallin, a site of modification in human cataract. *Exp. Eye. Res.* **161**, 163-173

92. Liu, Y., Zhang, X., Luo, L., Wu, M., Zeng, R., Cheng, G., Hu, B., Liu, B., Liang, J. J., and Shang, F. (2006) A novel αB-crystallin mutation associated with autosomal dominant congenital lamellar cataract. *Invest. Ophth. Vis. Sci.* **47**, 1069-1075

93. Hayes, V. H., Devlin, G., and Quinlan, R. A. (2008) Truncation of αB-crystallin by the myopathy-causing Q151X mutation significantly destabilizes the protein leading to aggregate formation in transfected cells. *J. Biol. Chem.* **283**, 10500-10512

94. Validandi, V., Sudhakar Reddy, V., Srinivas, P. N. B. S., Mueller, N. H., Bhagyalaxmi, S. G., Padma, T., Mark Petrash, J., and Bhanuprakash Reddy, G. (2011) Temperature-dependent structural and functional properties of a mutant (F71L) αA-crystallin: molecular basis for early onset of age-related cataract. *FEBS Lett.* **585**, 3884-3889

95. Makley, L. N., McMenimen, K. A., DeVree, B. T., Goldman, J. W., McGlasson, B. N., Rajagopal, P., Dunyak, B. M., McQuade, T. J., Thompson, A. D., Sunahara, R., Klevit, R. E., Andley, U. P., and Gestwicki, J. E. (2015) Pharmacological chaperone for α-crystallin partially restores transparency in cataract models. *Science* **350**, 674-677

96. Fu, L., and Liang, J. J. N. (2003) Enhanced stability of αB-crystallin in the presence of small heat shock protein Hsp27. *Biochem. Biophys. Res. Commun.* **302**, 710-714

97. Hussein, R. M., Benjamin, I. J., and Kampinga, H. H. (2015) Rescue of αB-crystallin (HSPB5) mutants associated protein aggregation by co-expression of HSPB5 partners. *PLOS ONE* **10**, e0126761

98. Björk, I. (1970) Studies on γ-crystallin from calf lens: III. Comparison of the main protein components by peptide mapping. *Exp. Eye. Res.* **9**, 152-157

99. Horwitz, J., Huang, Q.-L., Ding, L., and Bova, M. P. (1998) Lens α-crystallin: chaperone-like properties. *Methods Enzymol*. **290**, 365-383

100. Duff, A. P., Wilde, K. L., Rekas, A., Lake, V., and Holden, P. J. (2015) Chapter one - robust high-yield methodologies for 2H and 2H/15N/13C labeling of proteins for structural investigations using neutron scattering and NMR. *Methods Enzymol.* **565**, 3-25

101. Chiou, S.-H., Azari, P., Himmel, M. E., and Squire, P. G. (1979) Isolation and physical characterization of bovine lens crystallins. *Int. J. Pept. Protein Res.* **13**, 409-417

102. Liang, J. N., Bose, S. K., and Chakrabarti, B. (1985) Age-related changes in protein conformation in bovine lens crystallins. *Exp. Eye. Res.* **40**, 461-469

103. Wood, K., Mata, J. P., Garvey, C. J., Wu, C.-M., Hamilton, W. A., Abbeywick, P., Bartlett, D., Bartsch, F., Baxter, P., Booth, N., Brown, W., Christoforidis, J., Clowes, D., d'Adam, T., Darmann, F., Deura, M., Harrison, S., *et al*. (2018) QUOKKA, the pinhole small-angle neutron scattering instrument at the OPAL Research Reactor, Australia: design, performance, operation and scientific highlights. *J. Appl. Crystallogr.* **51**, 294-314

104. Kline, S. (2006) Reduction and analysis of SANS and USANS data using IGOR Pro. *J. Appl. Crystallogr.* **39**, 895-900

105. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H., and Svergun, D. I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **36**, 1277-1282

106. Svergun, D. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **25**, 495-503

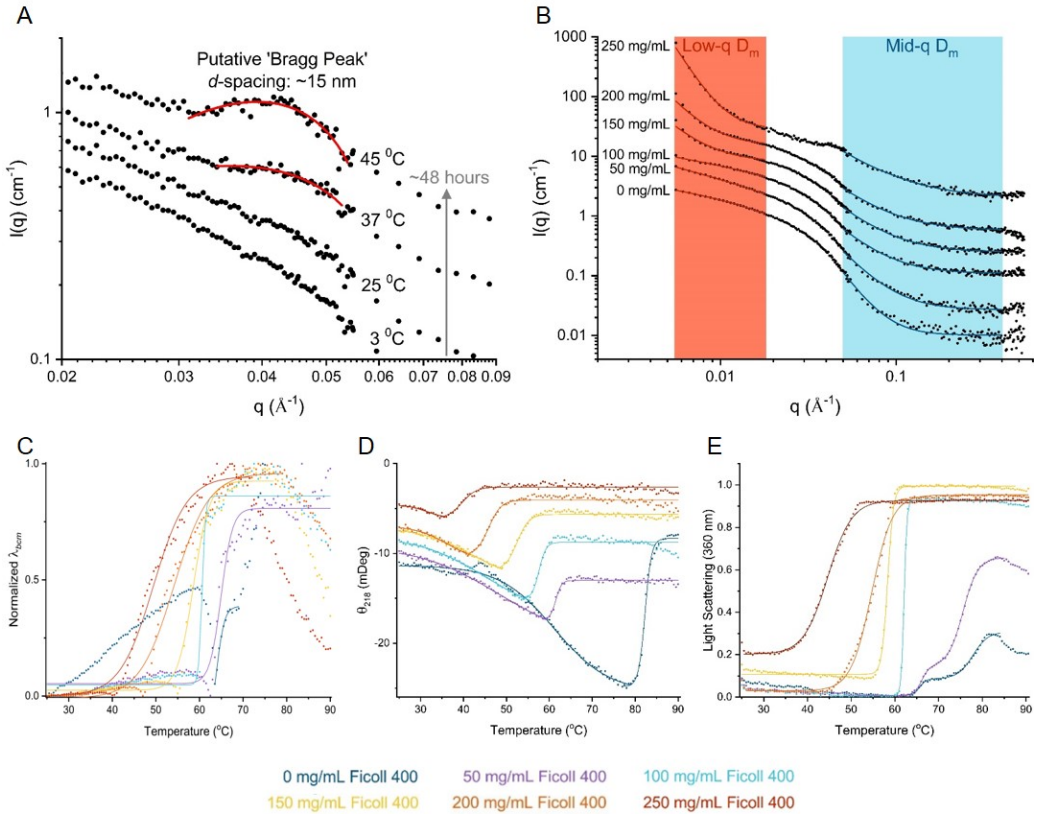
107. Doucet, M., Cho, J. H., Alina, G., Bakker, J., Bouwman, W., Butler, P., Campbell, K., Gonzales, M., Heenan, R., Jackson, A., Juhas, P., King, S., Kienzle, P., Krzywon, J., Markvardsen, A., Nielsen, T., O'Driscoll, L., *et al*. (2018) SasView. Version 4.2.0 Ed.

108. Cox, D., Selig, E., Griffin, M. D. W., Carver, J. A., and Ecroyd, H. (2016) Small heat-shock proteins prevent α-synuclein aggregation via transient interactions and their efficacy is affected by the rate of aggregation. *J. Biol. Chem.* **291**, 22618-2262

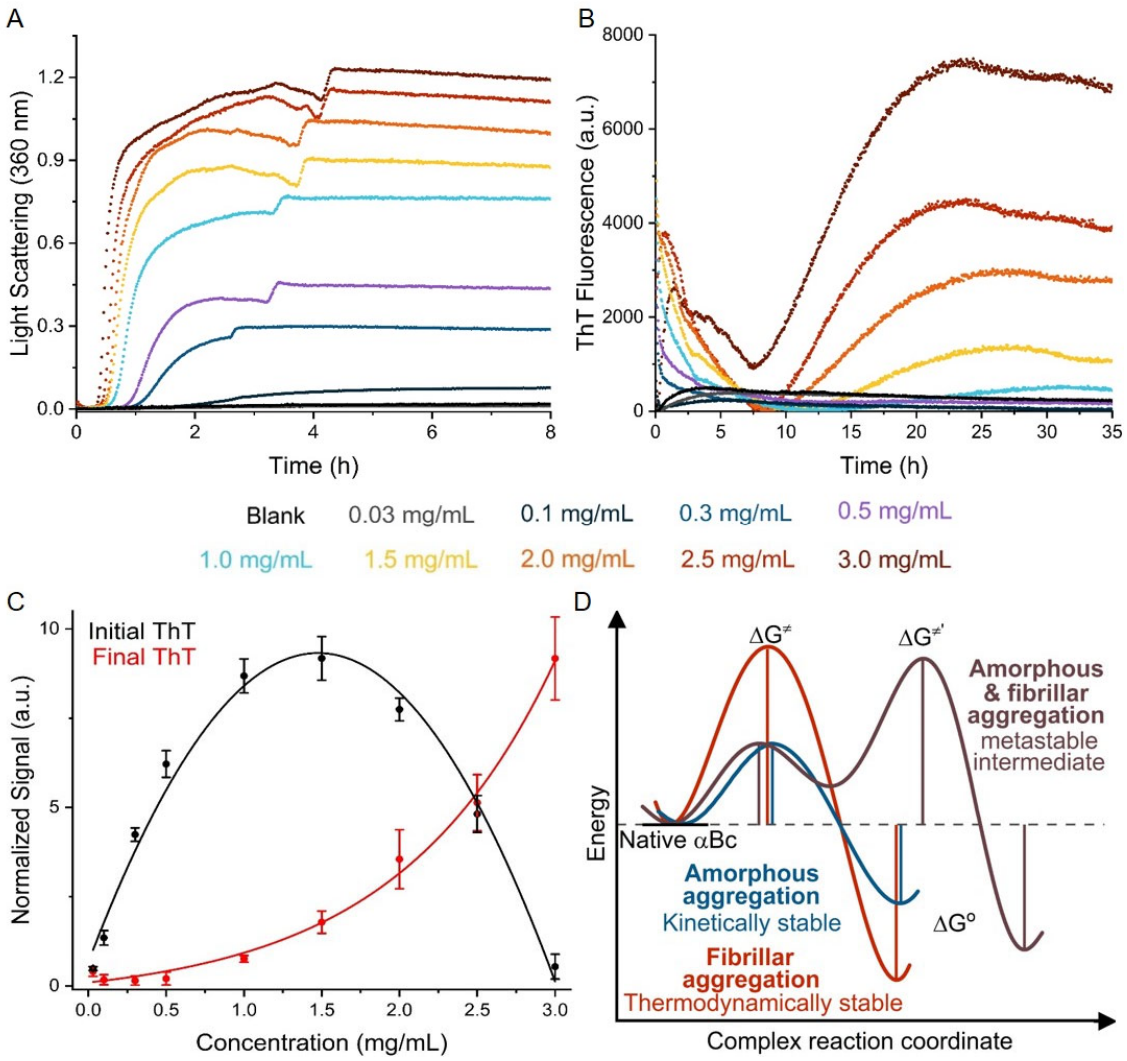
SUPPLEMENTAL INFORMATION

The aggregation of αB-crystallin under crowding conditions is prevented by αA-crystallin: Implications for α-crystallin stability and lens transparency

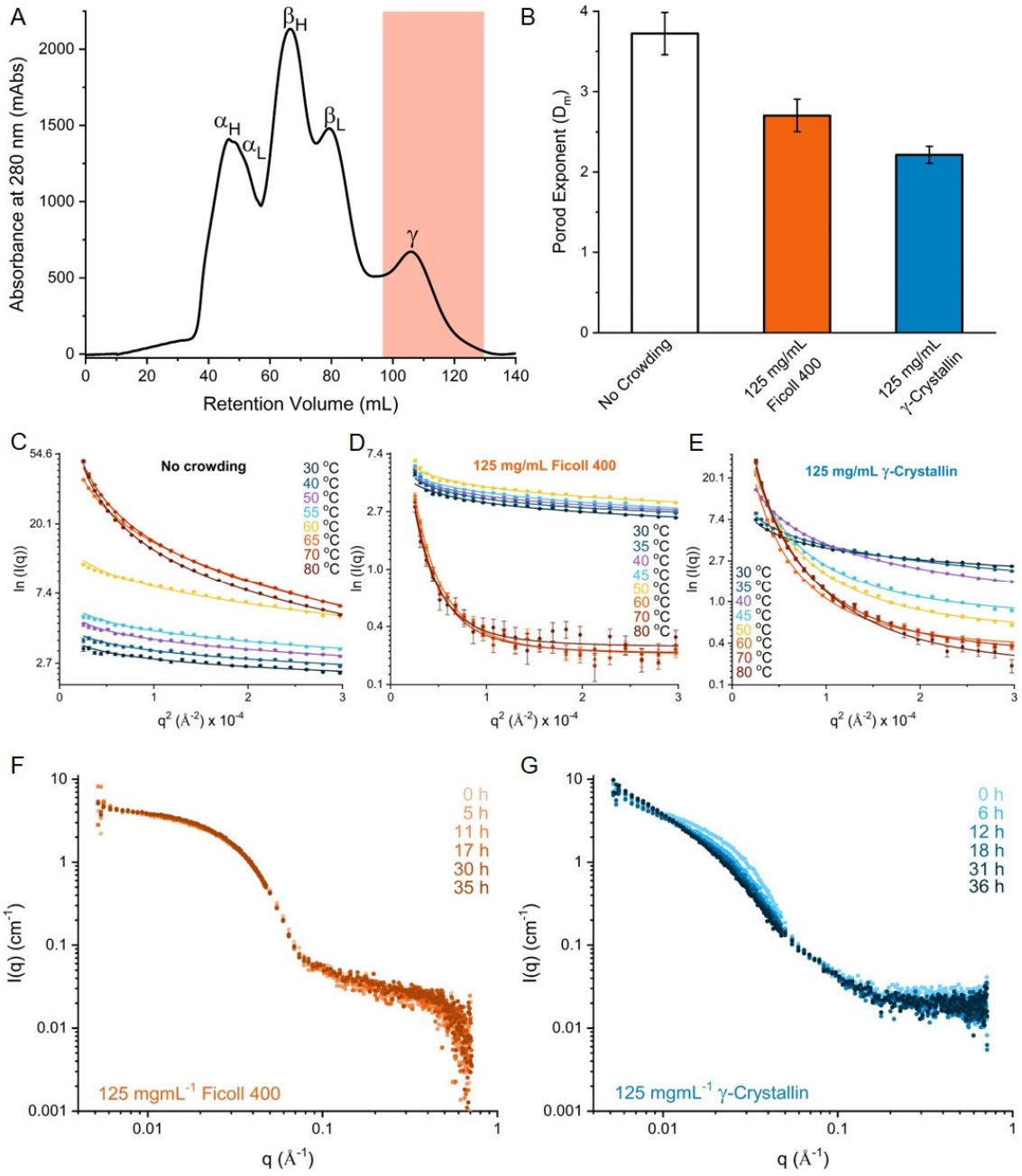
Aidan B. Grosas, Agata Rekas, Jitendra P. Mata, David C. Thorn and John A. Carver



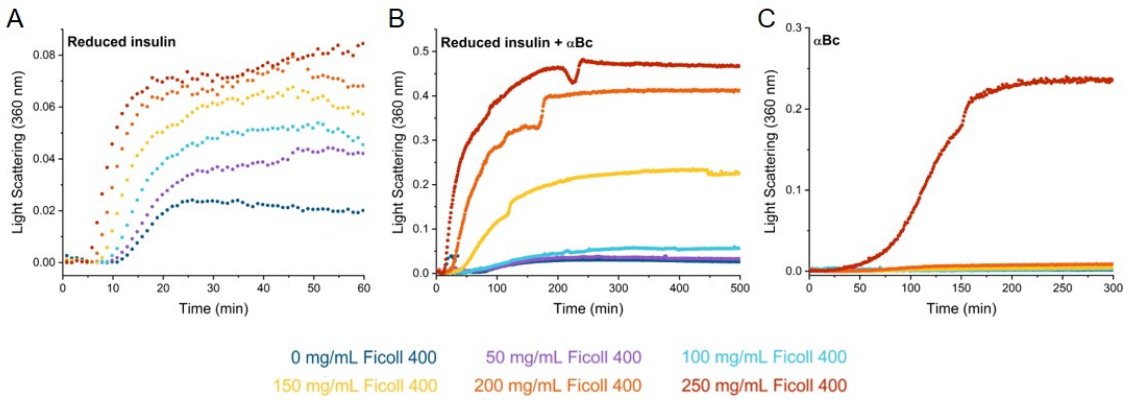
Supplemental Figure 1. Fitting of SANS and spectroscopic thermal stability data with increasing Ficoll 400 concentrations. *A*, expanded SANS scattering plot of 2.5 mg/mL αBc in 250 mg/mL Ficoll 400 at four different temperatures collected sequentially over ~ 48 hours. A putative Bragg peak at 37 °C and 45 °C is fit with a Gaussian function (*red*, *solid line*) giving a spacing of ~ 15 nm. *B*, scattering plot of 2.5 mg/mL αBc in increasing Ficoll 400 concentrations at 37 °C using 28 % D_2O. Scattering profiles are displaced along the *y*-axis with increasing Ficoll 400 concentration for comparison. The area of the mid-q (*blue* (0.05 $Å^{-1} < q < 0.40$ $Å^{-1}$)) and low-q (*red* (0.006 $Å^{-1} < q < 0.018$ $Å^{-1}$)) D_m is highlighted and fits to the power-law function are the solid lines through the scattering data at those respective mid-q and low-q limits. *C*-*E*, Thermal stability of 0.3 mg/mL αBc in increasing Ficoll 400 concentrations (colour key below) as monitored by *C*, tryptophan fluorescence analysed using barycentric mean fluorescence (λ_{bcm}), *D*, ellipticity at 218 nm (Θ_{218}) from circular dichroism and, *E*, light scattering at 360 nm. Data are presented as scatter and fits are solid lines. Thermal unfolding curves in *D* show two changes in Θ exemplified by negatively and positively sloped single Boltzmann sigmoidal functions. Thermal aggregation curves for 0 and 50 mg/mL Ficoll 400 in *E* are the best fitted to a double Boltzmann sigmoidal function. All other transitions were fitted to a single Boltzmann sigmoidal function.



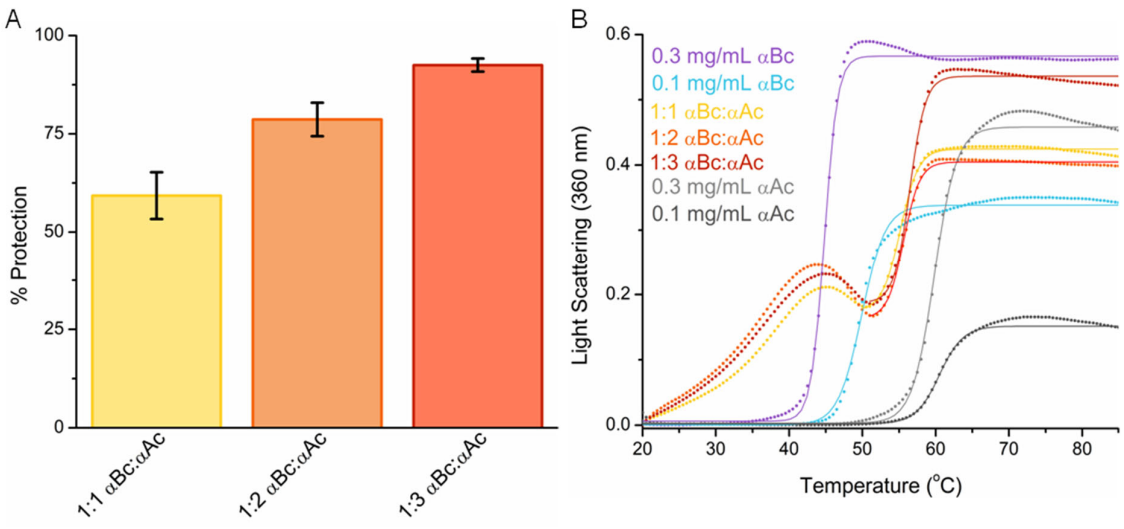
Supplemental Figure 2. Light scattering and ThT fluorescence kinetics of αBc aggregates formed in crowded conditions in the presence of 250 mg/mL Ficoll 400. *A*, representative light scattering (360 nm) and *B*, ThT fluorescence kinetics for αBc from 0.03 – 3.00 mg/mL in 250 mg/mL Ficoll 400 at 37 °C. The colours of the traces are related to the concentrations of αBc used and are given below panels *A* and *B*. *C*, average initial (*black*) and final (*red*) ThT fluorescence from three independent repeats of the ThT assay in panel *B*. Errors are given as the standard deviation of three independent repeats. Initial ThT values versus αBc concentration are fitted to a Gaussian function (R^2 = 0.99) while final ThT values are the same as those in Fig. 2, *H*, used here for comparison. This demonstrates that as αBc concentration increases at 250 mg/mL Ficoll 400 the initial and final ThT values are sensitive to two different structural aspects of the protein. *D*, conceptual reaction coordinate diagram to explain the distinct kinetics, via the first ($\Delta G^{\neq}$) or second ($\Delta G^{\neq'}$) activation energy, and differences in abundance with increasing concentration, via the change in free energy (ΔG^0), of amorphous and fibrillar aggregation. Three potential reaction pathways include distinct amorphous (*blue*) and fibrillar (*red*) aggregation pathways stemming from a native pool of native αBc or a single amorphous and fibrillar aggregation pathway (*purple*) with a metastable intermediate (likely a form of amorphous aggregate) implying a transition from an amorphous to a fibrillar aggregate.



Supplemental Figure 3. SANS data analysis of αBc using bovine γ-crystallin as a crowding agent. *A*, size-exclusion separation of bovine lens homogenate showing the γ-crystallin fraction extracted (*shaded*) and used as a biologically relevant crowding agent. *B*, Porod exponent of 2.5 mg/mL deuterated αBc at 30 °C under various crowding conditions as indicated. *C-E*, Guinier plots (*scatter*) and power-law fits (*solid line*) of 2.5 mg/mL deuterated αBc at temperatures from 30 – 80 °C under conditions of *C*, no crowding, *D*, 125 mg/mL Ficoll 450, and *E*, 125 mg/mL γ-crystallin. The 'low-q D_m' value from the power-law fit is used to indicate the presence and extent of αBc aggregation. *F and G*, Scattering plots of 2.5 mg/mL deuterated αBc at 37 °C from 0 – 36 hours for *F*, 125 mg/mL Ficoll 400, and *G*, 125 mg/mL γ-crystallin.



Supplemental Figure 4. Assessing the chaperone ability of αBc against reduced insulin aggregation with increasing Ficoll 400 concentrations. Representative light scattering curves of *A*, reduced insulin, *B*, reduced insulin with αBc in a 1:2 w:w ratio and *C*, αBc alone as monitored at 360 nm in increasing Ficoll 400 concentrations from 0 – 250 mg/mL. Ficoll 400 concentration is indicated by the color as per the key below the figures.



Supplemental Figure 5. Chaperone effectiveness and thermal stability for αBc, αAc and hetero-oligomers in 250 mg/mL of Ficoll 400. *A*, percentage protection for αAc's chaperone activity on αBc calculated from data in Fig. 5*A*. *B*, thermal aggregation curves for αBc, αAc and different ratios of hetero-oligomers in 250 mg/mL Ficoll 400. The initial increase in light scattering for 1:1, 1:2, and 1:3 αBc:αAc is likely due to a significant increase in size of the α-crystallin oligomer with temperature. This size increase is then reveresed, likely by αAc chaperone action, before finally completely aggregating at higher temperatures. The T_{agg} for 1:1, 1:2, and 1:3 αBc:αAc is measured as the transition after the initial increase and decrease in light scattering. All curves are an average of three independent repeats.

CHAPTER 4:

STUDYING CONFORMATIONAL HETEROGENEITY USING CAPILLARY ELECTROPHORESIS

"Real courage is when you know you're licked before you begin, but you begin anyway and see it through no matter what. You rarely win, but sometimes you do."

– Harper Lee

DECLARATION

The following manuscript is formatted in the style of *Analytical Chemistry (ACS)*.

All experimental work was carried out by the author, except where otherwise stated below, under the supervision of Professor John Carver and Dr Patrice Castignolles.

The contributions of other authors are as follows: The author conceived the study. Dr Patrice Castignolles and Dr Marianne Gaborieau provided advice on capillary electrophoresis experiments and data treatment. Ms Mar-dean Du Plessis and Dr Joel Thevarajah assisted in the collection of capillary electrophoresis data and its analysis. The author wrote the manuscript with input from Ms Mar-dean Du Plessis, Dr Joel Thevarajah, and Professor John Carver.

Using Capillary Electrophoresis to Investigate Protein Conformational Heterogeneity: a Comparative Study between the Distribution of Electrophoretic Mobilities and Molar Masses

Aidan B. Grosas,[1,*] Mar-dean Du Plessis,[2,3] Joel J. Thevarajah,[2,3,4] Marianne Gaborieau,[2,3] John A. Carver,[1] and Patrice Castignolles[2,*]

1 Research School of Chemistry, Australian National University, Acton, ACT 2601, Australia

2 Western Sydney University, School of Science and Health, Australian Centre for Research on Separation Science (ACROSS), Parramatta, NSW 2150, Australia

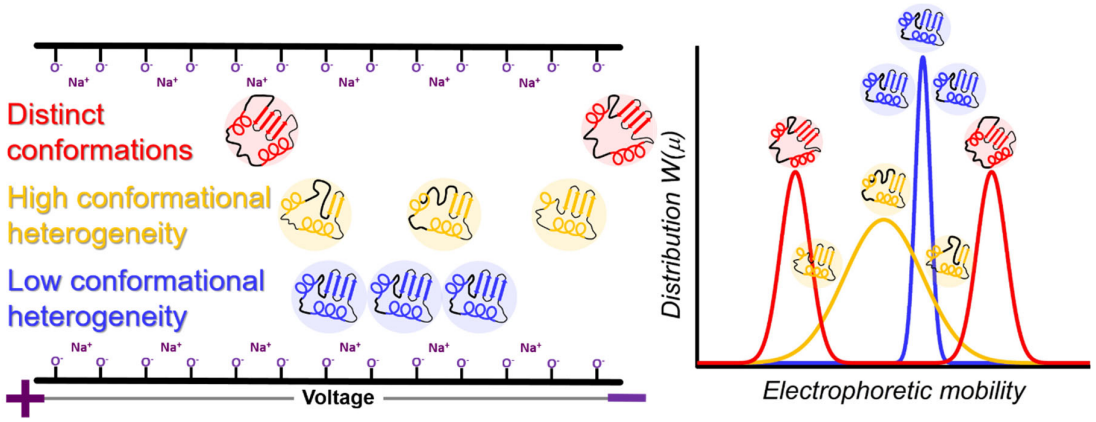
3 Western Sydney University, Molecular Medicine Research Group, School of Science and Health, Parramatta, NSW 2150, Australia

4 Climate Change Cluster (C3), University of Technology Sydney, Ultimo, New South Wales, Australia

Abbreviations: AA3, oligo(sodium acrylate); BGE, background electrolyte; BLA, bovine α-lactalbumin; BSA, bovine serum albumin; CD, circular dichroism; CE, free-solution capillary electrophoresis; CE-CC, capillary electrophoresis in the critical conditions; CE-MS, capillary electrophoresis with mass spectrometry detection; CE-UV-IM-MS, capillary electrophoresis with ultra violet, ion mobility, and mass spectrometry detection; DMSO, dimethyl sulfoxide; dRI, differential refractive index; DTT, dithiothreitol; $D(W(\mu))$, electrophoretic mobility dispersity; EDTA, ethylenediaminetetraacetic acid; EOF, electroosmotic flow; IM-MS, ion mobility mass spectrometry; M_w, weight-average molar mass; PM, pressure mobilization; RI, refractive index; SEC-MALS, size-exclusion chromatography with multi-angle light scattering detection; $W(M)$, distribution of molar masses; $W(\mu)$, distribution of electrophoretic mobilities; YADH, yeast alcohol dehydrogenase; $Đ$, molar mass dispersity; μ, electrophoretic mobility; μ_w, weight-average electrophoretic mobility

Keywords: Capillary electrophoresis, heterogeneity, protein, conformation, SEC-MALS, distribution, dispersity, structure, bovine serum albumin, alcohol dehydrogenase, bovine α-lactalbumin

ABSTRACT: Separation of pure protein using free-solution capillary electrophoresis (CE) can resolve distinct protein conformers and provide insight into conformational heterogeneity. However, the method is seldom used for protein conformational characterization. To clearly demonstrate the application of CE toward the characterization of distinct protein conformers and conformational heterogeneity, we analyzed three well-characterized proteins, bovine serum albumin (BSA), yeast alcohol dehydrogenase (YADH), and bovine α–lactalbumin (BLA) under conditions that led to their structural alteration. A distribution of electrophoretic mobilities was obtained from CE and was juxtaposed with the conventional distribution of molar masses obtained by size-exclusion chromatography coupled to light scattering (SEC-MALS). Oligomers as well as two monomeric conformations of BSA were separated by CE and the conformational heterogeneity of two different oligomeric preparations was compared. Analysis of YADH resolved two distinct monomeric conformers as well as multiple tetrameric species while changes to the protein's conformational heterogeneity could be followed upon removal of its intrinsic zinc ions and reduction of its disulfide bonds. While SEC-MALS profiles were indistinguishable for the apo- and holo-forms of BLA, CE resolved both species and also measured changes to the protein's heterogeneity upon metal ion binding and removal. Further, comparison of these structurally unrelated proteins revealed novel insights into relative differences in molar mass and conformational heterogeneity. Overall, this analysis has improved our understanding of how CE separates protein oligomers and distinct conformers and their relationship to conformational heterogeneity. CE can provide additional and complementary structural information on proteins with varying degrees of conformational complexity compared to other traditional techniques.



INTRODUCTION

Protein structural characterization provides fundamental insight into the physical and chemical properties that underlie conformational features and functions of these essential biological macromolecules. Free-solution capillary electrophoresis (CE) is one separation-based characterization technique that has been used extensively in proteomics[1,2], but has only recently begun to gain favor in relation to intact protein characterization. CE in the critical conditions (CE-CC) transcends traditional hydrodynamic based separation methods, i.e. separation dependent on hydrodynamic size, and instead separates on the basis of conformational and/or compositional feature(s) of a polyelectrolyte.[3] This molar mass independent separation method has been repeatedly demonstrated in the literature for various polyelectrolytes including denatured protein-SDS complexes,[4,5] DNA[6] and synthetic polymers.[7] It is clear that above a certain molar mass (> ~10,000 $g \cdot mol^{-1}$)[4] or degree of polymerization (> ~120 N),[7] unstructured charged macromolecules will have minimal variation in their electrophoretic mobility (μ), which is proportional to their velocity within an electric field.[8] For intact proteins, this behavior implies that any differences in μ is likely due to conformational and/or compositional differences based on a relationship between shape, size, effective charge, counter ion condensation and hydration that exist on a timescale amenable to CE separation.[9,10] To this end, the separation mechanism of CE-CC can be more aptly described as being effected by charge and electrostatic friction, sensitive to conformational differences rather than only hydrodynamic size.

The complex relationship between charge and conformation in CE-CC has been exploited to characterize various protein conformations. Folding/unfolding studies using either chemical denaturant or temperature have shown a significant distinction in the μ between folded, partially unfolded and unfolded populations.[9-14] Studies have interfaced CE with intact protein mass spectrometry (CE-MS)[15] to unambiguously assign different peaks, confirming the separation of conformers or post-translationally modified isoforms.[16] A folding transition has also been monitored using CE and has been validated through comparison with ion-mobility MS (IM-MS) which is sensitive to conformational differences via changes in protein cross-sectional area.[14] Both open and closed conformations of a protein in the presence of ligands have been studied with kinetic CE using UV-detection[17] and subsequently through coupling CE separation with IM-MS (CE-UV-IM-MS) to elucidate a more

definitive link between enzyme conformation and function.[18] While powerful, it is important to note that CE-MS faces significant challenges. It is difficult to select a suitable buffer system that can maintain a protein's native conformation, is amenable to CE separation, and can provide sufficient ionization capacity. This can potentially lead to a compromise that results in a loss of resolution.[16,19] Furthermore, while these studies focus on the separation and assignment of distinct conformations, it has also been recognized that the inherent broadness of a CE peak contains information related to multiple conformations with a similar μ.[13,17] Thus, the ability to measure the CE peak broadness provides previously unstudied information related to the conformational heterogeneity of a protein. However, the methodology that allows this measurement in CE was only recently developed.

CE data are acquired in the form of detector signal against migration time which can be transformed into a distribution of electrophoretic mobilities ($W(\mu)$).[20] This compensates for variations in electric field strength, electroosmotic flow (EOF) and peak area due to the different relative velocities of analytes in the detection window. Subsequently, due to the unique aforementioned separation mechanism of CE-CC, the $W(\mu)$ can be reflective of a distribution of conformational and/or compositional feature(s). Recently, the heterogeneity of the degree of acetylation of chitosan, a compositional component, and the heterogeneity of branching of poly(sodium acrylate), a conformational feature, have been quantified.[3] However, this novel methodology has not been applied to proteins. When using CE-CC to characterize a protein devoid of significant post-translational modifications, any variance in the $W(\mu)$ should be reflective of differences in that protein's conformation. This effectively allows quantification of the heterogeneity of conformation via a dispersity value ($D(W(\mu))$) calculated from the protein's $W(\mu)$.

Proteins can exhibit multiple forms of heterogeneity, namely with regard to molar mass, size, conformation and composition.[21] Traditional techniques are limited in the information they can provide on highly heterogeneous proteins. Currently the most common method to investigate heterogeneity is size exclusion chromatography coupled to a multi-angle light scattering detector (SEC-MALS). This technique yields the in-solution distribution of molar masses ($W(M)$) and the dispersity value ($Đ$) thereof, as well as the number- and weight-average (M_w) molar mass. While SEC-MALS probes the heterogeneity in molar mass, no method has been able to provide a direct in-solution measure of the

heterogeneity of conformational and/or compositional distributions. A better understanding of protein conformational heterogeneity inferred through changes in the $D(W(\mu))$ from a CE experiment could provide information on shifts in conformational equilibrium related to enzyme or chaperone function, structural plasticity, and thermal stability or lability.

Herein, we compare at physiological p*H* the respective $W(\mu)$ from CE with the $W(M)$ acquired from SEC-MALS for three structurally well-characterized proteins: bovine serum albumin (BSA), alcohol dehydrogenase from *Saccharomyces cerevisiae* (YADH) and α-lactalbumin from bovine milk (BLA) under various structurally altering treatments. Juxtaposition of these distributions for each protein demonstrates the striking complementary and orthogonal features of the two characterization methodologies. The relationship between dispersity values and conformational heterogeneity is examined. Also, correlations between dispersity and corresponding weight-average electrophoretic mobility (μ_w) and M_w, respectively, are used to better understand how protein oligomerization affects CE separation and conformational heterogeneity. The data are interpreted within the framework of the extensive existing literature regarding the conformational features of these proteins. Finally, some new conformational insights are derived from the data.

EXPERIMENTAL SECTION

Materials. All lyophilized protein samples including albumin from bovine serum (oligomeric BSA (A7906)), albumin, monomer bovine (monomeric BSA (A1900)), alcohol dehydrogenase from *Saccharomyces cerevisiae* (YADH (A7011 – 75KU)) and α-lactalbumin from bovine milk (BLA (Ca^{2+} depleted (dep)) (L6010)) were purchased from Sigma and used without further purification. All other reagents including ethylenediaminetetraacetic acid (EDTA) and dithiothreitol (DTT), boric acid, sodium borate, $CaCl_2$ and $MgCl_2$, unless otherwise stated, were purchased from Sigma. All solutions were prepared using MilliQ water (18 MΩ cm^{-1}, 0.22 μm membrane filter). A 200 mM stock solution of sodium phosphate buffer (NaPi) was produced by titrating acidic NaPi with NaOH to pH 7.4. It was then diluted to a working concentration of 5 mM to be used as the CE background electrolyte (BGE) and the SEC-MALS eluent for BSA and YADH. BLA experiments were conducted in 20 mM Tris(hydroxymethyl)methylamine-HCl (Tris-HCl) buffer, produced by titrating Tris base with HCl to

pH 7.4. Protein concentrations were quantified using the protein's extinction coefficient ($E^{1\%}$) value provided in the manufacturer's product documentation and measured using an ultra-low volume spectrophotometer (Thermo Fisher Scientific Inc.).

Determination of d*n*/d*c* values. For each protein sample, a concentration series consisting of 10 points between 0.5 $g \cdot L^{-1}$ and 5.0 $g \cdot L^{-1}$ was prepared from a stock solution. The relevant SEC-MALS eluent was sonicated for 5 minutes to remove dissolved gasses and then filtered (0.22 μm, Chromfilter) before use. The refractive index (RI) of the protein solution at each concentration was determined using an ultra-low volume refractometer (J357, Rudolph Research Analytical) at 25 °C in the aforementioned buffer. The data were plotted as differential RI (dRI) vs. concentration ($g \cdot mL^{-1}$) and the d*n*/d*c* value was taken as the slope of the linear fit.

Capillary Electrophoresis and Pressure Mobilization. Both CE and pressure mobilization (PM) experiments were performed on an Agilent CE 7100 instrument (Agilent Technologies) using an 80.0 cm total length (71.5 cm effective length) fused silica capillary (Polymicro, USA) with an internal diameter of 50 μm. Capillary pre-treatment consisted of a 10 minute flush with fresh 1 M NaOH, 5 minute flush with 0.1 M NaOH, 5 minute pre-treatment flush with Milli-Q water and a 5 minute flush with the BGE. This pre-treatment was also used before each new protein was analyzed. Between each individual experiment, the capillary was flushed for 5 minutes with the BGE. The capillary post-treatment was a flush for 5 minutes with 1 M NaOH, 10 minutes with Milli-Q water and 10 minutes with air for storage. For both CE and PM, BSA and YADH were prepared to a concentration of 1 $g \cdot L^{-1}$ while BLA was used at 0.5 $g \cdot L^{-1}$. Protein treatments (Table S1) were allowed to react for one hour before measurement. Protein samples were injected hydrodynamically into the capillary using 30 mbar of pressure for 10 seconds. CE separation was conducted at 30 kV while PM experiments were completed using 50 mbar of pressure. Oligo(sodium acrylate) (AA3) in 25 mM sodium borate buffer (pH 9.3) was used to validate the capillary (Figure S2).[22] Dimethyl sulfoxide (DMSO) was used as a mobility marker at 0.2 % (v/v). All data were recorded at 191 nm (bandwidth of 2 nm) and treated as previously described[3,20] using Origin 8.5.

Size-exclusion chromatography with Multi-Angle Light Scattering. SEC–MALS experiments were performed using a DAWN HELEOS 8 (Wyatt) laser light scattering device and an Optilab rEX (Wyatt) RI detector connected in series. Both detectors were set at 25 °C and the eluent was degassed online with a Waters In-Line Degasser AF. Protein samples were prepared at 3 $g \cdot L^{-1}$. Protein treatments (Table S1) were allowed to react for one hour before measurement. Sample separation was achieved using a Superdex 200 10/300 GL SEC column (GE Healthcare) for BSA and YADH while a Superdex 75 10/300 GL SEC column (GE Healthcare) was used for BLA. The columns were equilibrated using the same sample buffers as for CE experiments prior to use. Samples were injected onto the column via a 100 µL injection loop. Normalization of the MALS detectors was achieved using the monomeric fraction of BSA at 10 $g \cdot L^{-1}$ in the aforementioned equilibration buffer. Data acquisition and processing were completed using ASTRA (5.3.4) (Wyatt).

Circular Dichroism Spectroscopy (CD). CD experiments were performed using an Applied Photophysics Chirascan spectrophotometer attached to a Quantum Northwest TC 125 PELTIER temperature controller. Oligomeric and monomeric BSA was dissolved in 5 mM sodium phosphate, pH 7.0, at a concentration of 0.3 $g \cdot L^{-1}$. The spectrum was acquired in a 0.1 cm pathlength quartz cuvette from 180 – 250 nm at 25 °C. All spectra were acquired with a step width of 1 nm for 4 s and three repeats were taken. These parameters were used for three independent experiments which were averaged to produce the final spectrum.

RESULTS AND DISCUSSION

Analytical considerations. Initially, PM experiments were used to ensure there was no significant adsorption to the capillary wall that could affect electrophoretic separation and peak shape.[23] PM experiments were repeated five times for each protein along with each relevant treatment and yielded largely symmetrical Gaussian peaks (Figure S1 A-C). These results qualitatively indicate minimal adsorption between the proteins and the capillary wall. Adjusted R^2 values were obtained through fitting the peaks with a Gaussian function. The values, 0.951 - 0.993, indicate minimal peak asymmetry and thus confirmed adsorption was negligible across the experimental series (Table S2).

The separation efficiency of the capillary was tested using a complex synthetic polymer mixture, AA3. The AA3 peaks could easily be assigned from their respective μ values using the literature[24] (Figure S2) thereby indicating the effective separation of the capillary. Optimization of the separation conditions is a pivotal part of obtaining unambiguous separation data. While maintaining p*H* at the physiologically relevant value of 7.4, it was found that a low salt buffer and an 80 cm capillary were the ideal conditions for improved resolution. A comparison of separations at 40 cm and 80 cm capillary lengths (Figure S3) using the same injection volume illustrates the stark difference a longer capillary and sample overloading can make to the resolution of the separation. Transformation of raw CE data from migration time to μ requires taking into account the migration of a neutral molecule, DMSO, which is used as an EOF marker. This transformation accounts for variations between injections and ensures accuracy of the sample μ value. Post transformation, all relevant proteins and treatments were overlaid and visually inspected to ensure correct alignment of the EOF peak (Figure S4 A-C).

The SEC-MALS measurements require the determination of a protein's d*n*/d*c* value for accurate molar mass calculations. Each protein was dissolved in the relevant degassed buffer and the d*n*/d*c* value obtained from a linear fit ($R^2 > 0.99$) of dRI against a concentration series (Figure S5 A-C). Finally, the SEC chromatograms were examined to confirm that the treatments (Table S1) used had not greatly affected the retention time of the modified protein (Figure S6 A-C).

Conformational Heterogeneity and CE Separation of BSA Oligomers. BSA can form various oligomeric species (e.g. monomer, dimer, trimer etc.) but can also be purified to a high percentage of monomer only.[25,26] To understand how CE separates oligomeric proteins and assess the impact this has on conformational heterogeneity, we analyzed BSA in its oligomeric and monomeric forms and compared the results to those obtained by SEC-MALS.

The *W*(*M*) of the oligomeric BSA preparation shows a broad tailing distribution with unresolved oligomers at higher molar mass (Figure 1A, *black*) which are not apparent for monomeric BSA (Figure 1A, *red*), as is expected since the latter preparation is largely monomeric. Single peaks at ~63,000 g·mol^{-1} and ~120,000 g·mol^{-1} (peak max) correspond to the approximate molar masses for BSA monomer and dimer, respectively (Figure 1A). The *W*(μ) of oligomeric BSA (Figure 1B, *black*)

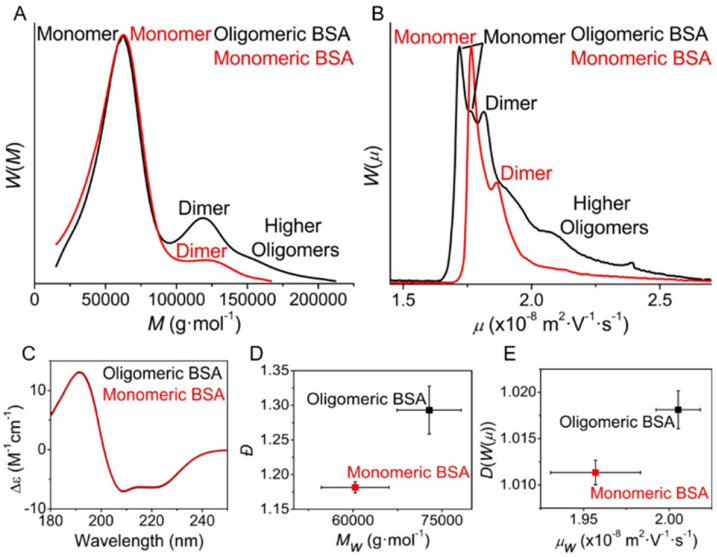


Figure 1. Separation and distribution analysis of oligomeric and monomeric BSA using SEC-MALS and CE. Representative distribution of (A) molar masses and (B) electrophoretic mobilities for oligomeric (black) and monomeric (red) BSA. (C) Far-UV CD spectra of oligomeric (black) and monomeric (red) BSA. (D) *Đ* plotted against the M_w for the whole sample. (E) $D(W(\mu))$ plotted against the μ_w for the whole sample.

also displays a broad tailing distribution with two distinct peaks but with a clear shoulder off the main peak while monomeric BSA (Figure 1B, *red*) has a far narrower distribution and two distinct peaks. The general shape of the $W(\mu)$ (Figure 1B) is qualitatively similar to that of the $W(M)$ (Figure 1A) for both oligomeric and monomeric BSA. Therefore, as the peaks in the $W(M)$ can be unambiguously assigned, we tentatively assigned similar peaks in the $W(\mu)$. In the oligomeric BSA preparation, the major peak with an μ of 1.72×10^{-8} $m^2 \cdot V^{-1} \cdot s^{-1}$ (peak max) was assigned as the monomeric form (Figure 1B, *black*). A clear shoulder on this monomeric peak suggests more than one population of monomeric conformers may exist in this sample. The distinct peak at an μ of 1.82×10^{-8} $m^2 \cdot V^{-1} \cdot s^{-1}$ (peak max) was assigned as the dimeric form, with the subsequent tailing being attributed to higher oligomers. The $W(\mu)$ of the monomeric BSA preparation shows a shift to a higher overall μ (Figure 1B, *red*). The assigned monomer peak at 1.77×10^{-8} $m^2 \cdot V^{-1} \cdot s^{-1}$ aligns with the aforementioned monomeric peak shoulder observed in the oligomeric BSA preparation (Figure 1B, *black*) further confirming its assignment as a monomeric conformer. Finally, the distinct peak at an μ of 1.87×10^{-8} $m^2 \cdot V^{-1} \cdot s^{-1}$ (peak max), which exhibits the same shift as the monomer relative to the corresponding peaks in oligomeric BSA, was assigned to a small amount of dimer which is also seen in the $W(M)$ for monomeric BSA.

To ensure the shift in μ is not due to non-native structural perturbations between the two BSA preparations, we examined far-UV CD spectra of oligomeric and monomeric BSA (Figure 1C). Both spectra exhibit a classic α-helical spectrum and are superimposable down to a wavelength of 180 nm, indicating the proteins are structurally indistinguishable in terms of their overall secondary structure. Therefore, the possibility that partial denaturation is the cause of the shift in μ between the two preparations is unlikely.

We quantitatively compared M_w and μ_w for the entire distribution to their respective $Đ$ and $D(W(\mu))$ values (Figures 1D and 1E, respectively). Consistent with the qualitative interpretations, oligomeric BSA is more disperse than monomeric BSA in both conformation and molar mass, with oligomeric BSA also having higher μ_w and M_w due to the presence of larger oligomeric species at higher μ and molar mass.

Previously, with the exception of one study that showed multiple peaks for a BSA separation using CE with a polyacrylamide coated capillary,[27] the separation of BSA by CE has largely been of very low resolution[28,29], or effected through the use of an in-solution molecular sieving agent like dextran or polyethylene glycol.[29,30] This is unsurprising given the low resolution obtained prior to optimisation (Figure S3 A). Considering CE separates on the basis of size to charge, it has been purported that the separation of BSA oligomers is not possible as both the size and charge are expected to scale uniformly upon oligomerization, essentially not providing any separable difference between oligomers in CE.[30] We have shown this to not be the case (Figure 1B), further validating the separation mechanism in CE-CC as being on the basis of charge to electrostatic friction. However, we acknowledge that our separation, particularly of higher oligomers, could be improved with further optimization of the separation parameters. Indeed, the higher oligomeric forms that are readily observed have a higher μ than lower oligomeric forms. A similar CE separation of glucosamine-6-phosphate synthase also observed that the hexameric oligomer had a higher μ than the dimeric form.[31] Together, these results indicate that oligomerization causes either a net increase in effective negative charge and/or a decrease in electrostatic friction allowing for separable differences related to the conformational changes inherent to oligomerization.

CE separation also reveals further heterogeneity in the oligomeric BSA sample where SEC-MALS did not. The heterogeneity of BSA has been studied in detail and can arise due to oligomerization, the formation of mixed disulfides, as well as the binding of metal ions or fatty acids.[25,26,32,33] The preparation of monomeric BSA evidently favors the less abundant monomer present as a shoulder in oligomeric BSA, causing the shift in μ noted for the monomeric preparation. Interestingly, the shift in μ is also noted for the small amount of dimer observable in the monomeric BSA preparation (Figure 1B, *red*) indicating the modification does not completely prohibit oligomerization. The shift in μ between the two monomer peaks in oligomeric BSA and that exhibited by monomeric BSA is smaller than what is usually expected for a full unit charge difference e.g. glutathionylation, strong metal binding or deamidation.[32,34] It therefore seems likely that either subtle conformational differences inherent to BSA's structure e.g. a difference in intramolecular disulfide bonding, or, given BSA's promiscuous binding nature, the loss of a bound molecule(s) e.g. fatty acids,[33] leads to the additional monomeric species with lower electrostatic friction. A more complete understanding of the exact nature of these differences would be furthered via coupling CE separation to MS detection.

CE separation of BSA and indeed other protein preparations generally are important for quality control in industry where the use of traditional characterization techniques render some conformational differences indiscernible. It is also important in research where BSA is used extensively as a model protein for the in-depth study of aggregation[35] potentially leading to erroneous interpretation if conformational aspects of a protein preparation are not known. Finally, for BSA, $D(W(\mu))$ from CE (Figure 1D) and $Đ$ from SEC-MALS (Figure 1E) correlate with one another, indicating that oligomeric BSA is more disperse in both conformation and molar mass. As such, using CE could provide a far more efficient alternative to SEC by providing more information in less time and with reduced sample volume and concentration.

Conformational Heterogeneity of Monomeric and Tetrameric Populations of YADH. The YADH monomer contains a catalytic domain, comprising one catalytic zinc atom and one structural zinc atom, and a coenzyme binding domain.[36,37] YADH monomers can adopt different conformations due to the

rotation of the two domains around a substrate binding cleft and, when associating to form the tetramer, can come together in a symmetric or asymmetric manner. Therefore, both the YADH monomer and the tetramer are subject to significant conformational heterogeneity.[36,37] YADH and structurally altered forms of the protein prepared by treatment with EDTA or DTT were separated by CE and SEC-MALS to assess the ability of CE to discern conformational heterogeneity inherent to oligomeric populations.

The $W(M)$ of YADH (Figure 2A, *left – bottom*) shows a small peak with a molar mass at ~36,000 $g \cdot mol^{-1}$ and a large peak at ~144,000 $g \cdot mol^{-1}$. These molar masses correspond with those expected for a monomeric and tetrameric species, respectively.[36] No other oligomeric species were observed. Changes in the oligomeric structure of YADH have been reported when using compounds that can chelate zinc from the protein such as EDTA or DTT, the latter can also reduce disulfide bonds.[38,39] After one hour of incubation with EDTA, the peak associated with YADH monomer is lost, suggesting oligomerization to the tetramer (Figure 2A, *left – middle*). After treatment with DTT for the same time, the YADH tetramer dissociates to produce an approximately even mixture of monomer and tetramer (Figure 2A, *left – top*). Comparison of the $W(M)$ with those of the $W(\mu)$ reveals complementarity that allowed monomer (highlighted in *blue*) and tetramer (highlighted in *red*) peak assignment when considering the peak alteration between the EDTA and DTT treatments.

The $W(\mu)$ of YADH reveals two well resolved, equally populated monomer peaks with an μ of 9.21 and 9.82×10^{-9} $m^2 \cdot V^{-1} \cdot s^{-1}$ (Figure 2A, *right – bottom*). Given the single peak seen in the $W(M)$, these two peaks can be ascribed to distinct monomeric conformers. They are likely the 'open' and 'closed' forms which are observed in the crystal structure of YADH (Figure 2B, *top*).[36] The YADH tetramer (Figure 2B, *bottom*) shows a broad $W(\mu)$ with multiple partially separated peaks at an μ of 1.09, 1.16, 1.21, and 1.28×10^{-8} $m^2 \cdot V^{-1} \cdot s^{-1}$. Treatment of YADH with EDTA leads to a loss of both monomer peaks in the CE separation (Figure 2A, *right – middle*), consistent with the $W(M)$. EDTA treatment also alters tetramer conformers as the peak at 1.09×10^{-8} $m^2 \cdot V^{-1} \cdot s^{-1}$ is lost, the peak at 1.16×10^{-8} $m^2 \cdot V^{-1} \cdot s^{-1}$ is less resolved and exhibits a new shouldering peak at 1.13×10^{-8} $m^2 \cdot V^{-1} \cdot s^{-1}$. Treatment with DTT shows a single large monomer peak at 9.33×10^{-9} $m^2 \cdot V^{-1} \cdot s^{-1}$ with minor species tailing up to 1.03×10^{-8} $m^2 \cdot V^{-1} \cdot s^{-1}$ (Figure 2A, *right – top*). This is consistent with an increase in the monomer peak in the $W(M)$ however, it indicates that only one of the YADH monomeric conformers predominates

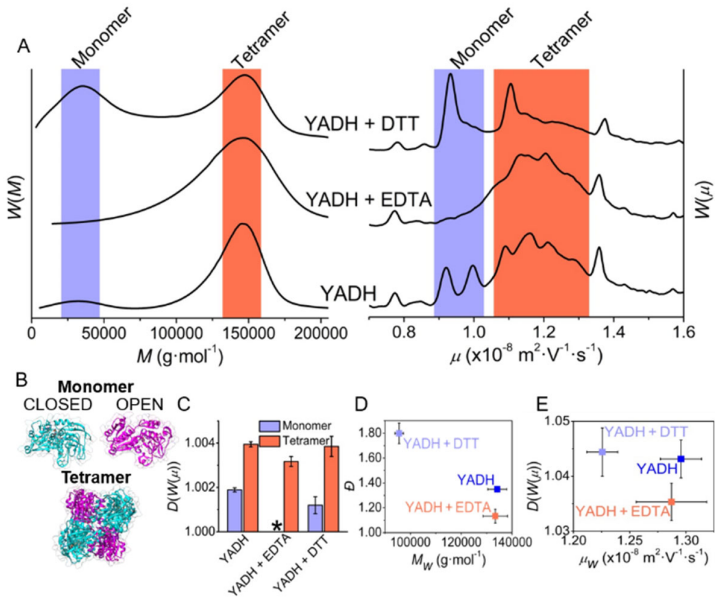


Figure 2. Separation and distribution analysis of YADH and two separate treatments with EDTA and DTT using SEC-MALS and CE. Representative distribution of (A) molar masses (*left*) and electrophoretic mobilities (*right*) for YADH (*bottom*), YADH + EDTA (*middle*) and YADH + DTT (*top*). (B) Crystal structure of YADH (PDB: 4W6Z) showing a closed (*cyan, top left*) and an open (*magenta, top right*) monomer conformation. The back-to-back YADH tetramer conformation (*bottom*) is shown with closed and open subunits coloured as above. (C) $D(W(\mu))$ calculated for the YADH monomer and tetramer during treatments. The limits used are approximated by the blue and red coloured boxes encompassing the monomer and tetramer respectively in panel A, *right*. The black asterisk in place of the YADH + EDTA monomer indicates no $D(W(\mu))$ value is applicable as there were no discernible monomeric peaks for this sample. (D) $Đ$ plotted against the M_w for the whole sample. (E) $D(W(\mu))$ plotted against the μ_w for the whole sample.

under such conditions. The tetrameric profile is altered drastically displaying a peak at 1.10×10^{-8} $m^2 \cdot V^{-1} \cdot s^{-1}$ with tailing species at higher μ, thus affirming that this peak is associated with a tetrameric conformer upon consideration of the corresponding $W(M)$. The peaks seen outside of the monomer and tetramer assignment in the $W(\mu)$ could not be confidently assigned to YADH. These peaks may be forms of YADH that remain unaffected by EDTA and DTT treatments or may be minor impurities undetected by SEC-MALS.

In order to quantitatively compare changes in the conformational heterogeneity for the monomer and tetramer upon treatment with EDTA and DTT, $D(W(\mu))$ values were calculated from the corresponding $W(\mu)$ for both species using the limits indicated by the blue and red boxes respectively (Figure 2A, *right*). The YADH monomer in the native state and the DTT-treated sample have a $D(W(\mu))$

of 1.00189 ± 0.00009, and 1.00120 ± 0.00038, respectively (Figure 2C, *blue*). A $D(W(\mu))$ value for the YADH monomer treated with EDTA could not be calculated due to a loss of the relevant peaks (Figure 2C, *black* asterisk). The native YADH contains two distinct monomeric populations, while the DTT-treated sample contains one. This is likely the reason for the higher $D(W(\mu))$ for native YADH, indicating greater monomer conformational heterogeneity relative to the DTT-treated sample (Figure 2C, *blue*). The $D(W(\mu))$ values for the tetrameric species were calculated as 1.00395 ± 0.00012, 1.00317 ± 0.00022, and 1.00385 ± 0.00045, for native YADH, EDTA and DTT-treated samples, respectively (Figure 2C, *red*). Native and DTT-treated YADH tetramers exhibit similar $D(W(\mu))$ values and are therefore similar in conformational heterogeneity despite the difference in peak profiles. The tetramer of EDTA-treated YADH has the lowest $D(W(\mu))$ value and is therefore less conformationally heterogeneous which could be rationalized by the loss of distinct peaks relative to native YADH. The tetramer has a higher $D(W(\mu))$ value than the monomer in all cases, indicating greater conformational heterogeneity for the higher oligomeric species (Figure 2C) in agreement with previous structural studies.[36]

Quantitative parameters were calculated for $W(M)$ and $W(\mu)$ for native YADH, EDTA and DTT-treated samples and plotted as the $Đ$ versus M_w (Figure 2D) and the $D(W(\mu))$ versus μ_w (Figure 2E). A positive correlation exists between M_w and μ_w which is consistent with the findings for BSA. $Đ$ and $D(W(\mu))$ for YADH and the EDTA-treated samples have similar relative positions on the scatter plot. However, the relative difference in $Đ$ between the DTT-treated sample and the native YADH and EDTA-treated samples is far more distinctive compared to the corresponding $D(W(\mu))$, indicating that molar mass and conformational heterogeneity are discernibly different in this instance.

Previously, CE studies involving YADH have focussed on monitoring the native enzyme's catalytic activity[40] or investigating its interaction with the surfactant sodium dodecyl sulfate.[5] However, no study has used CE to characterize its structure. CE analysis of YADH provides, to our knowledge, the first in-solution separation of the closed and open monomeric conformers (Figure 2A, *right – bottom*) evident in the crystal structure[36] and in native polyacrylamide gel electrophoresis.[41] A similar CE experiment has assigned open and closed conformations of human trans-glutaminase.[17,18] Given the open conformation for human trans-glutaminase has a lower μ, we suggest the peak at 9.21×10^{-9}

$m^2 \cdot V^{-1} \cdot s^{-1}$ be assigned to the open conformer of YADH and the peak at 9.82×10^{-9} $m^2 \cdot V^{-1} \cdot s^{-1}$ be assigned to the closed conformer (Figure 2A, *right – bottom, blue*). The open conformation is likely more permeable to counter ions and would therefore be subject to greater electrostatic friction relative to the closed conformation, thereby causing a lower mobility. However, amino acid charge differences due to alteration in tertiary structure may also be a contributing factor.

The presence of multiple tetrameric peaks in the $W(\mu)$ indicates several conformational species are present in the native protein (Figure 2A, *right – bottom, red*). This is consistent with the noted asymmetry in the crystal structure of YADH and the different monomeric conformers which can oligomerize to adopt numerous tetrameric conformers. A potential further contribution to the observed heterogeneity is the alternative coordination sphere of the catalytic zinc ion which can induce slight changes in the zinc ions effective charge due to changes in the proximity of the Glu-67 side chain.[36,37] Akin to the monomer, the conformers inherent to the tetrameric population of YADH at a lower μ, e.g. 1.09×10^{-8} $m^2 \cdot V^{-1} \cdot s^{-1}$, would likely have a more open conformation, while those at a higher μ, e.g. 1.28×10^{-8} $m^2 \cdot V^{-1} \cdot s^{-1}$, would likely be more closed. Such differences in μ can similarly be influenced by a change in net negative charge.

A previous study has examined the effect of EDTA and DTT on the catalytic function, zinc content, and heat-labile fraction of YADH.[38] The study indicated no significant changes in zinc content when YADH is treated with low concentrations of EDTA similar to that used here which is consistent with no major shifts to higher μ. However, the reason that the monomeric species is abolished and changes occur to the peaks associated with tetrameric conformers upon EDTA treatment is not fully understood (Figure 2A, *right – middle*). Counter-ions, particularly bivalent cations, have an effect on the stability and oligomeric state of YADH.[42] It is therefore possible that counter-ions in the sample[38] associate with native YADH to stabilize a small quantity of monomer. The addition of EDTA removes these ions and the monomeric subunits associate into the tetrameric form. Further, the loss of heterogeneity in the YADH tetrameric species could potentially be due to a change in inter-subunit electrostatic interactions upon removal of counter-ions.

The intense single monomer peak observed for the DTT-treated sample indicates dissociation from the tetrameric state (Figure 2A, *right – top*), as noted in previous studies.[36,39] This is likely the

open conformer with an increased net negative charge due to the loss of structural zinc[38], resulting in a slightly higher μ (Figure 2A, *right – top*). Similarly, the tetrameric population contains a peak at 1.10×10^{-8} $m^2 \cdot V^{-1} \cdot s^{-1}$ which is likely related the peak at 1.09×10^{-8} $m^2 \cdot V^{-1} \cdot s^{-1}$ in the native enzyme. While EDTA is a stronger chelator of zinc, DTT is able to abstract the more exposed structural zinc from YADH.[38] This is likely due to the removal of an obstructive disulfide bond but may also be helped by DTT providing a thiol ligand for interchange with the four cysteine thiol groups that coordinate the structural zinc.[36,38] Abstraction of zinc from the structural site ultimately disrupts important interfacial subunit interactions, thereby leading to the noted large-scale dissociation to YADH monomer.

This study of YADH suggests that the open monomer or tetramer conformer can exist without a structural zinc or disulfide bonds, leading to the conclusion that these structural features are pivotal in maintaining the closed conformation of YADH. Further, upon EDTA treatment, the monomer and the previously studied 'heat-labile fraction'[38] are lost while the opposite is noted for DTT treatment, providing a correlation between the YADH monomer and the enzyme's heat-labile fraction. $D(W(\mu))$ values have enabled quantitative comparisons of the conformational heterogeneity between oligomeric species within a sample, as well as between structurally altering treatments, thus aiding understanding in how $D(W(\mu))$ is linked to protein conformational heterogeneity. Finally, while a deeper discussion on the structural and functional implications of these results is beyond the scope of this study, the results presented herein provide a platform for further studies to examine the link between enzyme-cofactor interaction, catalytic function and conformational heterogeneity.

Metal Binding and Conformational Heterogeneity of Monodisperse BLA. BLA is a monomeric calcium-binding protein that undergoes a change in tertiary structure between the Ca^{2+} bound (holo-form (Figure 3B, *blue, top*)) and unbound (apo-form (Figure 3B, *yellow, bottom*)) forms.[43,44] To understand how changes in conformation and net charge affects a monomeric protein's conformational heterogeneity, we used CE and SEC-MALS to separate BLA (Ca^{2+} dep) in its untreated state, BLA + EDTA which formed apo-BLA, and both BLA + Ca^{2+} and BLA + Mg^{2+} which formed holo-BLA.

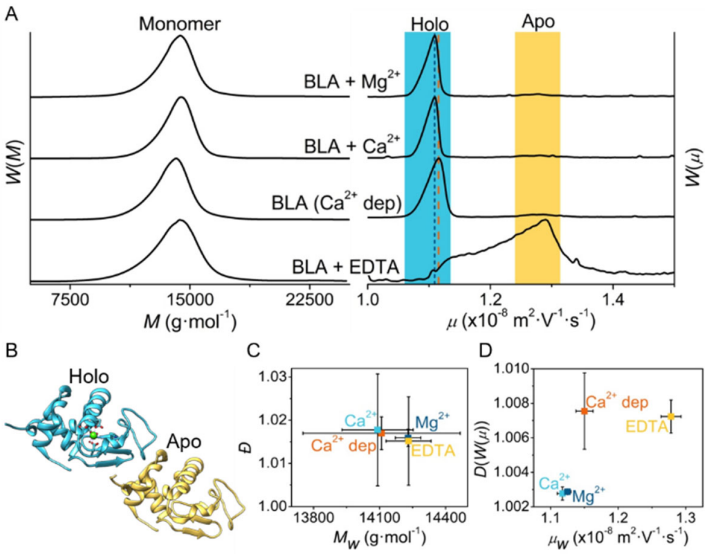


Figure 3. Separation and distribution analysis of BLA (Ca^{2+} dep) and corresponding individual treatments with EDTA, Ca^{2+} and Mg^{2+} using SEC-MALS and CE. Representative distribution of (A) molar masses (*left*) and electrophoretic mobilities (*right*) for BLA + EDTA (*bottom*), BLA (Ca^{2+} dep) (*middle*, *bottom*), BLA + Ca^{2+} (*middle, top*) and BLA + Mg^{2+} (*top*). The holo and apo peaks are highlighted in blue and yellow respectively as discerned from the $W(\mu)$. The small shift in μ between BLA (Ca^{2+} dep), orange large-dash line, and BLA + Ca^{2+} and BLA + Mg^{2+}, purple small-dash line, is indicated within the holo designation. (B) Crystal structure of holo-BLA (*blue, top,* PDB: 1F6S) with a Ca^{2+} ion coordinated (*green sphere*) and apo-BLA (*yellow, bottom,* PDB: 1F6R). (C) $Đ$ plotted against the M_w for the whole sample. (D) $D(W(\mu))$ plotted against the μ_w for the whole sample.

The $W(M)$ for all samples including BLA treated with EDTA, the Ca^{2+} depleted form, and the addition of Ca^{2+}, and Mg^{2+} (Figure 3A, *left – bottom, middle bottom, middle top, top,* respectively) display a single peak with a molar mass of 14,200 $g \cdot mol^{-1}$ in agreement with the monomeric mass of BLA.[45] In contrast, the $W(\mu)$ for BLA treated with EDTA (Figure 3A, *right – bottom*) displays a broad distribution with an μ at peak max of 1.29×10^{-8} $m^2 \cdot V^{-1} \cdot s^{-1}$. BLA ($Ca^{2+}$ dep) showed two peaks, the less intense of the two is more easily distinguishable on the unstacked $W(\mu)$ (Figure S4C). The more intense peak corresponds with an μ of 1.12×10^{-8} $m^2 \cdot V^{-1} \cdot s^{-1}$ (Figure 3A, *right – orange large-dash line*) while the lower intensity peak has an μ of 1.29×10^{-8} $m^2 \cdot V^{-1} \cdot s^{-1}$ (Figure 3A, *right – middle bottom*). BLA treated with Ca^{2+} (Figure 3A, *right – middle top*) and BLA treated with Mg^{2+} (Figure 3A, *right – top*) both also show two peaks (Figure S4C). For these treatments, the more intense peak is seen at an μ of

1.11×10^{-8} $m^2 \cdot V^{-1} \cdot s^{-1}$ (Figure 3A, *right – purple short-dash line*) and the less intense peak at 1.29×10^{-8} $m^2 \cdot V^{-1} \cdot s^{-1}$.

The $Đ$ versus M_w (Figure 3C) shows that, within error, there is no difference between the BLA samples, regardless of treatment. The $D(W(\mu))$ versus μ_w (Figure 3D) however, shows a more disperse plot. BLA treated with Ca^{2+} and Mg^{2+} have similar μ_w and $D(W(\mu))$, with both these values being lower than that of the other two samples. BLA (Ca^{2+} dep) and EDTA-treated samples have the same $D(W(\mu))$ within error, however BLA (Ca^{2+} dep) has a lower μ_w than EDTA-treated BLA.

BLA has been used to show the effectiveness of CE in separating bovine whey proteins,[46] to investigate how capillary temperature effects the μ of proteins,[11] to examine protein interactions with surfactant[5] and to demonstrate the ability of CE charge ladders to discern the effect of electrostatics on the thermodynamics of BLA unfolding.[12] While CE-based structural studies on BLA have been insightful, the assignment of the apo- and holo-forms and the corresponding conformational change that occurs due to metal binding has, to our knowledge, not been addressed using CE. This is despite two distinct peaks being observed in previous CE separation studies of BLA.[11,12] Considering Ca^{2+} and EDTA treatments of BLA, we have assigned two populations in the $W(\mu)$ at 1.11×10^{-8} $m^2 \cdot V^{-1} \cdot s^{-1}$ and 1.29×10^{-8} $m^2 \cdot V^{-1} \cdot s^{-1}$ to holo- and apo-BLA, respectively (Figure 3A, *right – blue and yellow,* respectively). This is consistent with the apo-form being more negatively charged than the cation-bound holo-form thus resulting in a higher μ for the apo-form.

Metal binding, most notably by Ca^{2+}, induces a conformational change in BLA[43,44,47] and it was anticipated that these conformational differences might be resolved in CE. However, as the conformational change is contingent on metal binding, it would seem any distinct conformational changes are essentially masked by the charge difference of these species. Despite this, CE separation still displayed numerous differences that can be linked to changes in charge or conformational heterogeneity that were not evident in SEC-MALS. The similarities in μ for the holo-peak as well as a low overall $D(W(\mu))$ for Ca^{2+} and Mg^{2+}-treated BLA samples indicate that holo-BLA is comparably low in conformational heterogeneity. Interestingly, the holo-peak of BLA (Ca^{2+} dep) shows a shift to a slightly higher μ indicating a slight increase in negative charge or decrease in electrostatic friction. Also, the $D(W(\mu))$ for BLA (Ca^{2+} dep) is increased relative to Ca^{2+} and Mg^{2+} treatments, which implies that

BLA charge and/or conformational heterogeneity is increased. These results could potentially indicate a shift in metal binding equilibrium of BLA (Ca^{2+} dep), or the binding of adventitious metal ions that contribute less positive change than either Ca^{2+} and Mg^{2+}, or the emergence of new BLA conformers in the Ca^{2+} dep-form.

EDTA-treated BLA shows a shift to the apo-form with a higher μ due to metal being abstracted from the protein. The broad $W(\mu)$ of EDTA-treated BLA encompasses some μ values that are not assigned to either apo- or holo-BLA. It is unclear why the distribution displays this broadness however it is possible that an intermediate conformational exchange rate exists between apo- and holo-BLA conformers[9] facilitated by metal interchange between chelator and BLA. The $D(W(\mu))$ is also increased due to this conformational heterogeneity and is unexpectedly comparable to BLA (Ca^{2+} dep). This indicates that a BLA species that is between the apo- and holo-states (e.g. BLA (Ca^{2+} dep) and BLA + EDTA) can be characterized as having maximum charge and/or conformational heterogeneity. However holo-BLA (e.g. BLA + Ca^{2+} and BLA + Mg^{2+}) and presumably pure apo-BLA can be characterized by low heterogeneity.

CE separation of monomeric BLA during modulation of apo- and holo-forms has provided a new dynamic to previously discussed oligomeric proteins, BSA and YADH, showing an obvious orthogonality between CE and SEC-MALS separation. Analysing CE data in terms of the $W(\mu)$ enables assignment and characterization of different metal-bound states of BLA and the ensuing charge and/or conformational heterogeneity via the $D(W(\mu))$ value. Such analyses are useful for industrial applications such as high-throughput quality control and research where it can be used to study sample heterogeneity and obtain metal binding affinities.

Comparison of Molar Mass and Conformational Dispersity for Different Proteins. A direct comparison of the overall $D(W(\mu))$ and $Đ$ for all protein samples studied reveals further insight into the interplay between molar mass and conformational heterogeneity (Figure 4). Relative to all proteins used, holo- and apo-BLA have the lowest conformational and molar mass heterogeneity. This may indicate that preferentially detected charge-based changes associated with a monomeric protein such as BLA will only alter $D(W(\mu))$ subtly in comparison to proteins with distinct oligomeric forms such as

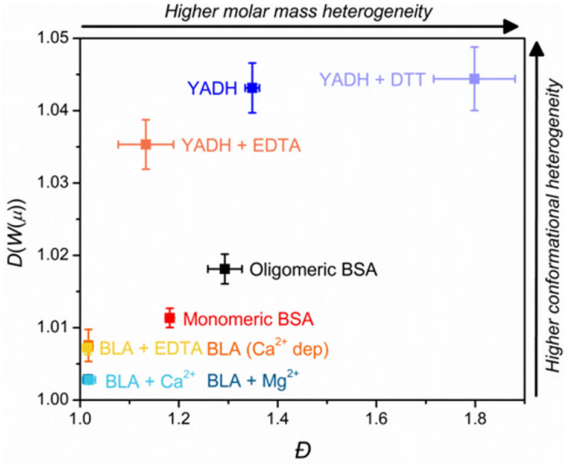


Figure 4. Comparison of the $D(W(\mu))$ and $Đ$, for BSA, YADH and BLA and related purified forms/treatments to discern comparative similarities and differences in conformational and molar mass heterogeneity.

oligomeric BSA or those of high conformational dispersity like YADH. Relative to BLA, $D(W(\mu))$ is increased in monomeric BSA and YADH + EDTA samples which both have similar $Đ$, the former being largely monomeric and the latter being largely tetrameric. However, both samples have significantly different $D(W(\mu))$, which is likely indicative of the noted conformational heterogeneity inherent to YADH whereas monomeric BSA is more homogenous in terms of conformation. $D(W(\mu))$ and $Đ$ are similarly increased for both oligomeric BSA and native YADH relative to monomeric BSA and YADH + EDTA, respectively, likely due to the presence of additional oligomeric forms in the BSA sample, and the presence of monomers and more distinct tetramers in native YADH. While there is no discernible difference in $D(W(\mu))$ between YADH and YADH + DTT, there is a significant increase in $Đ$, perhaps due to the mass difference between equally intense monomer and tetramer populations being far greater than the difference in μ.

Information inherent to specific groupings when comparing protein $D(W(\mu))$ and $Đ$ values may provide a platform from which to compare proteins displaying similarities in conformational and molar mass heterogeneity and subsequently discern any physical or functional commonalities within groupings (Figure 4). Our studies indicate that protein's with a high molar mass heterogeneity yet low conformational heterogeneity does not currently seem likely as high levels of oligomerization mostly occasioning high $Đ$ will still lead to some increase in $D(W(\mu))$ even if the individual oligomers are low in conformational heterogeneity, e.g. oligomeric BSA.

CONCLUDING REMARKS

The ability for CE to discern distinct protein conformers that are otherwise not accessible with SEC-MALS, and to infer the comparative level of conformational heterogeneity via the $D(W(\mu))$ from the $W(\mu)$, has been demonstrated here. New structural information from CE separations was also highlighted and assigned with the aid of complementary SEC-MALS analysis. The structural information was used to rationalize differences when comparing $D(W(\mu))$ with $Đ$ which yielded novel insight into the relationship between conformational and molar mass heterogeneity. Overall, this study has introduced the application of $D(W(\mu))$ to explore conformational heterogeneity of proteins and provided insight into what affects its quantification, leading to an additional important dimension of information for protein characterization by CE. Future studies will concern how these methods can be applied to examine the conformational heterogeneity of more complex proteins. This includes the α-crystallins, which are structurally heterogeneous due to the formation of complex oligomeric arrays, and post-translationally modified crystallins, which can exhibit charge (phosphorylation and deamidation), and conformational (truncation and racemization) heterogeneity.

CONTENT

Supporting Information

Representative PM elugrams; capillary validation via separation of AA3; comparing the separation of oligomeric proteins at 40 cm and 80 cm capillary lengths; overlaid representative $W(\mu)$ for all proteins and relevant treatments; protein dn/dc linear fits; overlaid representative chromatograms for all proteins and relevant treatments; table listing treatments used for each protein and their concentrations; table listing correlation coefficients for Gaussian fits to PM elugrams.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: aidan.grosas@anu.edu.au

*E-mail: p.castignolles@westernsydney.edu.au

ORCID

Aidan Grosas: 0000-0003-0069-7912

John Carver: 0000-0002-2441-8108

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

ABG thanks Dr. Nicholas J. Ray (Australian National University) for assistance with SEC-MALS and Dr. David C. Thorn (Australian National University) for thoughtful discussions. ABG acknowledges the financial support of an Australian Postgraduate Award. MDP acknowledges the financial support of the Australian Research Training Program. PC acknowledges an Academic Development Grant from Western Sydney University.

REFERENCES

(1) Grossman, P. D.; Colburn, J. C.; Lauer, H. H.; Nielsen, R. G.; Riggin, R. M.; Sittampalam, G. S.; Rickard, E. C. *Anal. Chem.* **1989**, *61*, 1186-1194.

(2) Fonslow, B. R.; Yates, J. R. *J. Sep. Sci.* **2009**, *32*, 1175-1188.

(3) Thevarajah, J. J.; Sutton, A. T.; Maniego, A. R.; Whitty, E. G.; Harrisson, S.; Cottet, H.; Castignolles, P.; Gaborieau, M. *Anal. Chem.* **2016**, *88*, 1674-1681.

(4) Karim, M. R.; Shinagawa, S.; Takagi, T. *Electrophoresis* **1994**, *15*, 1141-1146.

(5) Gudiksen, K. L.; Gitlin, I.; Whitesides, G. M. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 7968-7972.

(6) Stellwagen, N. C.; Gelfi, C.; Righetti, P. G. *Biopolymers* **1997**, *42*, 687-703.

(7) Cottet, H.; Gareil, P.; Theodoly, O.; Williams, C. E. *Electrophoresis* **2000**, *21*, 3529-3540.

(8) Hoagland, D. A.; Arvanitidou, E.; Welch, C. *Macromolecules* **1999**, *32*, 6180-6190.

(9) Hilser, V. J.; Freire, E. *Anal. Biochem.* **1995**, *224*, 465-485.

(10) Rochu, D.; Masson, P. *Electrophoresis* **2002**, *23*, 189-202.

(11) Rush, R. S.; Cohen, A. S.; Karger, B. L. *Anal. Chem.* **1991**, *63*, 1346-1350.

(12) Negin, R. S.; Carbeck, J. D. *J. Am. Chem. Soc.* **2002**, *124*, 2911-2916.

(13) de Kort, B. J.; ten Kate, G. A.; de Jong, G. J.; Somsen, G. W. *Anal. Chem.* **2011**, *83*, 6060-6067.

(14) Barr, J. D.; Shi, L. Q.; Russell, D. H.; Clemmer, D. E.; Holliday, A. E. *Anal. Chem.* **2016**, *88*, 10933-10939.

(15) Haselberg, R.; de Jong, G. J.; Somsen, G. W. *J. Chromatogr. A* **2007**, *1159*, 81-109.

(16) Bertoletti, L.; Schappler, J.; Colombo, R.; Rudaz, S.; Haselberg, R.; Dominguez-Vega, E.; Raimondi, S.; Somsen, G. W.; De Lorenzi, E. *Anal. Chim. Acta* **2016**, *945*, 102-109.

(17) Clouthier, C. M.; Mironov, G. G.; Okhonin, V.; Berezovski, M. V.; Keillor, J. W. *Angew. Chem.-Int. Edit.* **2012**, *51*, 12464-12468.

(18) Mironov, G. G.; Clouthier, C. M.; Akbar, A.; Keillor, J. W.; Berezovski, M. V. *Nat. Chem. Biol.* **2016**, *12*, 918-922.

(19) Eriksson, J. H. C.; Mol, R.; Somsen, G. W.; Hinrichs, W. L. J.; Frijlink, H. W.; de Jong, G. J. *Electrophoresis* **2004**, *25*, 43-49.

(20) Chamieh, J.; Martin, M.; Cottet, H. *Anal. Chem.* **2015**, *87*, 1050-1057.

(21) Baumann, M.; Meri, S. *Expert Rev. Proteomics* **2004**, *1*, 207-217.

(22) Gaborieau, M.; Causon, T. J.; Guillaneuf, Y.; Hilder, E. F.; Castignolles, P. *Aust. J. Chem.* **2010**, *63*, 1219-1226.

(23) de Jong, S.; Krylov, S. N. *Anal. Chem.* **2012**, *84*, 453-458.

(24) Castignolles, P.; Gaborieau, M.; Hilder, E. F.; Sprong, E.; Ferguson, C. J.; Gilbert, R. G. *Macromol. Rapid Commun.* **2006**, *27*, 42-46.

(25) Janatova, J.; Fuller, J. K.; Hunter, M. J. *J. Biol. Chem.* **1968**, *243*, 3612-3622.

(26) Hunter, A. K.; Carta, G. *J. Chromatogr. A* **2001**, *937*, 13-19.

(27) Kubo, K. *Anal. Biochem.* **1996**, *241*, 42-46.

(28) Pande, P. G.; Nellore, R. V.; Bhagat, H. R. *Anal. Biochem.* **1992**, *204*, 103-106.

(29) Tseng, W. L.; Lin, Y. W.; Chang, H. T. *Anal. Chem.* **2002**, *74*, 4828-4834.

(30) Zhu, M. D.; Hansen, D. L.; Burd, S.; Gannon, F. *J. Chromatogr.* **1989**, *480*, 311-319.

(31) Beneito-Cambra, M.; Gareil, P.; Badet, B.; Badet-Denisot, M. A.; Delaunay, N. *J. Chromatogr. B* **2018**, *1072*, 130-135.

(32) Andersson, L. O. *Biochim. Biophys. Acta* **1966**, *117*, 115-133.

(33) Kragh-Hansen, U. *Pharmacol. Rev.* **1981**, *33*, 17-53.

(34) Shi, Y. H.; Rhodes, N. R.; Abdolvahabi, A.; Kohn, T.; Cook, N. P.; Marti, A. A.; Shaw, B. F. *J. Am. Chem. Soc.* **2013**, *135*, 15897-15908.

(35) Borzova, V. A.; Markossian, K. A.; Chebotareva, N. A.; Kleymenov, S. Y.; Poliansky, N. B.; Muranov, K. O.; Stein-Margolina, V. A.; Shubin, V. V.; Markov, D. I.; Kurganov, B. I. *Plos One* **2016**, *11*, 1-29.

(36) Raj, S. B.; Ramaswamy, S.; Plapp, B. V. *Biochemistry* **2014**, *53*, 5791-5803.

(37) Plapp, B. V.; Charlier, H. A.; Ramaswamy, S. *Arch. Biochem. Biophys.* **2016**, *591*, 35-42.

(38) Magonet, E.; Hayen, P.; Delforge, D.; Delaive, E.; Remacle, J. *Biochem. J.* **1992**, *287*, 361-365.

(39) Debolle, X.; Vinals, C.; Prozzi, D.; Paquet, J. Y.; Leplae, R.; Depiereux, E.; Vandenhaute, J.; Feytmans, E. *Eur. J. Biochem.* **1995**, *231*, 214-219.

(40) Avila, L. Z.; Whitesides, G. M. *J. Org. Chem.* **1993**, *58*, 5508-5512.

(41) Nowakowski, A. B.; Wobig, W. J.; Petering, D. H. *Metallomics* **2014**, *6*, 1068-1078.

(42) DeBolle, X.; Vinals, C.; Fastrez, J.; Feytmans, E. *Biochem. J.* **1997**, *323*, 409-413.

(43) Kronman, M. J.; Sinha, S. K.; Brew, K. *J. Biol. Chem.* **1981**, *256*, 8582-8587.

(44) Chrysina, E. D.; Brew, K.; Acharya, K. R. *J. Biol. Chem.* **2000**, *275*, 37021-37029.

(45) Permyakov, E. A.; Berliner, L. J. *FEBS Lett.* **2000**, *473*, 269-274.

(46) Kinghorn, N. M.; Norris, C. S.; Paterson, G. R.; Otter, D. E. *J. Chromatogr. A* **1995**, *700*, 111-123.

(47) Murakami, K.; Andree, P. J.; Berliner, L. J. *Biochemistry* **1982**, *21*, 5488-5494.

Supporting Information

Using Capillary Electrophoresis to Investigate Protein Conformational Heterogeneity: a Comparative Study between the Distribution of Electrophoretic Mobilities and Molar Masses

Aidan B. Grosas, Mar-dean Du Plessis, Joel J. Thevarajah, Marianne Gaborieau, John A. Carver, and Patrice Castignolles

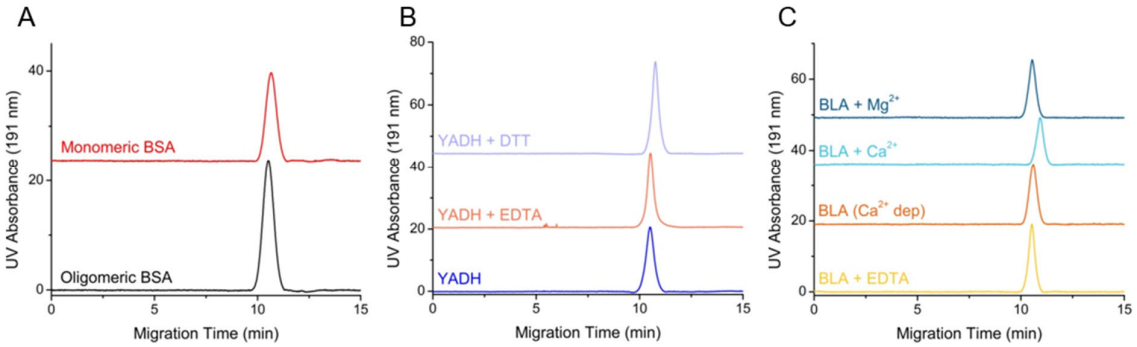


Figure S1. Baseline corrected representative elugrams obtained using PM which ensures that the protein has minimal adsorption on the capillary wall thereby validating separation in CE. Instrument and sample conditions for PM are detailed in the *Materials and Methods*. All comparable samples have been displaced along the y-axis for comparison. Slight differences in migration time are likely due to small deviations in pressure during the PM experiment. (A) BSA and mBSA. (B) YADH and two treatments, with EDTA and DTT. (C) BLA (Ca^{2+} dep) and three treatments using EDTA, Ca^{2+} and Mg^{2+} respectively.

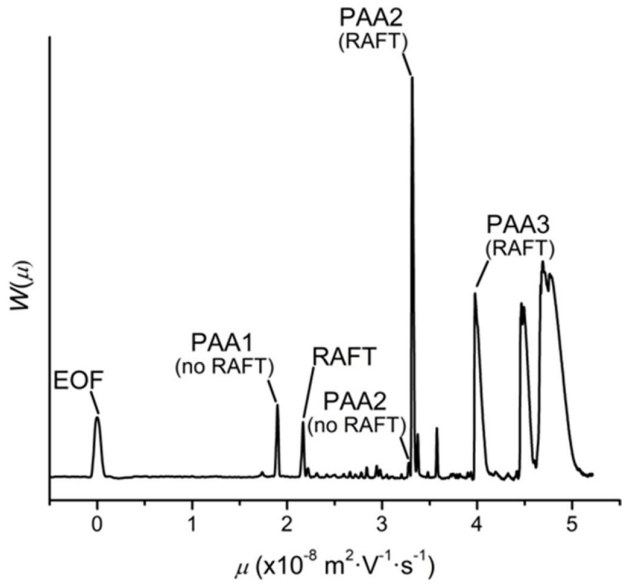


Figure S2. Electropherogram of synthetic polymer mix AA3 used to validate the fused silica capillary before commencing experiments. Separation of this complex mixture containing different sized oligomers of synthetic acrylic acid (PAA) with or without reversible addition-fragmentation chain transfer (RAFT) agent allows the CE user to determine if separation efficiency is maintained.

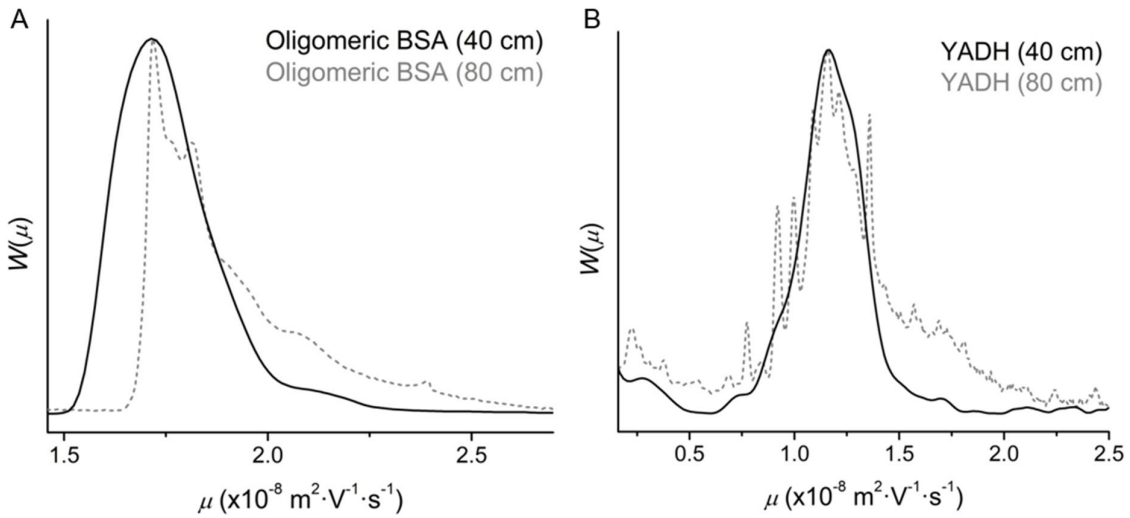


Figure S3. Distributions of electrophoretic mobilities obtained using CE of (A) oligomeric BSA and (B) YADH illustrating the increase in the separation resolution between a 40 cm (*black, solid line,* 31.5 cm to detector) and an 80 cm (*grey, dashed line,* 71.5 cm to detector) fused silica capillary. The injection volume remained the same between both capillary lengths which may have led to overloading at the 40 cm length. All other sample conditions and concentrations for the respective proteins are as outlined in the *Materials and Methods*.

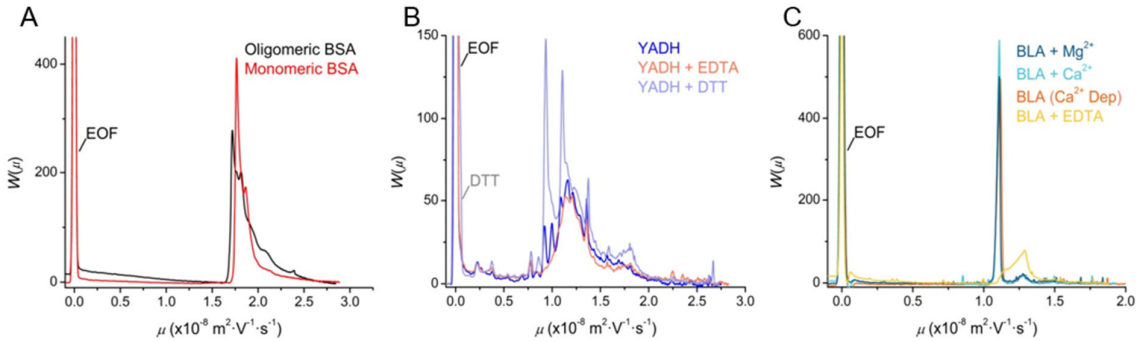


Figure S4. Overlaid $W(\mu)$. (A) Oligomeric BSA and monomeric BSA, (B) YADH only and also treated with EDTA and DTT, and (C) BLA (Ca^{2+} dep) and also treated with EDTA, Ca^{2+} and Mg^{2+}. The peak from the neutral marker 0.2 % DMSO is labelled as the EOF. The presence of DTT is attributed to the shouldering seen on the EOF peak (B).

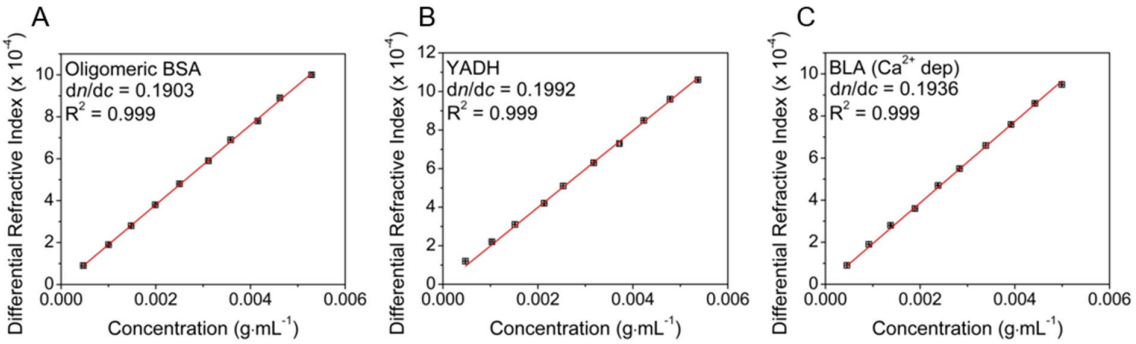


Figure S5. Calculation of protein d*n*/d*c*. The refractive index of (A) oligomeric BSA, (B) YADH, and (C) BLA (Ca^{2+} dep) are analyzed over 10 concentrations and fitted to a linear equation through point 0,0. The gradient of the linear fit is given as the protein's d*n*/d*c* at 25 °C in 5 mM NaPi, pH 7.4 for BSA and YADH and in 20 mM tris-HCl, pH 7.4 for BLA. The d*n*/d*c* and R^2 fit value are given on each respective plot.

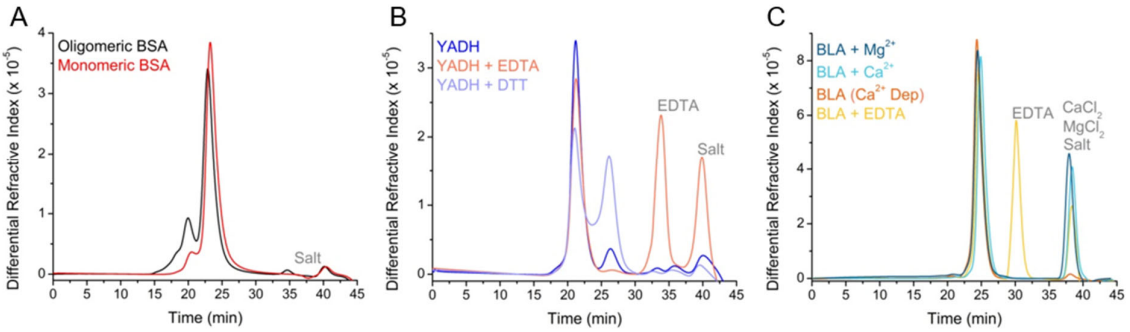


Figure S6. Overlaid chromatograms used for $W(M)$ determination obtained using SEC with refractive index detection. (A) Oligomeric BSA and monomeric BSA, (B) YADH only and also treated with EDTA and DTT, and (C) BLA (Ca^{2+} dep) and also treated with EDTA, Ca^{2+} and Mg^{2+}. Salts and reagents added for treatments have been assigned based on the appropriate blank e.g. treatment only in relevant buffer.

Table S1. Model protein and corresponding treatments for one hour at 25 °C, p*H* 7.4.

CE Experiments		SEC-MALS Experiments	
Protein	Treatment	Protein	Treatment
1 $g \cdot L^{-1}$ BSA	None	3 $g \cdot L^{-1}$ BSA	None
1 $g \cdot L^{-1}$ YADH	2 mM EDTA	3 $g \cdot L^{-1}$ YADH	6 mM EDTA
	2 mM DTT		6 mM DTT
0.5 $g \cdot L^{-1}$ BLA	1 mM $CaCl_2$	3 $g \cdot L^{-1}$ BLA	6 mM $CaCl_2$
	1 mM $MgCl_2$		6 mM $MgCl_2$
	1 mM EDTA		6 mM EDTA

Table S2. Adjusted R^2 for the fit of a Gaussian function to pressure mobilization data. Errors given as the standard deviation of five repeats.

Sample	Average Adjusted R^2
BSA (oligomeric)	0.958 ± 0.014
BSA (monomeric)	0.976 ± 0.009
YADH	0.978 ± 0.002
YADH + EDTA	0.971 ± 0.006
YADH + DTT	0.951 ± 0.017
BLA	0.976 ± 0.013
BLA + Ca^{2+}	0.987 ± 0.018
BLA + Mg^{2+}	0.992 ± 0.005
BLA + EDTA	0.993 ± 0.001

CHAPTER 5:

CONCLUSIONS AND FUTURE DIRECTIONS

"Everyone has a hidden agenda. Except me!"

– Michael Crichton

The work presented in this thesis aims to further our understanding of the molecular mechanisms inherent to cataract and find new methods of studying protein heterogeneity that may provide a novel perspective on lens crystallins. Changes in the lens with age, such as oxidation, and the impacts of significant macromolecular crowding can affect the crystallin proteins that maintain lens transparency. Therefore, the structure, function and aggregation-propensity of prominent lens crystallin proteins γS- (γSc) and αB-crystallin (αBc) under these conditions are of interest and have been studied herein. Further, this work examined the use of capillary electrophoresis (CE) to study protein conformational heterogeneity via a proof-of-concept study on well-characterised proteins that are also used as model aggregation-prone proteins.

Chapter 2 provided insight into how oxidation of lens proteins can contribute to age-related cataract (1). The crystal structure of the disulfide-linked human γSc dimer provided a clear view of the intra- and inter-molecular disulfide bond arrangement and, in conjunction with small angle X-ray scattering, the overall structure of the protein. This was an important step for structural characterisation of crystallins given full-length human γSc's recalcitrance toward crystallisation (2, 3). The loss of the principal lens reductant, glutathione, with age leads to elevated oxidative conditions in the aging lens which is further worsened at the onset of cataract. Importantly, it was established that the C24-C24' disulfide-linked γSc dimer was stable at glutathione concentrations akin to that in aged and cataractous lenses, but not in those of healthy/young lenses where the disulfide bond was reduced giving rise to the monomer. Biophysical and thermal stability studies provided a rationale for the greater aggregation propensity of the γSc dimer. The dimer was not structurally perturbed relative to the monomer but had an increased aggregation propensity and exhibited non-cooperative domain unfolding. These characteristics are markedly similar to that of the well-studied cataract-associated mutant G18V γSc, indicating that congenital and age-related forms of cataract may share similar developmental pathways on a molecular level. This work also supports the development of approaches, for example small molecule therapeutics, that stabilise the reducing environment of the lens to stave off oxidative events that lead to the destabilisation of the crystallins.

The role of disulfide formation in the lens and its relation to opacity is a pertinent field of investigation (4). In addition to the work described in Chapter 2 (1), a recent study has demonstrated

that γD-crystallin (γDc) has the ability to act as an oxidoreductase (5), while a lens proteomic analysis showed a shift from intra- to inter-molecular disulfides in crystallins of younger versus older lenses, respectively (6). These findings particularly underscore the importance of the characterisation of disulfide-linked crystallins and their contribution to cataract etiology. To this end, the formation of a C110-C110' disulfide-linked γDc dimer has been associated with instability and heightened aggregation propensity when in the presence of Cu(II) (7, 8). However, no studies have purified this disulfide-linked crystallin to homogeneity and compared it to the reduced (monomeric) form as was undertaken in Chapter 2 of this thesis (1). It would be of interest to perform this analysis in order to unambiguously determine if disulfide-linked dimerisation has similar effects on γDc as for γSc, notably because the purported C110-C110' intermolecular disulfide in γDc dimer would result in a C-terminal-to-C-terminal domain linkage, a significantly different dimer architecture to that of N-terminal-to-N-terminal domain linked γSc dimer (via C24 of each monomer). An additional comparison to the disulfide-linked dimer formation of cataract-associated mutant R14C γDc, yielding a N-terminal-to-N-terminal domain γDc dimer-linkage, given R14 exhibits greater solvent exposure relative to C110, would also be of interest (9). Further, investigation into the formation of disulfide-linked dimers of γC-crystallin formed via its N-terminal domain in the presence of high equivalents of Hg(II) (10) would help to refine the hypothesis that disulfide-based dimerisation is ultimately deleterious in age-related cataract.

Chapter 3 investigated the structure and function of small heat-shock protein αBc in highly crowded conditions analogous to the eye lens using either high concentrations of a polysaccharide, Ficoll 400, or a protein, bovine γ-crystallin. Contrast variation small angle neutron scattering (SANS) and other biophysical techniques showed that αBc under crowded conditions became destabilized, unfolded, increased in size/oligomeric state, exhibited reduced chaperone ability and formed amorphous and fibrillar aggregates. The inherent stability of the lens partner protein αA-crystallin (αAc) stabilised αBc and prevented its aggregation in the crowded environment when mixed together at physiologically relevant ratios.

Studies in the dilute environment *in vitro* have noted the difference in stability between αBc and αAc and demonstrated the ability of αAc to prevent the kinetic aggregation of αBc using heat stress (11, 12). Knockout studies *in vivo* showed that αAc knockout (i.e. αBc only) mice developed early-

onset, αBc-rich cataract (13) while αBc knockout (i.e. αAc alone) mice developed a relatively normal lens phenotype (14). The results presented in Chapter 3 fill a gap in the literature by demonstrating the changes that αBc undergoes before aggregating and note that the stability difference between αBc and αAc is exaggerated under conditions of macromolecular crowding. These findings suggest that age-related cataract associated with α-crystallin aggregation is related with post-translational modifications that cause αBc to lose stability, thereby encouraging aggregation, and/or cause αAc to reduce its chaperone ability, thereby providing ineffective prevention of αBc aggregation. Cataract-associated mutants R120G and D109A αBc destabilise the central β-sheet-rich α-crystallin domain (ACD) dimer interface and show very similar structural and functional features under dilute conditions to that of wild-type αBc under crowded conditions (15-18). These correlations suggest a common mechanism for αBc aggregation *in vivo* which involves the loss of ACD stability and structure preceding aggregation. It also suggests that mutants such as R120G and D109A will be severely impacted in a highly crowded environment, making it difficult for these αBc mutants to be chaperoned by αAc as evidenced *in vivo* by their congenital cataract phenotypes. In addition, the congenital cataract-associated mutants of αBc that have been characterised all exhibit a loss of stability despite showing other differences in structure and oligomeric state. In light of this, therapeutics that are efficacious in treating cataract by stabilising αBc and its mutants, such as sterols that bind to the ACD dimer interface (19), would seem to negate this proposed mechanism of αBc aggregation under crowded conditions.

While the formation of cataract has various mechanisms, αBc aggregation has been implicated in numerous congenital and age-related forms of cataract (13, 17, 20-22). Performing similar studies to those detailed in Chapter 3 on congenital cataract-associated mutants of αBc or αAc could provide the means to confirm the hypothesis that aggregation in a highly crowded environment such as the eye lens is largely associated with αBc while its prevention is largely the job of αAc. Given the mechanism put forth for αBc aggregation, the use of sterol-based compounds under conditions of macromolecular crowding could provide a means to not only to quell its aggregation, but also to determine if the stability of αBc is improved in this environment, as is observed for mixtures with increasing αAc. This could be useful as part of a screening assay for novel anti-cataract compounds as the stability change for αBc under crowded conditions is likely to be far more relevant to the physiological environment of the eye

lens. Further, an understanding of the crowding-induced unfolding of the αBc ACD could be important to the development of therapeutic compounds given this event precedes αBc aggregation. If crowding affects the excised αBc ACD, which forms a 20 kDa dimer (23), in a similar manner to that of the ACD in full-length αBc, then the preparation of ^{15}N, ^{13}C labelled ACD and characterisation by NMR spectroscopy under crowded conditions can provide an atomic understanding of the crowding-induced ACD unfolding. Finally, SANS can be employed to determine the impacts that high concentrations of β-crystallins (overall repulsive) have on αBc stability as opposed to γ-crystallins (overall attractive) which were used herein (24). These studies could also be extended to isoforms of β- or γ-crystallin with the overall aim being to provide a better understanding of the impacts of the lens make-up on αBc structure and stability. The proposed experiments provide a foundation to further expand our understanding of the role of macromolecular crowding in αBc aggregation. In addition, studies using macromolecular crowding *in vitro* can help to reconcile findings from the dilute *in vitro* and *in vivo* environment for other lens crystallins (25).

Chapter 4 investigated the use of an uncommon protein characterisation technique, CE, to reveal novel insights into protein conformational heterogeneity. CE can separate proteins in-solution based on conformation and charge with distinct differences (i.e. open and closed conformers) being resolved into separate peaks, while more subtle differences result in peak broadening. As such, the width of a CE peak or peaks can provide information on the heterogeneity of conformations. Recently, analytical techniques were developed to calculate a distribution of electrophoretic mobilties (26) (which is sensitive to protein conformation) and analyse the distribution to obtain a dispersity value i.e. an analytical measure of peak broadness (27). As this technique had not yet been applied to proteins, a proof-of-concept study was conducted on structurally different, well-characterised proteins. Both monomeric and oligomeric proteins were used to discern how CE separates oligomeric species and how this could affect the measure of conformational heterogeneity. To this end, bovine serum albumin (BSA), yeast alcohol dehydrogenase (YADH), and bovine α-lactalbumin (BLA), along with various treatments and preparations that change the proteins' conformation were separated and compared using CE. The raw data were transformed to a distribution of electrophoretic mobilities and were compared to the distribution of molar masses from common separation technique size-exclusion chromatography

coupled to multiple angle light scattering (SEC-MALS). The respective dispersity values were calculated and compared.

BSA, which contains monomeric and oligomeric forms, was compared to a largely monomeric preparation. CE separated BSA oligomers and showed that as oligomeric state increases, so does electrophoretic mobility implying larger oligomers have less electrostatic friction or more effective negative charge or a combination thereof. CE was able to elucidate two different BSA monomeric populations what were not discerned by SEC-MALS. Overall, CE showed that conformational and molar mass heterogeneity scaled similarly, indicating that dispersity derived from the distribution of electrophoretic mobilities is also sensitive to changes in protein conformation due to oligomerisation. Wild-type YADH is largely a tetramer with a small amount of monomer. However, it was found that treatment with the metal chelator EDTA removes monomer while treatment with reducing agent DTT breaks down the tetramer. The monomeric and tetrameric populations of YADH were detected in the distribution of electrophoretic mobilities based on comparisons with the distribution of molar masses and on changes induced by the aforementioned treatments. CE showed two distinct monomer conformations, most likely arising from the open and closed forms of YADH which is observed in the crystal structure (28). Multiple partially separated peaks were observed for the tetrameric species. Overall, the tetramer was more conformationally heterogeneous than the monomer consistent with the multiple peaks present in the CE profile. Results from CE experiments with YADH showed that CE was able to discern and monitor changes to YADH conformers from two different oligomeric populations during a single experiment. BLA showed two peaks, most likely from apo- and holo-BLA, when separated by CE while only one peak was observed by SEC-MALS. Comparison of dispersity values from CE indicated apo-BLA was more conformationally heterogeneous than holo-BLA, however, molar mass heterogeneity for all forms of BLA was unchanged, consistent with BLA being monomeric. Comparing the dispersity from the distributions of electrophoretic mobilities and molar masses of all proteins and treatments in this study provided a comparison of the conformational and molar mass heterogeneity between these three unrelated proteins and their treatments. The proteins that grouped together within this plot could be rationalised based on the preceding experiments and the extensive literature available.

These CE experiments and results provide a platform for the investigation of conformationally heterogeneous proteins such as the crystallins which form complex oligomeric structures, e.g. αBc and αAc, or are post-translationally modified, e.g. isolated from an *in vivo* source. Recently, microfluidic electrophoresis was used to investigate the heterogeneity of αBc in-solution by separating its dominant oligomeric species according to their different electrophoretic mobilities (29). The basis of separation (electrophoretic mobility) is the same as CE, however the timescale of separation was in the order of seconds for the microfluidic device as opposed to minutes for CE. This is an important consideration as αBc subunit exchange at room temperature is in the order of minutes (30) and to effectively resolve inter-converting species, the separation would need to occur faster than subunit exchange (31). While CE separation is comparable to the timescale of αBc subunit exchange at room temperature, this potential impediment can be overcome via reducing the temperature to slow subunit exchange to an order of hours thereby facilitating a higher resolution CE separation. However, if complete separation of αBc was not achieved by CE, the application of the dispersity of electrophoretic mobilities would allow a relative measure of the in-solution heterogeneity of αBc. CE also has the added benefit of allowing the resolving power to be further tuned through simply increasing capillary length, while a similar modification would be more difficult to achieve with a prefabricated microfluidic device. Nevertheless, the successful identification of oligomeric species of αBc in solution via their electrophoretic mobilities provides evidence that the future goal of this proof-of-concept CE study to apply CE to crystallin proteins such as αBc is feasible. In addition to αBc, the heterogeneity of other oligomeric small heat-shock proteins such as αAc and Hsp27 could also be investigated by CE, particularly in comparing modifications that change their oligomeric state, e.g. phosphorylation (30, 32).

With age, lens crystallins undergo extensive post-translational modification, particularly deamidation (33, 34). CE coupled to mass spectrometry (CE-MS) might provide a means to identify deamidated crystallins unambiguously, as the additional negative charge(s) changes the electrophoretic mobility of the protein (33). CE-MS would also enable quantitative measurement of the concentration of proteins or digested peptide fragments using UV detection rather than determining this information from the relative MS intensity (34, 35) which can be inherently inaccurate due to differences in the

ionisation efficiencies of species. In addition to numerous analytical research opportunities, and given the level of detail determined from CE protein separation, this method could potentially be translated to industrial applications. They include quality control in protein production or analysis of the long-term viability of protein therapeutics, where a high-throughput, ultra-low volume technique that is easily automated would provide complementary and orthogonal information on protein heterogeneity that is otherwise not accessible by other separation techniques e.g. SEC-MALS.

This thesis provides novel insight into a protein aggregation disorder, cataract, and a technique (CE) that yielded novel perspective on protein heterogeneity. Future work will aim to acquire an even deeper understanding of the structural changes that lens crystallins undergo with age, to aid in the development of treatments to prevent cataract. With funding for basic science on the decline in Australia, the chance that basic research will lead to groundbreaking new insights into the molecular underpinnings of disease is diminished. It is critical this is redressed so that astute investigation is allowed unencumbered to provide the linkages, some of which are detailed herein, leading to the treatment of cataract and other protein folding disorders that undoubtedly exists.

References

1. Thorn DC, Grosas AB, Mabbitt PD, Ray NJ, Jackson CJ, Carver JA. The structure and stability of the disulfide-linked γS-crystallin dimer provide insight into oxidation products associated with lens cataract formation. J. Mol. Biol. **2019**;431(3):483-97.

2. Purkiss AG, Bateman OA, Goodfellow JM, Lubsen NH, Slingsby C. The x-ray crystal structure of human γS-crystallin c-terminal domain. J. Biol. Chem. **2002**;277(6):4199-205.

3. Sagar V, Chaturvedi SK, Schuck P, Wistow G. Crystal structure of chicken γS-crystallin reveals lattice contacts with implications for function in the lens and the evolution of the βγ-crystallins. Structure. **2017**;25(7):1068-78

4. Truscott RJW. Age-related nuclear cataract - oxidation is the key. Exp. Eye. Res. **2005**;80(5):709-25.

5. Serebryany E, Yu S, Trauger SA, Budnik B, Shakhnovich EI. Dynamic disulfide exchange in a crystallin protein in the human eye lens promotes cataract-associated aggregation. J. Biol. Chem. **2018**;293(46):17997-8009.

6. Fan X, Zhou S, Wang B, Hom G, Guo M, Li B, *et al.* Evidence of highly conserved β-crystallin disulfidome that can be mimicked by in vitro oxidation in age-related human cataract and glutathione depleted mouse lens. Mol. Cell. Proteom. **2015**;14(12):3211-23.

7. Ramkumar S, Fan X, Wang B, Yang S, Monnier VM. Reactive cysteine residues in the oxidative dimerization and Cu2+ induced aggregation of human γD-crystallin: implications for age-related cataract. Biochim. Biophys. Acta - Mol. Basis Dis. **2018**;1864(11):3595-604.

8. Quintanar L, Domínguez-Calva JA, Serebryany E, Rivillas-Acevedo L, Haase-Pettingell C, Amero C, *et al.* Copper and zinc ions specifically promote nonamyloid aggregation of the highly stable human γD-crystallin. ACS Chem. Biol. **2016**;11(1):263-72.

9. Pande A, Pande J, Asherie N, Lomakin A, Ogun O, King JA, *et al.* Molecular basis of a progressive juvenile-onset hereditary cataract. Proc. Natl. Acad. Sci. U.S.A. **2000**;97(5):1993-8.

10. Domínguez-Calva JA, Pérez-Vázquez ML, Serebryany E, King JA, Quintanar L. Mercury-induced aggregation of human lens γ-crystallins reveals a potential role in cataract disease. J. Biol. Inorg. Chem. **2018**;23(7):1105-18.

11. Liang J, Sun T-X, Akhtar NJ. Heat-induced conformational change of human lens recombinant alphaA-and alphaB-crystallins. Mol. Vis. **2000**;6:10-4.

12. Srinivas P, Narahari A, Petrash JM, Swamy MJ, Reddy GB. Importance of eye lens α-crystallin heteropolymer with 3: 1 αA to αB ratio: stability, aggregation, and modifications. IUBMB Life. **2010**;62(9):693-702.

13. Brady JP, Garland D, Duglas-Tabor Y, Robison WG, Groome A, Wawrousek EF. Targeted disruption of the mouse αA-crystallin gene induces cataract and cytoplasmic inclusion bodies containing the small heat shock protein αB-crystallin. Proc. Natl. Acad. Sci. U.S.A. **1997**;94(3):884-9.

14. Brady JP, Garland DL, Green DE, Tamm ER, Giblin FJ, Wawrousek EF. αB-crystallin in lens development and muscle integrity: a gene knockout approach. Invest. Ophth. Vis. Sci. **2001**;42(12):2924-34.

15. Bova MP, Yaron O, Huang Q, Ding L, Haley DA, Stewart PL, *et al.* Mutation R120G in αB-crystallin, which is linked to a desmin-related myopathy, results in an irregular structure and defective chaperone-like function. Proc. Natl. Acad. Sci. U.S.A. **1999**;96(11):6137-42.

16. Treweek TM, Rekas A, Lindner RA, Walker MJ, Aquilina JA, Robinson CV, *et al.* R120G αB-crystallin promotes the unfolding of reduced α-lactalbumin and is inherently unstable. FEBS J. **2005**;272(3):711-24.

17. Fichna JP, Potulska-Chromik A, Miszta P, Redowicz MJ, Kaminska AM, Zekanowski C, *et al.* A novel dominant D109A CRYAB mutation in a family with myofibrillar myopathy affects αB-crystallin structure. Biochim. Biophys. Acta, Clin. **2017**;7:1-7.

18. Lyon YA, Collier MP, Riggs DL, Degiacomi MT, Benesch JLP, Julian RR. Structural and functional consequences of age-related isomerization in α-crystallins. J. Biol. Chem. **2019**;294(19):7546-55.

19. Makley LN, McMenimen KA, DeVree BT, Goldman JW, McGlasson BN, Rajagopal P, *et al.* Pharmacological chaperone for α-crystallin partially restores transparency in cataract models. Science. **2015**;350(6261):674-7.

20. Truscott RJW, Chen YC, Shaw DC. Evidence for the participation of αB-crystallin in human age-related nuclear cataract. Int. J. Biol. Macromol. **1998**;22(3):321-30.

21. Vicart P, Caron A, Guicheney P, Li Z, Prévost M-C, Faure A, *et al.* A missense mutation in the αB-crystallin chaperone gene causes a desmin-related myopathy. Nature Genet. **1998**;20(1):92-5.

22. Pras E, Frydman M, Levy–Nissenbaum E, Bakhan T, Raz J, Assia EI, *et al.* A nonsense mutation (W9X) in CRYAA causes autosomal recessive cataract in an inbred Jewish Persian family. Invest. Ophth. Vis. Sci. **2000**;41(11):3511-5.

23. Hochberg GKA, Ecroyd H, Liu C, Cox D, Cascio D, Sawaya MR, *et al.* The structured core domain of αB-crystallin can prevent amyloid fibrillation and associated toxicity. Proc. Natl. Acad. Sci. U.S.A. **2014**;111(16):1562-70.

24. Tardieu A, Vérétout F, Krop B, Slingsby C. Protein interactions in the calf eye lens: interactions between β-crystallins are repulsive whereas in γ-crystallins they are attractive. Eur. Biophys. J. **1992**;21(1):1-12.

25. Huang K-Y, Kingsley CN, Sheil R, Cheng C-Y, Bierma JC, Roskamp KW, *et al.* Stability of protein-specific hydration shell on crowding. J. Am. Chem. Soc. **2016**;138(16):5392-402.

26. Chamieh J, Martin M, Cottet H. Quantitative analysis in capillary electrophoresis: transformation of raw electropherograms into continuous distributions. Anal. Chem. **2015**;87(2):1050-7.

27. Thevarajah JJ, Sutton AT, Maniego AR, Whitty EG, Harrisson S, Cottet H, *et al.* Quantifying the heterogeneity of chemical structures in complex charged polymers through the dispersity of their distributions of electrophoretic mobilities or of compositions. Anal. Chem. **2016**;88(3):1674-81.

28. Raj SB, Ramaswamy S, Plapp BV. Yeast alcohol dehydrogenase structure and catalysis. Biochemistry. **2014**;53(36):5791-803.

29. Wright MA, Ruggeri FS, Saar KL, Challa PK, Benesch JL, Knowles TP. Analysis of αB-crystallin polydispersity in solution through native microfluidic electrophoresis. Analyst. **2019**;144(14):4413-24.

30. Peschek J, Braun N, Rohrberg J, Back KC, Kriehuber T, Kastenmüller A, *et al.* Regulated structural transitions unleash the chaperone activity of αB-crystallin. Proc. Natl. Acad. Sci. U.S.A. **2013**;110(40): 3780-9.

31. Hilser VJ, Freire E. Quantitative analysis of conformational equilibrium using capillary electrophoresis: applications to protein folding. Anal. Biochem. **1995**;224(2):465-85.

32. Jovcevski B, Kelly Megan A, Rote Anthea P, Berg T, Gastall Heidi Y, Benesch Justin LP, *et al.* Phosphomimics destabilize Hsp27 oligomeric assemblies and enhance chaperone activity. Chem. Biol. **2015**;22(2):186-95.

33. Shi Y, Rhodes NR, Abdolvahabi A, Kohn T, Cook NP, Marti AA, *et al.* Deamidation of asparagine to aspartate destabilizes Cu, Zn superoxide dismutase, accelerates fibrillization, and mirrors ALS-linked mutations. J. Am. Chem. Soc. **2013**;135(42):15897-908.

34. Wilmarth PA, Tanner S, Dasari S, Nagalla SR, Riviere MA, Bafna V, *et al.* Age-related changes in human crystallins determined from comparative analysis of post-translational modifications in young and aged lens: does deamidation contribute to crystallin insolubility? J. Proteome Res. **2006**;5(10):2554-66.

35. Hains PG, Truscott RJ. Age-dependent deamidation of lifelong proteins in the human lens. Invest. Ophth. Vis. Sci. **2010**;51(6):3107-14.

"Oh, there are pictures. I keep them where I need the most cheering up."

– Homer J. Simpson

APPENDIX 1:

ROLE OF UNSTRUCTURED N- AND C-TERMINAL REGIONS IN αB-CRYSTALLIN

"If you really want something in this life, you have to work for it. Now quiet, they're about to announce the lottery numbers!"

– Homer J. Simpson

DECLARATION

The following article is peer-reviewed and published in the journal *Cell Stress and Chaperones*.

Carver JA, Grosas AB, Ecroyd H, Quinlan RA. The functional roles of the unstructured N- and C-terminal regions in αB-crystallin and other mammalian small heat-shock proteins. Cell Stress and Chaperones. 2017;22(4):627-38. doi: 10.1007/s12192-017-0789-6

The contributions of authors are as follows: Professor John Carver conceived the study. The author used software *ZipperDB*, *TANGO*, and *Zyggregtor* to predict β-sheet aggregation propensity of ten human and two *E. coli* small heat-shock protein sequences and prepared all figures. Professor John Carver wrote the publication with input from the co-authors.

Cell Stress and Chaperones
DOI 10.1007/s12192-017-0789-6

SMALL HEAT SHOCK PROTEINS

The functional roles of the unstructured N- and C-terminal regions in αB-crystallin and other mammalian small heat-shock proteins

John A. Carver[1] · **Aidan B. Grosas**[1] · **Heath Ecroyd**[2] · **Roy A. Quinlan**[3]

Received: 10 January 2017 / Revised: 6 March 2017 / Accepted: 16 March 2017
© Cell Stress Society International 2017

Abstract Small heat-shock proteins (sHsps), such as αB-crystallin, are one of the major classes of molecular chaperone proteins. In vivo, under conditions of cellular stress, sHsps are the principal defence proteins that prevent large-scale protein aggregation. Progress in determining the structure of sHsps has been significant recently, particularly in relation to the conserved, central and β-sheet structured α-crystallin domain (ACD). However, an understanding of the structure and functional roles of the N- and C-terminal flanking regions has proved elusive mainly because of their unstructured and dynamic nature. In this paper, we propose functional roles for both flanking regions, based around three properties: (i) they act in a localised crowding manner to regulate interactions with target proteins during chaperone action, (ii) they protect the ACD from deleterious amyloid fibril formation and (iii) the flexibility of these regions, particularly at the extreme C-terminus in mammalian sHsps, provides solubility for sHsps under chaperone and non-chaperone conditions. In the eye lens, these properties are highly relevant as the crystallin proteins, in particular the two sHsps αA- and αB-crystallin, are present at very high concentrations.

Keywords AlphaB-crystallin · Small heat-shock proteins · Molecular chaperone · Structure · Function · Unstructured regions

Abbreviations

ACD	α-Crystallin domain
αAc	αA-Crystallin
αBc	αB-Crystallin
IF	Intermediate filament
NAC	Non-amyloid-β component
sHsps	Small heat-shock proteins

Introduction

The intimate relationship between a protein's structure and its function is a basic tenet of biology (Bagowski et al. 2010; Worth et al. 2009). Major advances in structural biology techniques over the past 20 or so years have led to the determination of the higher order structures of a wide range of globular proteins. However, amongst the most elusive of proteins whose structures have yet to be fully elucidated are the mammalian small heat-shock proteins (sHsps). Recent progress in this quest has, however, been significant and is summarised in a variety of review articles (Bakthisaran et al. 2015; Basha et al. 2012; Hochberg and Benesch 2014; Treweek et al. 2015). X-ray crystallographic studies have determined the atomic-level structure of the excised α-crystallin domain (ACD) of sHsps which encompasses approximately the central 80 amino acids of the protein. Solid state NMR studies have also provided such information about the ACD, in addition to some detail about the structural arrangement of the N-terminal region (Jehle et al. 2009; Jehle et al. 2011). However, the atomic-level structure of the N- and C-terminal regions and their relationship to the

Electronic supplementary material The online version of this article (doi:10.1007/s12192-017-0789-6) contains supplementary material, which is available to authorized users.

✉ John A. Carver
john.carver@anu.edu.au

[1] Research School of Chemistry, The Australian National University, Acton, ACT 2601, Australia

[2] School of Biological Sciences and the Illawarra Health and Medical Research Institute, University of Wollongong, Wollongong, NSW 2522, Australia

[3] Department of Biosciences, Durham University, Durham DH1 3LE, UK

Published online: 08 April 2017

ACD in determining the overall quaternary arrangement have proved refractory to accurate determination. Mass spectrometry, X-ray solution scattering, cryo-electron microscopy and molecular modelling have also provided important structural information that, when combined, has enabled the construction of a variety of models for the oligomeric structure of αB-crystallin (αBc, HspB5), the principal sHsp (Baldwin et al. 2011b; Jehle et al. 2010). However, a detailed understanding of the structural and functional roles of the N- and C-terminal flanking regions of mammalian sHsps represents a significant gap in our knowledge.

sHsps are ubiquitous and numerous intracellular proteins and are one of the major classes of molecular chaperone proteins. Under conditions of stress (elevated temperature, infection, oxidation, etc.), they selectively interact, in an ATP-independent manner, with target proteins that are destabilised, for example, to form intermediate states that are prone to association and subsequent large-scale aggregation. Accordingly, sHsps exhibit specificity for target proteins that enter off-folding pathways towards either an amorphous or amyloid fibrillar aggregated form (Fig. 1) (Kulig and Ecroyd 2012; Treweek et al. 2015). As a result, sHsp levels are up-regulated markedly under cellular stress conditions, although significant constitutive expression can occur in some cell types presumably to regulate the correct conformation and levels of target proteins. Accordingly, sHsps play an important role in the maintenance of cellular protein homeostasis or proteostasis. Indeed, in a recent study of the proteome of aged nematodes (*C. elegans*), sHsp levels (and not those of other molecular chaperones) were elevated markedly to compensate for enhanced general protein aggregation that occurs with ageing, highlighting the crucial role of sHsps in maintaining cellular proteostasis (Walther et al. 2015).

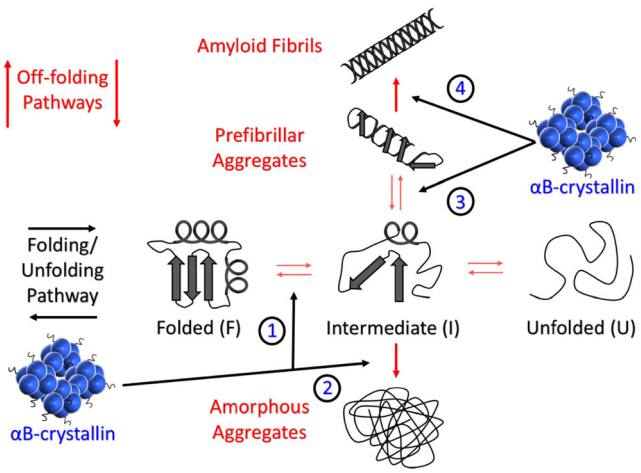


Fig. 1 Schematic of the protein folding/unfolding pathway (*horizontal*) and the off-folding pathways (*vertical*). The folding/unfolding pathway depicts both folded (*F*) and unfolded (*U*) states whilst the intermediate (*I*) represents the partially folded state(s) between these two extremes which is characterised by having elements of secondary structure and has the potential to enter the off-folding pathways to form amorphous aggregates and/or amyloid fibrils. The oligomeric sHsp αBc is depicted with small black 'squiggly' protrusions representing the solvent-exposed and flexible C-terminal extension. The various junctures at which αBc interacts on the folding/unfolding and off-folding pathways (Cox et al. 2014; Treweek et al. 2015) are indicated with *black arrows* labelled *1*, *2*, *3* and *4*. Briefly, *1*. Interaction with destabilised native-like species, *2*. Interaction with intermediate species to prevent or delay amorphous aggregation, *3*. Interaction with intermediate species to prevent or delay the generation of prefibrillar aggregates and also to dissociate prefibrillar/oligomeric species back to an intermediate or native state, *4*. Binding to amyloid fibrils in order to stabilise them and prevent further elongation and fragmentation which may lead to secondary nucleation

In humans, there are ten sHsps. The most populous and widespread (in terms of tissue distribution) of the sHsps is αBc which is also found at high levels, along with its closely related partner αA-crystallin (αAc, HspB4), in the eye lens. Both sHsps naturally co-associate and function in a chaperone manner to maintain lens transparency via inhibition of crystallin protein aggregation. Most of the recent sHsp structural work has been undertaken on αBc and it will be the focus of much of the discussion in this paper.

sHsps, in particular the mammalian ones, are difficult to characterise structurally because of their heterogeneous and dynamic nature. Such behaviour is not conducive to their crystallisation and hence structural characterisation by X-ray crystallography. For example, αBc has a monomeric subunit mass of around 20 kDa but exists as an ensemble of oligomers (Aquilina et al. 2003) with a mass distribution of 420 to 980 kDa and an average mass of 650 kDa under physiological conditions (Haley et al. 1998). The conserved central ACD in sHsps is highly β-sheet in character with the β-strands arranged in an immunoglobulin-like fold (Bagneris et al. 2009; Laganowsky et al. 2010). The ACD is flanked by N- and C-terminal regions that are variable in length and lack sequence similarity between sHsps. Apart from short sections of transient helical and β-sheet structure in the N-terminal region (Jehle et al. 2011), the flanking regions adopt little ordered secondary structure and exhibit significant dynamism. Indeed, the last 12 amino acids of αBc (along with similar regions in other mammalian sHsps) have been recognised for over 20 years as a C-terminal extension that is unstructured and has great flexibility, comparable to isolated peptides of the same length (Carver 1999; Carver et al. 1992; Esposito et al. 1998). Figure 2 provides a schematic of the various structural regions in αBc.

The polydisperse nature of αBc (and some other mammalian sHsps) is intimately related to the exchange of individual αBc subunits, a process that is highly temperature dependent with subunit exchange rate increasing significantly at higher temperatures (Baldwin et al.

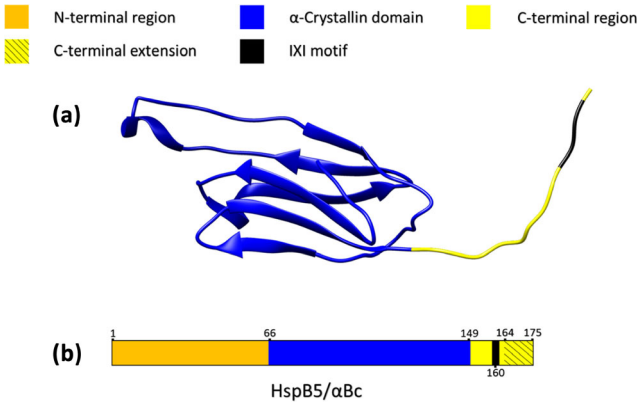


Fig. 2 Schematic diagram of the three structural regions in αBc. The colour scheme for (**a**) and (**b**) is indicated at the top of the figure. (**a**) The crystal structure of a truncated αBc monomer (PDB: 3L1G) incorporating residues 68–162 which is the majority of the structured ACD and part of the unstructured C-terminal region including the conserved IXI sequence (I159-P160-I161). The highly flexible C-terminal extension of 12 amino acids and the N-terminal region were not included in order to facilitate crystallisation (Laganowsky et al. 2010). (**b**) A linearised αBc sequence displaying the relative lengths of the ACD and the flanking N- and C-terminal regions as well the location of the conserved IXI motif and the flexible C-terminal extension

2011c; Bova et al. 1997; Hilton et al. 2013; Sobott et al. 2002). Much insight into the mechanism of subunit exchange has come from recent combined structural investigations of αBc using X-ray crystallography, NMR spectroscopy and mass spectrometry (Baldwin et al. 2011a; Baldwin et al. 2011c; Hochberg and Benesch 2014). The important role of the conserved IXI sequence in the C-terminal region (I159-P160-I161 in αBc) in interacting with an adjacent subunit is now well recognised. Subunit exchange may play an important role in oligomeric sHsp chaperone action by facilitating dissociation from the oligomer and interaction with the target protein.

The role of the various sHsp regions in chaperone action is unclear, more so of late with the observation that the isolated ACD of αBc has significant chaperone ability to prevent the aggregation of target proteins (Hochberg et al. 2014), implying that the chaperone activity (and hence target protein interaction site(s)) is encompassed within this central domain. However, other studies have implicated the N- and C-terminal regions of sHsps in interaction and binding with amorphously aggregating target proteins during chaperone action (Mainz et al. 2015; McDonald et al. 2012; Rajagopal et al. 2015). The obvious question to ask, therefore, is what are the structural and functional roles of the N- and C-terminal flanking regions in sHsps? In this paper, we address this question in a general context in terms of the role of unstructured regions in proteins and specifically in relation to flanking regions in sHsps. The role of these regions has been discussed previously in a review of existing structural data (Sudnitsyna et al. 2012).

Do the flanking regions in sHsps facilitate initial target protein interaction and act as localised crowding agents to regulate interactions with target proteins?

Earlier work by Hall (Hall 2006; Hall and Dobson 2006) examined the effect of conformational changes inherent to an inert biopolymer, through either association/dissociation or undergoing a shape transition from an expanded to a compact form, in regulating macromolecular crowding by altering the excluded volume component of the solution. It was concluded that folding of destabilised proteins was promoted under the conditions that maximised molecular crowding, i.e. when greater excluded volume of the solution occurred. In agreement with this, random polymer chains undergo significant compaction under conditions of macromolecular crowding (Le Coeur et al. 2009).

Conceivably, sHsps could utilise such means to regulate the excluded volume within the crowded environment of the cell, i.e. the highly malleable nature of their unstructured flanking regions would lead to conversion between structural compaction and expansion whilst the extensive subunit exchange would oscillate the proteins between smaller (e.g. dimer) and larger oligomeric species. Subunit exchange in sHsps may simply be a way of facilitating the initial interaction with the target protein by enabling enhanced malleability in the terminal regions, since the dynamic nature of these regions would be no longer relatively constrained within the oligomer. Within the lens fibre cells, where αAc and αBc are present at high concentrations and are by far the predominant species, these properties would be exacerbated. Phosphorylation of large oligomeric sHsps such as Hsp27 (HspB1) and αBc (HspB5) may enable this to occur as well, as has been investigated via the use of phosphomimics of these two sHsps. In these two cases, the phosphomimics have altered oligomeric size and/or mass distribution (Ecroyd et al. 2007; Peschek et al. 2013; Hayes et al. 2009) and, depending on the particular target protein and type of aggregation (amorphous or fibrillar), they exhibit enhanced sHsp chaperone ability (Ecroyd et al. 2007; Jovcevski et al. 2015).

The unfolded flanking regions, particularly when they are associated to form large heterogeneous oligomers as in the mammalian sHsps, would increase molecular crowding in the vicinity of, and when interacting with, intermediately folded (I) target proteins (Fig. 1). In doing so, the sHsps could stabilise these target proteins, stop their unfolding and thereby facilitate their refolding back to the native state via transient interactions. One could conceive of this as a localised molecular crowding phenomenon arising from the close proximity of the two proteins. Hall (2006; Hall and Dobson 2006) has shown that increasing the concentration of a partially folded crowding agent (e.g. a protein) leads to greater structure in the crowding agent, a process that could be applicable to how sHsps function in their

initial interaction to stabilise aggregation-prone target proteins. Thus, the unstructured terminal regions of sHsps initially act as akin to a 'lasso' to capture the unfolding target protein. The subsequent step of compaction of the sHsp and interaction of the target protein with the structured ACD leads to more intimate association of the two proteins and stabilisation of the intermediately folded target protein. The rationale above provides an explanation for the observation that the ACD is all that is required for the chaperone action of αBc in vitro (Hochberg et al. 2014). Of course, this is an artificial and simple system compared to the crowded nature of the cell where numerous competing interactions are possible with a diversity of cellular components. Finally, interaction and binding of the intermediately folded target protein with the sHsp during chaperone action, and its subsequent refolding, would couple folding to binding (Ganguly and Chen 2011; Shammas et al. 2016). Thereby, sHsps, either as individuals or in partnership with each other and other molecular chaperones (the latter potentially also utilising ATP hydrolysis), would contribute to the maintenance of cellular proteostasis (Jeng et al. 2015).

The importance of the N-terminal region of αBc in capturing amorphously aggregating lysozyme was demonstrated by Mainz et al. (2015). They used a combination of truncation mutants and chaperone assays to show that truncation of the N-terminal region leads to a marked loss of chaperone activity. Furthermore, specific interactions of the N-terminal region in intact αBc were inferred by solid-state NMR through chemical shift changes and alterations in dynamics of resonances in the N-terminal region (Mainz et al. 2015). Similarly, the N-terminal region of Hsp20 (HspB6) has multiple sites of interaction with a target protein as well as a role in regulating chaperone activity (Heirbaut et al. 2014) whilst also being important in the formation of a hetero-oligomer with Hsp27 (Heirbaut et al. 2016). There is evidence from interactome studies that the N-terminal region of plant sHsps is involved in interacting with amorphously aggregating target proteins (Jaya et al. 2009). Another example of the unstructured N-terminal region of sHsps' involvement in interacting with other proteins comes from the recent work of Sluchanko et al. (2017) who determined the crystal structure of a complex between a phosphorylated form (at Ser16) of the dimeric sHsp, Hsp20 (van de Klundert et al. 1998; Weeks et al. 2014) and the 14-3-3σ dimer, i.e. the two proteins form a 2:2 complex. Phosphoserine 16 in Hsp20 interacts with the binding grove of 14-3-3σ via a long loop containing the N-terminal region of the former protein. The dimeric ACD region of Hsp20, in an immunoglobulin fold conformation, binds in an asymmetric manner to one of the 14-3-3σ monomers.

One key target for sHsps is the intermediate filament (IF) cytoskeleton as evidenced by the range of diseases (cataract, myopathies, neuropathies), caused by mutations in Hsp27, HspB3, αAc, αBc and Hsp22 (HspB8) (Perng and Quinlan 2015), which in all cases cause characteristic histopathological aggregates into which IFs are also concentrated. The mutations span the primary sequence of the sHsps involved, including the N- and C-terminal regions, but there does not seem to be any clustering. Suffice to say that when the C-terminal region is completely removed, as with the cardiomyopathy-causing mutation Q151X in αBc, it is as, if not more, efficient in binding to desmin filaments and also in modulating their assembly. Indeed, data from pin array studies show that IF proteins are bound by multiple sequences throughout αBc (Ghosh et al. 2007). There also appears to be no particular requirement for phosphorylation of sHsps for them to associate with IFs (Nicholl and Quinlan 1994). IF proteins all possess intrinsically disordered domains located on the filament surface (Herrmann and Aebi 2016) and the possibility of synergy (Landsbury et al. 2010) with similarly structured N- and C-terminal regions in sHsps when bound to the filaments has not been explored. Germane to this discussion is that IFs also provide binding sites for other molecular chaperones such as Hsp70 (Perng et al. 1999) as well as the proteasome (Olink-Coux et al. 1994), so the proteostatic machinery is appropriately partitioned on IFs, structures that are integral to the cellular stress response.

The sHsps interact with all elements of the cytoskeleton (Landsbury et al. 2010; Quinlan 2002). Microfilaments, microtubules and IFs are all dependent upon sHsps for their competence (Quinlan 2002), and all are modulated by them. There appear to be multiple binding sequences across the primary sequences of all the sHsps, both in the ACD and in the N- and C-terminal regions (Ghosh et al. 2007). Indeed, deletion of the N-terminal region does not prevent sHsps from binding to actin (Guo and Cooper 2000) or to tubulin and from chaperoning microtubules (Ohto-Fujita et al. 2007). Whilst there is evidence that sequences from the N-terminal region and the ACD of both Hsp27 and αBc are effective inhibitors of actin assembly in vitro (Wieske et al. 2001), the topic is contentious since Hsp27 mutants can stimulate actin polymerisation (Butt et al. 2001) whereas another study using wild-type Hsp27 reported little to no significant change (Graceffa 2011). A common theme emerges from these studies: the assembly and dynamics of all three major cytoskeletal elements in the cell are modulated by sHsps, including Hsp27, αBc, HspB7 and Hsp22. Whilst studies sometimes focus on one specific element of the cytoskeleton (Almeida-Souza et al. 2011; Shimizu et al. 2016), it is obvious that both the competence and the integration of the different elements of the cytoskeleton rely on sHsps.

Do the unstructured flanking (terminal) regions prevent deleterious aggregation of the structured, central α-crystallin domain?

Hall (Hall and Hirota 2009; Hall et al. 2005) and Abeln and Frenkel (Abeln and Frenkel 2008; Abeln and Frenkel 2011) have examined the effect of unstructured flanking polypeptide regions on the aggregation propensity (to form both amyloid fibrillar and amorphous aggregates) of central regions. They conclude that the flanking regions have a marked propensity to prevent the central regions from aggregating; they do so by 'frustrating the encounter event' that, of course, is the crucial event in the aggregation process. Furthermore, the presence of the flanking regions on both the N- and C-terminal ends prevents aggregation to an enhanced degree, i.e. the location of the aggregating region in the middle of an unstructured polypeptide chain is most advantageous for the suppression of aggregation.

We have undertaken a survey of the regions of the ten human sHsps (HspB1 to HspB10) with a propensity to form amyloid fibrils via the algorithm ZipperDB, which determines the presence of so-called amyloid zipper sequences within the amino acid sequence of a particular protein (Goldschmidt et al. 2010). Figure 3 (along with Fig. S1 and Supplementary Table 1) summarises the results of these analyses for the sHsps. It is readily apparent that all sHsps contain significant regions of fibril-forming propensity that is mainly found in their ACD. In general, in silico analysis with other fibril prediction algorithms (TANGO and Zyggregator) (Fernandez-Escamilla et al. 2004; Tartaglia et al. 2008) gives similar results to the ZipperDB analysis (Fig. 3 and Fig. S1) in implying a significant tendency for the ACD to have more fibril-prone residues than the two other (terminal) regions (Supplementary Table 1).

Specifically, the ACD of αAc and αBc have large portions of their sequences that are predicted to form amyloid fibrils (18.1 and 20.2% respectively via ZipperDB analysis, Fig. 3, Supplementary Table 1). There is experimental evidence to support this. The isolated peptide encompassing K70 to K88 in αAc (and the corresponding region in αBc, D73 to K92) has marked chaperone ability to prevent amorphous target protein aggregation (Sharma et al. 2000). The peptide has been named 'mini-chaperone' by Sharma and co-workers. In addition, F71-K88 αAc forms amyloid fibrils (Raju et al. 2016; Tanaka et al. 2008). However, addition of the last ten amino acids of αAc, i.e. the flexible C-terminal extension, to the C-terminus of the αAc 'mini-chaperone' prevents the peptide from forming amyloid fibrils but retains its chaperone ability (Raju et al. 2014). Furthermore, Laganowsky et al. found that the 11-amino acid fragment K90-V100 in αBc (encompassing a loop region between two strands of antiparallel β-sheet) was highly amyloidogenic based on ZipperDB and experimental analyses (Laganowsky et al. 2012). Indeed, K90-V100, readily formed classic amyloid fibrils, in addition to a β-sheet oligomer whose structure was determined by X-ray crystallography. Furthermore, the oligomer was cytotoxic.

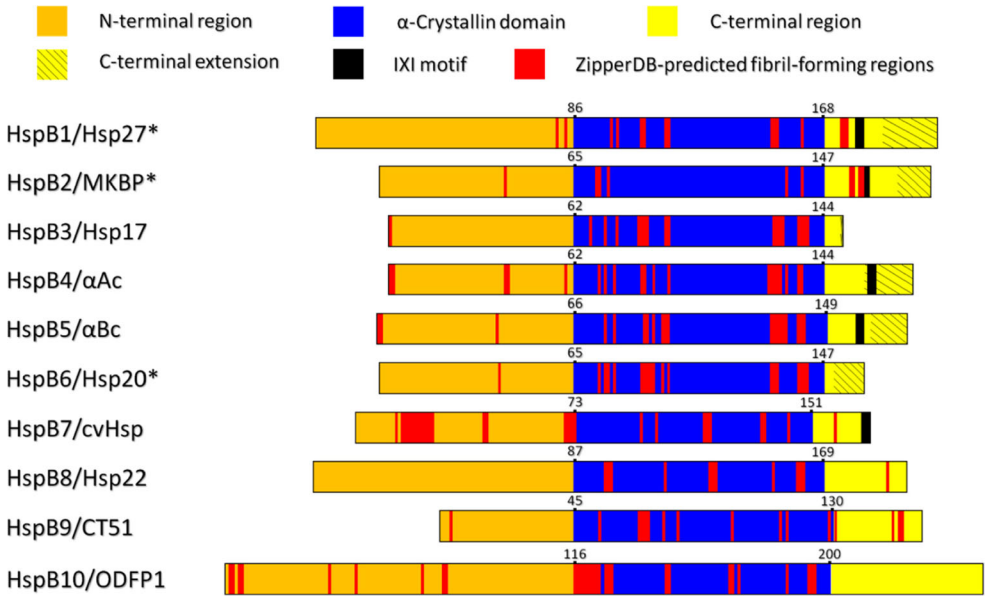


Fig. 3 ZipperDB analysis (Goldschmidt et al. 2010) of the amino acid sequences of the ten human sHsps displayed as linearised sequences aligned at the N-terminal end of the ACD. The *colour scheme* is indicated at the *top* of the figure. The position of the published IXI sequences (Delbecq et al. 2015) and the C-terminal extensions (Carver 1999) are displayed in the C-terminal region. The *red regions* are those residues associated with the start of a hexapeptide that has a high propensity to form amyloid fibrils as determined by ZipperDB, i.e. these residues have a Rosetta energy that is less than or equal to the threshold energy of −23 kcal/mol. (*)The C-terminal extensions for HspB1/Hsp27, HspB2/MKBP and HspB6/Hsp20 were identified from ^{1}H NMR spectroscopic studies of sHsps from other mammals, i.e. mouse, rat and rat respectively

Our work with αAc and αBc has shown that they form amyloid fibrils under slightly destabilising conditions, for example, in the presence of low concentrations of denaturant and elevated temperature (Meehan et al. 2004; Meehan et al. 2007). The *E. coli* sHsp IbpA forms amyloid fibrils under physiological conditions in vitro, which is prevented by the presence of its co-chaperone IbpB (Ratajczak et al. 2010). IbpB, the other *E. coli* sHsp, shares 48% sequence identity with IbpA but does not form fibrils under physiological conditions. The ACD of IbpA is slightly more aggregation-prone than the ACD of IbpB by ZipperDB analysis (Supplementary Table 1).

Furthermore, as stated above, the ACD of sHsps adopts an immunoglobulin-like fold, a motif that is prone to amyloid fibril formation, possibly because of its highly β-sheet character which is primed for conversion into the amyloid fold. Thus, immunoglobulin light chains, or its fragments, form amyloid fibrils in amyloid light chain amyloidosis. Likewise, superoxide dismutase 1 and β2-microglobulin both adopt the immunoglobulin fold and are the principal components of the amyloid fibrillar deposits associated with amyotrophic lateral sclerosis and haemodialysis-related amyloidosis (Knowles et al. 2014).

Thus, the ACD of αBc (or at least part(s) of it) is prone to form fibrils which, if were also true for the intact protein, would be highly deleterious to the protein's functionality in vivo. As αBc (and other sHsps) do not form fibrils under normal physiological conditions, the implication is that the flanking regions have a modulating effect on the amyloidogenicity of the ACD.

The dynamic nature of the flanking regions, particularly the polar, flexible C-terminal extension in mammalian sHsps, acts as a solubilising agent for the protein under chaperone and non-chaperone conditions

The unstructured and dynamic nature of the N- and C-terminal regions is well recognised, particularly so for the C-terminal extension of mammalian sHsps, located at the extreme C-terminal end of the protein that has flexibility comparable to small peptides of comparable length and hence is amenable to observation in solution by NMR spectroscopy (Carver 1999; Carver and Lindner 1998; Treweek et al. 2010). Figures 2b and 3 provide a comparison of the length of the C-terminal extension observed by solution phase NMR spectroscopy for Hsp27 (HspB1) through to Hsp20 (HspB6). It is of note that the remaining four mammalian sHsps (HspB7 to HspB10) have not been studied by NMR spectroscopy to ascertain whether they also possess a flexible C-terminal extension. The structure and function of this region have been well characterised, as summarised in various reviews (Carver 1999; Carver and Lindner 1998). It is suffice to state that removal of the C-terminal extension leads to destabilisation of the protein and reduces its chaperone effectiveness (Lindner et al. 2000) whilst replacement of charged amino acids with uncharged alanine in this region leads to similar effects (Morris et al. 2008; Treweek et al. 2007). It is concluded that the C-terminal extension has an important solubilising role for at least some mammalian sHsps, and is required to offset the inherent exposed hydrophobicity, a factor that is probably of importance for the proteins' chaperone function. The same role for this extension is utilised under chaperone conditions to solubilise the complex that sHsps form with amorphously aggregating target proteins.

Discussion

In this paper, we have proposed a variety of functional roles for the N- and C-terminal regions of sHsps:

1. They regulate the interaction and stabilisation of target proteins during chaperone action via localised molecular crowding action.
2. They effectively shield the central ACD from potential aggregation to form amyloid fibrils.
3. Their dynamic nature, particularly from the C-terminal extension, acts as solubilising agents for the protein under normal physiological conditions and during chaperone action.

The N- and C-terminal flanking regions in sHsps

The role and importance of structural disorder in molecular chaperone action have been considered by others (Bardwell and Jakob 2012; Tompa and Csermely 2004; Tompa et al. 2015), along with the realisation that many proteins are unstructured in their native state, or have large regions of their polypeptide chain that are disordered (Tompa 2012). Unstructured proteins are classified as intrinsically disordered proteins (Dunker et al. 2008). Mammalian sHsps have many properties of intrinsically disordered proteins because of their mainly unstructured flanking terminal regions.

The flanking regions in mammalian sHsps share very little sequence similarity, are highly variable in length (Fig. 3) and are present in all sHsps (Kappé et al. 2003). The proposed roles of these regions are consistent with the absence of conserved sequence. Thus, all that is required are regions of polypeptide that lack structure and are flexible, malleable and predominantly hydrophilic in character. Many sequences of amino acids can satisfy these requirements.

Our data regarding the unfolding of αAc and αBc in the presence of urea showed that the ACD is more exposed to solution than the N-terminal region (Carver et al. 1993). Likewise, the highly mobile C-terminal extension is very exposed to solution (Treweek et al. 2010). Thus, for the sHsp oligomer, the unstructured C-terminal region most likely undergoes the initial interaction with target proteins, prior to more intimate association with the ACD and/or the N-terminal region, which may be coupled with subunit dissociation.

The mammalian sHsps, Hsp20 and Hsp22 (HspB8) do not form large oligomeric assemblies but exist as smaller species, for example, dimers in the case of Hsp20 (van de Klundert et al. 1998; Weeks et al. 2014; Shemetov et al. 2008). Because of their dissociated nature, it is conceivable that these sHsps largely have their N- and C-terminal regions exposed and, as such, may have some basal level of activity that supports the proteostasis network.

The well-defined oligomeric sHsps, e.g. wheat Hsp16.9 and *Methanococcus* Hsp16.5, have no flexible C-terminal extension, nor does yeast Hsp26, yet they undergo subunit exchange (Benesch et al. 2010). However, they all have the conserved IXI sequence which facilitates subunit exchange. Hence, they can still potentially act as a 'lasso' via their N- or C-terminal regions, as per mammalian sHsps, during chaperone action.

The arguments relating to unstructured regions acting as localised crowding agents could be applied to other molecular chaperones. For example, Hsp70 has large regions of disorder which could function in a similar manner to the terminal regions in sHsps in the initial interaction with an unstructured target protein, prior to the instigation of protein folding along with ATP hydrolysis. For Hsp60 (GroEL), encapsulation of the target protein within the protein cage leads to crowding of the target protein and hence facilitates folding, via an ATP-dependent mechanisms. The concept of 'molecular shields' has been proposed to account for the chaperone action of unstructured Late Embryogenesis Abundant (LEA) proteins (Chakrabortee et al. 2012). Their mode of action is via transient interactions that shield the hydrophobic regions of target proteins from association to prevent aggregation. This behaviour is comparable with that of localised molecular crowding proposed for the flanking terminal regions in sHsps. In sHsps, these transient interactions encourage the structurally destabilised target proteins, e.g. potentially amyloid fibril-forming α-synuclein, to return to its natively unfolded (intrinsically disordered) state (Cox et al. 2014; Treweek et al. 2015). From our other studies, it is apparent that the unrelated molecular chaperones, clusterin, caseins and 14-3-3ζ, all exhibit a very similar mechanism of ATP-independent sHsp-like chaperone action (Carver et al. 2003; Holt et al. 2013; Thorn et al. 2015; Williams et al. 2011). In *E. coli*, curli proteins (e.g. CsgA) form functional amyloid extracellularly. Intracellularly, specific molecular chaperones (e.g. CsgC) prevent inappropriate curli fibril formation via a sHsp-like mechanism (Taylor et al. 2016).

Under stress conditions in vivo, e.g. heat shock, large-scale protein unfolding and potential aggregation occurs. sHsps, such as αBc, are activated (which may involve structural change and/or dissociation from the oligomer to form the dimer species) to interact with and bind to destabilised target proteins to form a high molecular weight complex (Lindner et al. 1998; Stamler et al. 2005). By contrast, under non-heat shock (i.e. constitutive) conditions, transient interaction of destabilised target proteins with non-activated sHsps occurs which does not lead to complex formation (Cox et al. 2014; Kulig and Ecroyd 2012; Treweek et al. 2015). The interaction of αBc with amyloid fibril-forming proteins, e.g. α-synuclein, ataxin-3, apolipoprotein C-II, kappa-casein and β2-microgobulin, is such a situation (Cox et al. 2016; Esposito et al. 2013; Hatters et al. 2001; Rekas et al. 2004; Rekas et al. 2007; Robertson et al. 2010). The variation in sHsp chaperone mechanism depending on conditions and the degree of unfolding of the target protein is consistent with various studies. We have shown that under mild stress conditions, i.e. slightly elevated temperature, target proteins such as malate dehydrogenase and α-lactalbumin interact with the α-crystallin oligomer via complex formation that is consistent with intercalation into the porous surface of the oligomer. The target proteins are readily accessible to interaction with molecular chaperones (e.g. Hsp70) that are capable of refolding target proteins, coupled to ATP hydrolysis (Regini et al. 2010). However, under conditions of significant stress, i.e. high temperature, the target protein (in this case γE-crystallin) is inserted into the central cavity of the α-crystallin oligomer (Clarke et al. 2010). Mchaourab's work on the chaperone action of sHsps with T4 lysozyme mutants of varying stability also implies sHsp activation during chaperone action that is directly related to the degree of unfolding, and hence binding affinity, of the particular T4 lysozyme mutant (Mchaourab et al. 2002; Shashidharamurthy et al. 2005).

The central α-crystallin domain of sHsps

Goldschmidt et al. (2010) noted that fibril-forming regions in globular, structured proteins are buried and therefore not exposed to solution and any subsequent potential interaction with other similar regions. For unstructured peptides and proteins, the presence of fibril-forming regions in the middle of unstructured peptides and proteins is a general phenomenon (Goldschmidt et al. 2010). Thus, the non-amyloid-β component (NAC), fibril-forming region in α-synuclein, is embedded in the middle of the unstructured protein. Our results (Rekas et al. 2012) showed that when the first 60 amino acids of α-synuclein were absent, i.e. the region immediately N-terminal to the NAC region, fibril formation occurred rapidly.

Others have shown that deletion of portions of the C-terminal region leads to enhanced fibril formation of α-synuclein (Hoyer et al. 2004). Consistent with these data, the fibril-forming region of unstructured κ-casein is in the middle of the protein (Ecroyd et al. 2008). Likewise, the crucial fibril-forming residues of amyloid β (Glu11 to Ala21) are in the middle of the peptide (Serpell 2000). Recently, we have shown that a four amino acid tract in the centre of the sequence of SEVI, a peptide which potentiates HIV infection, is crucial in promoting fibril formation (Elias et al. 2014). Finally, the observation that addition of the C-terminal extension to the αAc 'mini-chaperone' prevents its fibril formation (Raju et al. 2014) is also consistent with the ability of flexible, unstructured peptide flanking regions to prevent core regions from fibril formation.

Oligomerisation of sHsps

The role of subunit oligomerisation, and the associated subunit exchange, in sHsp chaperone action is an unresolved matter of debate within the literature (Haslbeck et al. 2005). There is evidence that chaperone action is enhanced under conditions of faster subunit exchange, for example, at higher temperature (Carver et al. 2002). However, the cross-linked oligomeric form of α-crystallin is chaperone-active (Augusteyn 2004), as is an immobilised form of the protein (Garvey et al. 2011). One explanation for sHsp oligomerisation is that it may protect against fibril formation, for example, within the ACD, in addition to the protection provided by the unstructured flanking regions. Other protein oligomers associate for such a reason, as we have shown for the unstructured milk casein proteins in which micelle (oligomer) formation by the four unrelated caseins (either with themselves individually, or with other caseins, or all of them to form the casein micelle in milk) prevents amyloid fibril formation by κ- and α_{s2}-casein via mutual chaperone interaction and also (principally within the casein micelle) by the chaperone action of the β- and α_{s1}-caseins (Holt et al. 2013; Holt and Carver 2012; Thorn et al. 2015). Similarly, transthyretin fibril formation requires initial dissociation from a tetrameric species prior to a conformational change within the monomer which leads to an aggregation-prone intermediate (Colon and Kelly 1992). Furthermore, methionine oxidation of apolipoprotein A1 reduces its oligomerisation and leads to enhanced amyloid fibril formation of the protein (Wong et al. 2010).

Relevance to crystallin proteins in the eye lens

The short, flexible C-terminal extensions in mammalian sHsps impart heterogeneity and enhance solubility to the proteins. Via their chaperone action in the eye lens, the α-crystallin subunits also prevent aggregation and precipitation of the β- and γ-crystallins and hence lens opacification. The two α-crystallin subunits, αAc and αBc (in a 3:1 ratio in the human lens), are the predominant lens proteins. The unrelated β-crystallin subunits also have highly flexible and unstructured terminal extensions (at both termini in the basic β-crystallins, but only at the N-terminus in the acidic β-crystallins) and are oligomers (dimers to octamers). The γ-crystallins are structurally related to the β-crystallins and form similar two-domain, Greek key motif, highly β-sheet structures, but are monomers. The major γ-crystallin, γs, has a short, flexible, four amino acid N-terminal extension (Cooper et al. 1994). The other γ-crystallins do not have terminal extensions. In support of the role of unstructured, highly flexible terminal extensions in preventing aggregation of the crystallins, truncation mutants of α- and β-crystallins without terminal extensions (and parts thereof) are prone to aggregation and potential precipitation (Lampi et al. 2002; Treweek et al. unpublished results). In the same vein, deletion of the C-terminal region of αBc (i.e. removal of residues 151 to 175 inclusive) leads to insolubilisation of the protein and the formation of inclusion bodies (Asomugha et al. 2011), although the protein retains some chaperone activity despite losing much of its secondary structure, and has a reduction in its oligomeric status (Hayes et al. 2008).

Goto and co-workers have described the amorphous, glassy state of supersaturated protein solutions and compared it to the amyloid fibril state (Yoshimura et al. 2012). Their conclusions have direct relevance to the arrangement of crystallin proteins in the eye lens. The glassy protein state is present in the lens; the very high concentration of crystallin proteins (up to 400 mg/mL in the centre) is highly stable and maintains solubility (and hence transparency) for tens of years without forming crystals or amyloid fibrils. The amorphous, glassy state of crystallin protein arrays or aggregates that are responsible for lens transparency arise because the proteins are "highly flexible and various intermolecular interactions are possible" (Yoshimura et al. 2012). Specifically, with respect to the lens, the flexibility of the terminal regions in αAc and αBc, the β-crystallins and the N-terminal extension in γs-crystallin, along with extensive subunit exchange of αAc and αBc, ensure that the lens crystallin protein mixture does not crystallise or form amyloid fibrils, occurrences that would be highly deleterious to lens transparency. It is the crystallin mixture that behaves as such, because individual, isolated β- and γ-crystallin subunits readily form well-ordered crystals whose structures have been determined by X-ray crystallography (Lapatto et al. 1991; Moreau and King 2012; Norledge et al. 1996). Thus, from a simple consideration of the highly dynamic nature of the crystallin proteins and supersaturation of a concentrated protein (crystallin) solution, the transparency of the lens can be explained. Transparency occurs despite a very high lens

crystallin concentration, a situation that normally favours large-scale aggregation, for example, to form crystals or amyloid fibrils.

Concluding comments

We have proposed that the largely unstructured N- and C-terminal regions of mammalian sHsps have multi-faceted roles: (i) they perform the initial interaction with target proteins during chaperone action, (ii) they protect the structured and sHsp-defining ACD from the possibility of misfolding into potentially non-functional and toxic amyloid fibrils and (iii) because of their dynamic, polar and unstructured nature, they act as solubilising agents for sHsps under chaperone and non-chaperone conditions. Experimentally, (iii) has been shown, in general, to be correct, at least for the C-terminal extension. However, there is plenty of scope and opportunity to undertake experiments to test (i) and (ii) and thereby determine the veracity or not of these two hypotheses, and whether they could be expanded to non-mammalian sHsps.

Acknowledgements The Australian National Health and Medical Research Council is thanked for financial support via a project grant to JAC. ABG acknowledges the financial support of an Australian Postgraduate Award. We thank Prof. Yuji Goto, Osaka University, Prof. Roger Truscott, University of Wollongong and Dr. Nicholas Ray and Dr. David Thorn, Australian National University, for the helpful discussions relating to crystallin protein interactions in the eye lens. JAC is indebted to Simon Tognetti whose creative drawings and pottery inspired some of the ideas presented herein.

References

Abeln S, Frenkel D (2008) Disordered flanks prevent peptide aggregation. PLoS Comput Biol 4(12):e1000241. doi:10.1371/journal.pcbi.1000241

Abeln S, Frenkel D (2011) Accounting for protein-solvent contacts facilitates design of nonaggregating lattice proteins. Biophys J 100(3):693–700. doi:10.1016/j.bpj.2010.11.088

Almeida-Souza L, Asselbergh B, d'Ydewalle C et al (2011) Small heat-shock protein HSPB1 mutants stabilize microtubules in Charcot-Marie-Tooth neuropathy. J Neurosci 31(43):15320–15328. doi:10.1523/JNEUROSCI.3266-11.2011

Aquilina JA, Benesch JL, Bateman OA, Slingsby C, Robinson CV (2003) Polydispersity of a mammalian chaperone: mass spectrometry reveals the population of oligomers in αB-crystallin. Proc Natl Acad Sci U S A 100(19):10611–10616. doi:10.1073/pnas.1932958100

Asomugha C, Gupta R, Srivastava O (2011) Structural and functional roles of deamidation of N146 and/or truncation of NH_2- or COOH-termini in human αB-crystallin. Mol Vis 17(262):2407–2420

Augusteyn R (2004) Dissociation is not required for α-crystallin's chaperone function. Exp Eye Res 79(6):781–784. doi:10.1016/j.exer.2004.08.010

Bagneris C, Bateman OA, Naylor CE, Cronin N, Boelens W, Keep NH, Slingsby C (2009) Crystal structures of α-crystallin domain dimers of αB-crystallin and Hsp20. J Mol Biol 392(5):1242–1252. doi:10.1016/j.jmb.2009.07.069

Bagowski CP, Bruins W, te Velthuis AJW (2010) The nature of protein domain evolution: shaping the interaction network. Curr Genomics 11(5):368–376. doi:10.2174/138920210791616725

Bakthisaran R, Tangirala R, Rao CM (2015) Small heat shock proteins: role in cellular functions and pathology. BBA-Proteins Proteom 1854(4):291–319. doi:10.1016/j.bbapap.2014.12.019

Baldwin AJ, Hilton GR, Lioe H, Bagnéris C, Benesch JL, Kay LE (2011a) Quaternary dynamics of αB-crystallin as a direct consequence of localised tertiary fluctuations in the C-terminus. J Mol Biol 413(2):310–320. doi:10.1016/j.jmb.2011.07.017

Baldwin AJ, Lioe H, Hilton GR, Baker LA, Rubinstein JL, Kay LE, Benesch JL (2011b) The polydispersity of αB-crystallin is rationalized by an interconverting polyhedral architecture. Structure 19(12):1855–1863. doi:10.1016/j.str.2011.09.015

Baldwin AJ, Lioe H, Robinson CV, Kay LE, Benesch JL (2011c) αB-crystallin polydispersity is a consequence of unbiased quaternary dynamics. J Mol Biol 413(2):297–309. doi:10.1016/j.jmb.2011.07.016

Bardwell JC, Jakob U (2012) Conditional disorder in chaperone action. Trends Biochem Sci 37(12):517–525. doi:10.1016/j.tibs.2012.08.006

Basha E, O'Neill H, Vierling E (2012) Small heat shock proteins and α-crystallins: dynamic proteins with flexible functions. Trends Biochem Sci 37(3):106–117. doi:10.1016/j.tibs.2011.11.005

Benesch JL, Aquilina JA, Baldwin AJ et al (2010) The quaternary organization and dynamics of the molecular chaperone HSP26 are thermally regulated. Chem Biol 17(9):1008–1017. doi:10.1016/j.chembiol.2010.06.016

Bova MP, Ding L-L, Horwitz J, Fung BK-K (1997) Subunit exchange of αA-crystallin. J Biol Chem 272(47):29511–29517

Butt E, Immler D, Meyer HE, Kotlyarov A, Laass K, Gaestel M (2001) Heat shock protein 27 is a substrate of cGMP-dependent protein kinase in intact human platelets: phosphorylation-induced actin polymerization caused by HSP27 mutants. J Biol Chem 276(10):7108–7113. doi:10.1074/jbc.M009234200

Carver JA (1999) Probing the structure and interactions of crystallin proteins by NMR spectroscopy. Prog Retin Eye Res 18(4):431–462. doi:10.1016/S1350-9462(98)00027-5

Carver JA, Aquilina JA, Truscott RJ (1993) An investigation into the stability of α-crystallin by NMR spectroscopy; evidence for a two-domain structure. BBA-Protein Struct Mol Enzymol 1164(1):22–28. doi:10.1016/0167-4838(93)90107-3

Carver JA, Aquilina JA, Truscott RJ, Ralston GB (1992) Identification by ^{1}H NMR spectroscopy of flexible C-terminal extensions in bovine lens α-crystallin. FEBS Lett 311(2):143–149. doi:10.1016/0014-5793(92)81386-Z

Carver JA, Lindner RA (1998) NMR spectroscopy of α-crystallin. Insights into the structure, interactions and chaperone action of small heat-shock proteins. Int J Biol Macromol 22(3–4):197–209. doi:10.1016/S0141-8130(98)00017-8

Carver JA, Lindner RA, Lyon C, Canet D, Hernandez H, Dobson CM, Redfield C (2002) The interaction of the molecular chaperone α-crystallin with unfolding α-lactalbumin: a structural and kinetic spectroscopic study. J Mol Biol 318(3):815–827. doi:10.1016/S0022-2836(02)00144-4

Carver JA, Rekas A, Thorn DC, Wilson MR (2003) Small heat-shock proteins and clusterin: intra- and extracellular molecular chaperones with a common mechanism of action and function? IUBMB life 55(12):661–668. doi:10.1080/15216540310001640498

Chakrabortee S, Tripathi R, Watson M et al (2012) Intrinsically disordered proteins as molecular shields. Mol BioSyst 8(1):210–219. doi:10.1039/c1mb05263b

Clarke MJ, Artero JB, Moulin M et al (2010) Investigation of γE-crystallin target protein binding to bovine lens alpha-crystallin by

small-angle neutron scattering. BBA-Gen Subjects 1800(3):392–397. doi:10.1016/j.bbagen.2009.12.001

Colon W, Kelly JW (1992) Partial denaturation of transthyretin is sufficient for amyloid fibril formation in vitro. Biochemistry 31(36): 8654–8660. doi:10.1021/bi00151a036

Cooper PG, Carver JA, Aquilina JA, Ralston GB, Truscott RJ (1994) A ^{1}H NMR spectroscopic comparison of γ_S- and γ_B-crystallins. Exp Eye Res 59(2):211–220. doi:10.1006/exer.1994.1099

Cox D, Carver JA, Ecroyd H (2014) Preventing α-synuclein aggregation: the role of the small heat-shock molecular chaperone proteins. BBA-Mol Basis Dis 1842(9):1830–1843. doi:10.1016/j.bbadis.2014.06.024

Cox D, Selig E, Griffin MD, Carver JA, Ecroyd H (2016) Small heat-shock proteins prevent α-synuclein aggregation via transient interactions and their efficacy is affected by the rate of aggregation. J Biol Chem 291(43):22618–22629. doi:10.1074/jbc.M116.739250

Delbecq SP, Rosenbaum JC, Klevit RE (2015) A mechanism of subunit recruitment in human small heat shock protein oligomers. Biochemistry 54(28):4276–4284. doi:10.1021/acs.biochem.5b00490

Dunker AK, Oldfield CJ, Meng J et al (2008) The unfoldomics decade: an update on intrinsically disordered proteins. BMC Genomics 9(Suppl 2):S1. doi:10.1186/1471-2164-9-S2-S1

Ecroyd H, Koudelka T, Thorn DC, Williams DM, Devlin G, Hoffmann P, Carver JA (2008) Dissociation from the oligomeric state is the rate-limiting step in fibril formation by κ-casein. J Biol Chem 283(14): 9012–9022. doi:10.1074/jbc.M709928200

Ecroyd H, Meehan S, Horwitz J et al (2007) Mimicking phosphorylation of αB-crystallin affects its chaperone activity. Biochem J 401(1): 129–141. doi:10.1042/BJ20060981

Elias AK, Scanlon D, Musgrave IF, Carver JA (2014) SEVI, the semen enhancer of HIV infection along with fragments from its central region, form amyloid fibrils that are toxic to neuronal cells. BBA-Proteins Proteom 1844(9):1591–1598. doi:10.1016/j.bbapap.2014.06.006

Esposito G, Garvey M, Alverdi V et al (2013) Monitoring the interaction between β_2-microglobulin and the molecular chaperone αB-crystallin by NMR and mass spectrometry: αB-crystallin dissociates β_2-microglobulin oligomers. J Biol Chem 288(24):17844–17858. doi:10.1074/jbc.M112.448639

Esposito G, Viglino P, Fogolari F, Gaestel M, Carver JA (1998) Selective NMR experiments on macromolecules: implementation and analysis of QUIET-NOESY. J Magn Reson 132(2):204–213. doi:10.1006/jmre.1998.1430

Fernandez-Escamilla A-M, Rousseau F, Schymkowitz J, Serrano L (2004) Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. Nat Biotechnol 22(10): 1302–1306. doi:10.1038/nbt1012

Ganguly D, Chen J (2011) Topology-based modeling of intrinsically disordered proteins: balancing intrinsic folding and intermolecular interactions. Proteins: Struct, Funct, Bioinf 79(4):1251–1266. doi:10.1002/prot.22960

Garvey M, Griesser SS, Griesser HJ et al (2011) Enhanced molecular chaperone activity of the small heat-shock protein αB-crystallin following covalent immobilization onto a solid-phase support. Biopolymers 95(6):376–389. doi:10.1002/bip.21584

Ghosh JG, Houck SA, Clark JI (2007) Interactive sequences in the stress protein and molecular chaperone human αB crystallin recognize and modulate the assembly of filaments. Int J Biochem Cell Biol 39(10): 1804–1815. doi:10.1016/j.biocel.2007.04.027

Goldschmidt L, Teng PK, Riek R, Eisenberg D (2010) Identifying the amylome, proteins capable of forming amyloid-like fibrils. Proc Natl Acad Sci U S A 107(8):3487–3492. doi:10.1073/pnas.0915166107

Graceffa P (2011) Hsp27-actin interaction. Biochem Res Int 2011: 901572. doi:10.1155/2011/901572

Guo Z, Cooper LF (2000) An N-terminal 33-amino-acid-deletion variant of hsp25 retains oligomerization and functional properties. Biochem Bioph Res Co 270(1):183–189. doi:10.1006/bbrc.2000.2401

Haley DA, Horwitz J, Stewart PL (1998) The small heat-shock protein, αB-crystallin, has a variable quaternary structure. J Mol Biol 277(1): 27–35. doi:10.1006/jmbi.1997.1611

Hall D (2006) Protein self-association in the cell: a mechanism for fine tuning the level of macromolecular crowding? Eur Biophys J 35(3): 276–280. doi:10.1007/s00249-005-0016-8

Hall D, Dobson CM (2006) Expanding to fill the gap: a possible role for inert biopolymers in regulating the extent of the 'macromolecular crowding' effect. FEBS Lett 580(11):2584–2590. doi:10.1016/j.febslet.2006.04.005

Hall D, Hirota N (2009) Multi-scale modelling of amyloid formation from unfolded proteins using a set of theory derived rate constants. Biophys Chem 140(1–3):122–128. doi:10.1016/j.bpc.2008.11.013

Hall D, Hirota N, Dobson CM (2005) A toy model for predicting the rate of amyloid formation from unfolded protein. J Mol Biol 351(1): 195–205. doi:10.1016/j.jmb.2005.05.013

Haslbeck M, Franzmann T, Weinfurtner D, Buchner J (2005) Some like it hot: the structure and function of small heat-shock proteins. Nat Struct Mol Biol 12(10):842–846. doi:10.1038/nsmb993

Hatters DM, Lindner RA, Carver JA, Howlett GJ (2001) The molecular chaperone, α-crystallin, inhibits amyloid formation by apolipoprotein C-II. J Biol Chem 276(36):33755–33761. doi:10.1074/jbc.M105285200

Hayes VH, Devlin G, Quinlan RA (2008) Truncation of αB-crystallin by the myopathy-causing Q151X mutation significantly destabilizes the protein leading to aggregate formation in transfected cells. J Biol Chem 283(16):10500–10512. doi:10.1074/jbc.M706453200

Hayes D, Napoli V, Mazurkie A, Stafford WF, Graceffa P (2009) Phosphorylation dependence of Hsp27 multimeric size and molecular chaperone function. J Biol Chem 284(28):18801–18807. doi:10.1074/jbc.M109.011353

Heirbaut M, Beelen S, Strelkov SV, Weeks SD (2014) Dissecting the functional role of the N-terminal domain of the human small heat shock protein HSPB6. PLoS One 9(8):e105892. doi:10.1371/journal.pone.0105892

Heirbaut M, Lermyte F, Martin EM, Beelen S, Verschueren T, Sobott F, Strelkov SV, Weeks SD (2016) The preferential heterodimerization of human small heat shock proteins HSPB1 and HSPB6 is dictated by the N-terminal domain. Arch Biochem Biophys 610:41–50. doi:10.1016/j.abb.2016.10.002

Herrmann H, Aebi U (2016) Intermediate filaments: structure and assembly. Cold Spring Harb Perspect Biol 8(11):a018242. doi:10.1101/cshperspect.a018242

Hilton GR, Hochberg GK, Laganowsky A, McGinnigle SI, Baldwin AJ, Benesch JL (2013) C-terminal interactions mediate the quaternary dynamics of αB-crystallin. Phil Trans R Soc B 368(1617): 20110405. doi:10.1098/rstb.2011.0405

Hochberg GK, Benesch JL (2014) Dynamical structure of αB-crystallin. Prog Biophys Mol Biol 115(1):11–20. doi:10.1016/j.pbiomolbio.2014.03.003

Hochberg GK, Ecroyd H, Liu C et al (2014) The structured core domain of αB-crystallin can prevent amyloid fibrillation and associated toxicity. Proc Natl Acad Sci U S A 111(16):E1562–E1570. doi:10.1073/pnas.1322673111

Holt C, Carver JA (2012) Darwinian transformation of a 'scarcely nutritious fluid' into milk. J Evolution Biol 25(7):1253–1263. doi:10.1111/j.1420-9101.2012.02509.x

Holt C, Carver JA, Ecroyd H, Thorn DC (2013) Invited review: caseins and the casein micelle: their biological functions, structures, and behavior in foods. J Dairy Sci 96(10):6127–6146. doi:10.3168/jds.2013-6831

Hoyer W, Cherny D, Subramaniam V, Jovin TM (2004) Impact of the acidic C-terminal region comprising amino acids 109-140 on α-

synuclein aggregation in vitro. Biochemistry 43(51):16233–16242. doi:10.1021/bi048453u

Jaya N, Garcia V, Vierling E (2009) Substrate binding site flexibility of the small heat shock protein molecular chaperones. Proc Natl Acad Sci U S A 106(37):15604–15609. doi:10.1073/pnas.0902177106

Jehle S, Rajagopal P, Bardiaux B et al (2010) Solid-state NMR and SAXS studies provide a structural basis for the activation of αB-crystallin oligomers. Nat Struct Mol Biol 17(9):1037–1042. doi:10.1038/nsmb.1891

Jehle S, van Rossum B, Stout JR et al (2009) αB-crystallin: a hybrid solid-state/solution-state NMR investigation reveals structural aspects of the heterogeneous oligomer. J Mol Biol 385(5):1481–1497. doi:10.1016/j.jmb.2008.10.097

Jehle S, Vollmar BS, Bardiaux B et al (2011) N-terminal domain of αB-crystallin provides a conformational switch for multimerization and structural heterogeneity. Proc Natl Acad Sci U S A 108(16):6409–6414. doi:10.1073/pnas.1014656108

Jeng W, Lee S, Sung N, Lee J, Tsai FT (2015) Molecular chaperones: guardians of the proteome in normal and disease states. F1000Research 4(F1000 Faculty Rev):1448. doi:10.12688/f1000research.7214.1

Jovcevski B, Kelly MA, Rote AP, Berg T, Gastall HY, Benesch JL, Aquilina JA, Ecroyd H (2015) Phosphomimics destabilize Hsp27 oligomeric assemblies and enhance chaperone activity. Chem Biol 22(2):186–195. doi:10.1016/j.chembiol.2015.01.001

Kappé G, Franck E, Verschuure P, Boelens WC, Leunissen JA, de Jong WW (2003) The human genome encodes 10 α-crystallin–related small heat shock proteins: HspB1–10. Cell Stress Chaperones 8(1):53–61. doi:10.1379/1466-1268(2003)8<53:THGECS>2.0.CO;2

Knowles TP, Vendruscolo M, Dobson CM (2014) The amyloid state and its association with protein misfolding diseases. Nat Rev Mol Cell Biol 15(6):384–396. doi:10.1038/nrm3810

Kulig M, Ecroyd H (2012) The small heat-shock protein αB-crystallin uses different mechanisms of chaperone action to prevent the amorphous versus fibrillar aggregation of α-lactalbumin. Biochem J 448(3):343–352. doi:10.1042/BJ20121187

Laganowsky A, Benesch JL, Landau M et al (2010) Crystal structures of truncated alphaA and alphaB crystallins reveal structural mechanisms of polydispersity important for eye lens function. Protein Sci 19(5):1031–1043. doi:10.1002/pro.380

Laganowsky A, Liu C, Sawaya MR et al (2012) Atomic view of a toxic amyloid small oligomer. Science 335(6073):1228–1231. doi:10.1126/science.1213151

Lampi KJ, Kim YH, Bächinger HP et al (2002) Decreased heat stability and increased chaperone requirement of modified human βB1-crystallins. Mol Vis 8:359–366

Landsbury A, Perng MD, Pohl E, Quinlan RA (2010) Functional symbiosis between the intermediate filament cytoskeleton and small heat shock proteins. In: Arrigo AP, Simon S (eds) Small stress proteins and human diseases. Nova Science, New York, pp 55–87

Lapatto R, Nalini V, Bax B, Driessen H, Lindley P, Blundell T, Slingsby C (1991) High resolution structure of an oligomeric eye lens β-crystallin: loops, arches, linkers and interfaces in βB2 dimer compared to a monomeric γ-crystallin. J Mol Biol 222(4):1067–1083. doi:10.1016/0022-2836(91)90594-V

Le Coeur C, Demé B, Longeville S (2009) Compression of random coils due to macromolecular crowding. Phys Rev E 79(3 Pt 1):031910. doi:10.1103/PhysRevE.79.031910

Lindner RA, Carver JA, Ehrnsperger M et al (2000) Mouse Hsp25, a small heat shock protein. Eur J Biochem 267(7):1923–1932. doi:10.1046/j.1432-1327.2000.01188.x

Lindner RA, Kapur A, Mariani M, Titmuss SJ, Carver JA (1998) Structural alterations of α-crystallin during its chaperone action. Eur J Biochem 258(1):170–183. doi:10.1046/j.1432-1327.1998.2580170.x

Mainz A, Peschek J, Stavropoulou M et al (2015) The chaperone αB-crystallin uses different interfaces to capture an amorphous and an amyloid client. Nat Struct Mol Biol 22(11):898–905. doi:10.1038/nsmb.3108

McDonald ET, Bortolus M, Koteiche HA, Mchaourab HS (2012) Sequence, structure, and dynamic determinants of Hsp27 (HspB1) equilibrium dissociation are encoded by the N-terminal domain. Biochemistry 51(6):1257–1268. doi:10.1021/bi2017624

Mchaourab HS, Dodson EK, Koteiche HA (2002) Mechanism of chaperone function in small heat shock proteins: two-mode binding of the excited states of T4 lysozyme mutants by αA-crystallin. J Biol Chem 277(43):40557–40566. doi:10.1074/jbc.M206250200

Meehan S, Berry Y, Luisi B, Dobson CM, Carver JA, MacPhee CE (2004) Amyloid fibril formation by lens crystallin proteins and its implications for cataract formation. J Biol Chem 279(5):3413–3419. doi:10.1074/jbc.M308203200

Meehan S, Knowles TPJ, Baldwin AJ et al (2007) Characterisation of amyloid fibril formation by small heat-shock chaperone proteins human αA-, αB- and R120G αB-crystallins. J Mol Biol 372(2):470–484. doi:10.1016/j.jmb.2007.06.060

Moreau KL, King JA (2012) Protein misfolding and aggregation in cataract disease and prospects for prevention. Trends Mol Med 18(5):273–282. doi:10.1016/j.molmed.2012.03.005

Morris AM, Treweek TM, Aquilina JA, Carver JA, Walker MJ (2008) Glutamic acid residues in the C-terminal extension of small heat shock protein 25 are critical for structural and functional integrity. FEBS J 275(23):5885–5898. doi:10.1111/j.1742-4658.2008.06719.x

Nicholl I, Quinlan R (1994) Chaperone activity of alpha-crystallins modulates intermediate filament assembly. EMBO J 13(4):945–953

Norledge B, Mayr E-M, Glockshuber R, Bateman O, Slingsby C, Jaenicke R, Driessen H (1996) The X-ray structures of two mutant crystallin domains shed light on the evolution of multi-domain proteins. Nat Struct Mol Biol 3(3):267–274. doi:10.1038/nsb0396-267

Ohto-Fujita E, Fujita Y, Atomi Y (2007) Analysis of the αB-crystallin domain responsible for inhibiting tubulin aggregation. Cell Stress Chaperones 12(2):163–171. doi:10.1379/CSC-255.1

Olink-Coux M, Arcangeletti C, Pinardi F, Minisini R, Huesca M, Chezzi C, Scherrer K (1994) Cytolocation of prosome antigens on intermediate filament subnetworks of cytokeratin, vimentin and desmin type. J Cell Sci 107(Pt 3):353–366

Perng MD, Cairns L, van den IJssel P, Prescott A, Hutcheson AM, Quinlan RA (1999) Intermediate filament interactions can be altered by HSP27 and αB-crystallin. J Cell Sci 112(13):2099–2112

Perng MD, Quinlan RA (2015) The dynamic duo of small heat proteins and IFs maintain cell homeostasis, resist cellular stress and enable evolution in cells and tissues. In: Hightower LE, Tanguay RM (eds) The big book on small heat shock proteins. Springer International Publishing, New York, pp 401–434

Peschek J, Braun N, Rohrberg J et al (2013) Regulated structural transitions unleash the chaperone activity of αB-crystallin. Proc Natl Acad Sci U S A 110(40):E3780–E3789. doi:10.1073/pnas.1308898110

Quinlan RA (2002) Cytoskeletal competence requires protein chaperones. In: Arrigo AP, Müller WEG (eds) Small stress proteins. Springer, Berlin Heidelberg, pp 219–233

Rajagopal P, Tse E, Borst AJ, Delbecq SP, Shi L, Southworth DR, Klevit RE (2015) A conserved histidine modulates HSPB5 structure to trigger chaperone activity in response to stress-related acidosis. elife 4:e07304. doi:10.7554/eLife.07304

Raju M, Santhoshkumar P, Sharma KK (2016) Alpha-crystallin-derived peptides as therapeutic chaperones. BBA-Gen Subjects 1860(1 Pt B):246–251. doi:10.1016/j.bbagen.2015.06.010

Raju M, Santhoshkumar P, Xie L, Sharma KK (2014) Addition of αA-crystallin sequence 164–173 to a mini-chaperone DFVIFLDVKHFSPEDLT alters the conformation but not the

chaperone-like activity. Biochemistry 53(16):2615–2623. doi:10.1021/bi4017268

Ratajczak E, Stróżecka J, Matuszewska M, Ziętkiewicz S, Kuczyńska-Wiśnik D, Laskowska E, Liberek K (2010) IbpA the small heat shock protein from *Escherichia coli* forms fibrils in the absence of its cochaperone IbpB. FEBS Lett 584(11):2253–2257. doi:10.1016/j.febslet.2010.04.060

Regini JW, Ecroyd H, Meehan S et al (2010) The interaction of unfolding α-lactalbumin and malate dehydrogenase with the molecular chaperone αB-crystallin: a light and X-ray scattering investigation. Mol Vis 16:2446–2456

Rekas A, Adda CG, Aquilina JA et al (2004) Interaction of the molecular chaperone αB-crystallin with α-synuclein: effects on amyloid fibril formation and chaperone activity. J Mol Biol 340(5):1167–1183. doi:10.1016/j.jmb.2004.05.054

Rekas A, Ahn KJ, Kim J, Carver JA (2012) The chaperone activity of α-synuclein: utilizing deletion mutants to map its interaction with target proteins. Proteins: Struct, Funct, Bioinf 80(5):1316–1325. doi:10.1002/prot.24028

Rekas A, Jankova L, Thorn DC, Cappai R, Carver JA (2007) Monitoring the prevention of amyloid fibril formation by α-crystallin: temperature dependence and the nature of the aggregating species. FEBS J 274(24):6290–6304. doi:10.1111/j.1742-4658.2007.06144.x

Robertson AL, Headey SJ, Saunders HM, Ecroyd H, Scanlon MJ, Carver JA, Bottomley SP (2010) Small heat-shock proteins interact with a flanking domain to suppress polyglutamine aggregation. Proc Natl Acad Sci U S A 107(23):10424–10429. doi:10.1073/pnas.0914773107

Serpell LC (2000) Alzheimer's amyloid fibrils: structure and assembly. BBA-Mol Basis Dis 1502(1):16–30. doi:10.1016/S0925-4439(00)00029-6

Shammas SL, Crabtree MD, Dahal L, Wicky BI, Clarke J (2016) Insights into coupled folding and binding mechanisms from kinetic studies. J Biol Chem 291(13):6689–6695. doi:10.1074/jbc.R115.692715

Sharma KK, Kumar RS, Kumar GS, Quinn PT (2000) Synthesis and characterization of a peptide identified as a functional element in αA-crystallin. J Biol Chem 275(6):3767–3771. doi:10.1074/jbc.275.6.3767

Shashidharamurthy R, Koteiche HA, Dong J, Mchaourab HS (2005) Mechanism of chaperone function in small heat shock proteins: dissociation of the HSP27 oligomer is required for recognition and binding of destabilized T4 lysozyme. J Biol Chem 280(7):5281–5289. doi:10.1074/jbc.M407236200

Shemetov AA, Seit-Nebi AS, Gusev NB (2008) Structure, properties, and functions of the human small heat-shock protein HSP22 (HspB8, H11, E2IG1): a critical review. J Neurosci Res 86(2):264–269. doi:10.1002/jnr.21441

Shimizu M, Tanaka M, Atomi Y (2016) Small heat shock protein αB-crystallin controls shape and adhesion of glioma and myoblast cells in the absence of stress. PLoS One 11(12):e0168136. doi:10.1371/journal.pone.0168136

Sluchanko NN, Beelen S, Kulikova AA, Weeks SD, Antson AA, Gusev NB, Strelkov SV (2017) Structural basis for the interaction of a human small heat shock protein with the 14-3-3 universal signaling regulator. Structure 25(2):305–316. doi:10.1016/j.str.2016.12.005

Sobott F, Benesch JL, Vierling E, Robinson CV (2002) Subunit exchange of multimeric protein complexes. Real-time monitoring of subunit exchange between small heat shock proteins by using electrospray mass spectrometry. J Biol Chem 277(41):38921–38929. doi:10.1074/jbc.M206060200

Stamler R, Kappé G, Boelens W, Slingsby C (2005) Wrapping the α-crystallin domain fold in a chaperone assembly. J Mol Biol 353(1):68–79. doi:10.1016/j.jmb.2005.08.025

Sudnitsyna MV, Mymrikov EV, Seit-Nebi AB, Gusev NB (2012) The role of intrinsically disordered regions in the structure and functioning of small heat shock proteins. Curr Protein Pept Sci 13(1):76–85. doi:10.2174/138920312799277875

Tanaka N, Tanaka R, Tokuhara M, Kunugi S, Lee Y-F, Hamada D (2008) Amyloid fibril formation and chaperone-like activity of peptides from αA-crystallin. Biochemistry 47(9):2961–2967. doi:10.1021/bi701823g

Tartaglia GG, Pawar AP, Campioni S, Dobson CM, Chiti F, Vendruscolo M (2008) Prediction of aggregation-prone regions in structured proteins. J Mol Biol 380(2):425–436. doi:10.1016/j.jmb.2008.05.013

Taylor JD, Hawthorne WJ, Lo J et al (2016) Electrostatically-guided inhibition of Curli amyloid nucleation by the CsgC-like family of chaperones. Sci Rep 6:24656. doi:10.1038/srep24656

Thorn DC, Ecroyd H, Carver JA, Holt C (2015) Casein structures in the context of unfolded proteins. Int Dairy J 46:2–11. doi:10.1016/j.idairyj.2014.07.008

Tompa P (2012) Intrinsically disordered proteins: a 10-year recap. Trends Biochem Sci 37(12):509–516. doi:10.1016/j.tibs.2012.08.004

Tompa P, Csermely P (2004) The role of structural disorder in the function of RNA and protein chaperones. FASEB J 18(11):1169–1175. doi:10.1096/fj.04-1584rev

Tompa P, Schad E, Tantos A, Kalmar L (2015) Intrinsically disordered proteins: emerging interaction specialists. Curr Opin Struc Biol 35:49–59. doi:10.1016/j.sbi.2015.08.009

Treweek TM, Ecroyd H, Williams DM, Meehan S, Carver JA, Walker MJ (2007) Site-directed mutations in the C-terminal extension of human αB-crystallin affect chaperone function and block amyloid fibril formation. PLoS One 2(10):e1046. doi:10.1371/journal.pone.0001046

Treweek TM, Meehan S, Ecroyd H, Carver JA (2015) Small heat-shock proteins: important players in regulating cellular proteostasis. Cell Mol Life Sci 72(3):429–451. doi:10.1007/s00018-014-1754-5

Treweek TM, Rekas A, Walker MJ, Carver JA (2010) A quantitative NMR spectroscopic examination of the flexibility of the C-terminal extensions of the molecular chaperones, αA- and αB-crystallin. Exp Eye Res 91(5):691–699. doi:10.1016/j.exer.2010.08.015

van de Klundert FA, Smulders RH, Gijsen ML, Lindner RA, Jaenicke R, Carver JA, de Jong WW (1998) The mammalian small heat-shock protein Hsp20 forms dimers and is a poor chaperone. Eur J Biochem 258(3):1014–1021. doi:10.1046/j.1432-1327.1998.2581014.x

Walther DM, Kasturi P, Zheng M et al (2015) Widespread proteome remodeling and aggregation in aging *C. elegans*. Cell 161(4):919–932. doi:10.1016/j.cell.2015.03.032

Weeks SD, Baranova EV, Heirbaut M, Beelen S, Shkumatov AV, Gusev NB, Strelkov SV (2014) Molecular structure and dynamics of the dimeric human small heat shock protein HSPB6. J Struct Biol 185(3):342–354. doi: 10.1016/j.jsb.2013.12.009

Wieske M, Benndorf R, Behlke J, Dölling R, Grelle G, Bielka H, Lutsch G (2001) Defined sequence segments of the small heat shock proteins HSP25 and αB-crystallin inhibit actin polymerization. Eur J Biochem 268(7):2083–2090. doi:10.1046/j.1432-1327.2001.02082.x

Williams DM, Ecroyd H, Goodwin KL et al (2011) NMR spectroscopy of 14-3-3ζ reveals a flexible C-terminal extension: differentiation of the chaperone and phosphoserine-binding activities of 14-3-3ζ. Biochem J 437(3):493–503. doi:10.1042/BJ20102178

Wong YQ, Binger KJ, Howlett GJ, Griffin MD (2010) Methionine oxidation induces amyloid fibril formation by full-length apolipoprotein AI. Proc Natl Acad Sci U S A 107(5):1977–1982. doi:10.1073/pnas.0910136107

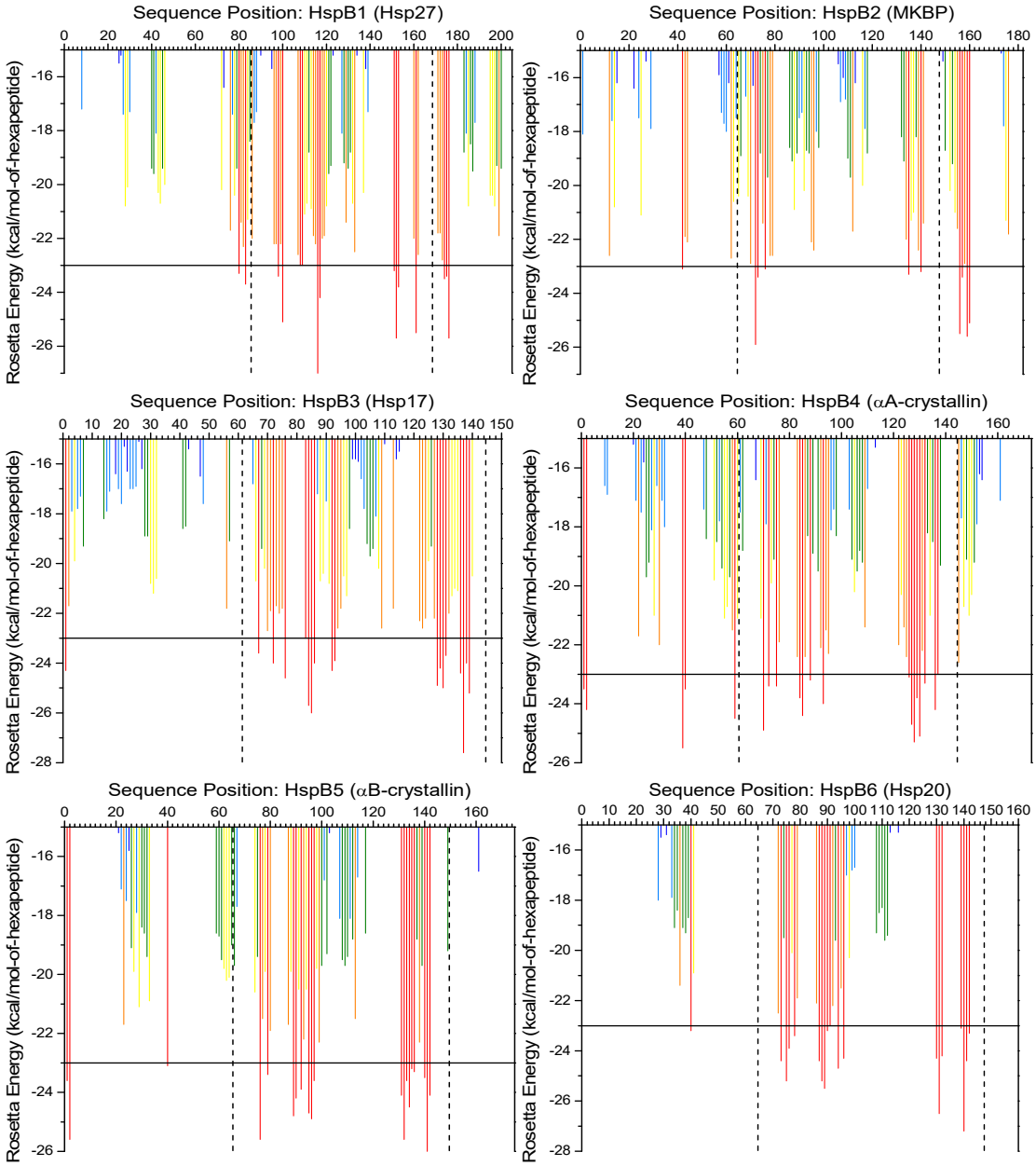
Worth CL, Gong S, Blundell TL (2009) Structural and functional constraints in the evolution of protein families. Nat Rev Mol Cell Biol 10(10):709–720. doi:10.1038/nrm2762

Yoshimura Y, Lin Y, Yagi H et al (2012) Disthingusihing crystal-like amyloid fibrils and glass-like amorphous aggregates from their kinetics of formation. Proc Natl Acad Sci U S A 109(36):14446–14451. doi:10.1073/pnas.1208228109

Supplementary Material

The functional roles of the unstructured N- and C-terminal regions in small heat-shock proteins

John A. Carver, Aidan B. Grosas, Heath Ecroyd, and Roy A. Quinlan



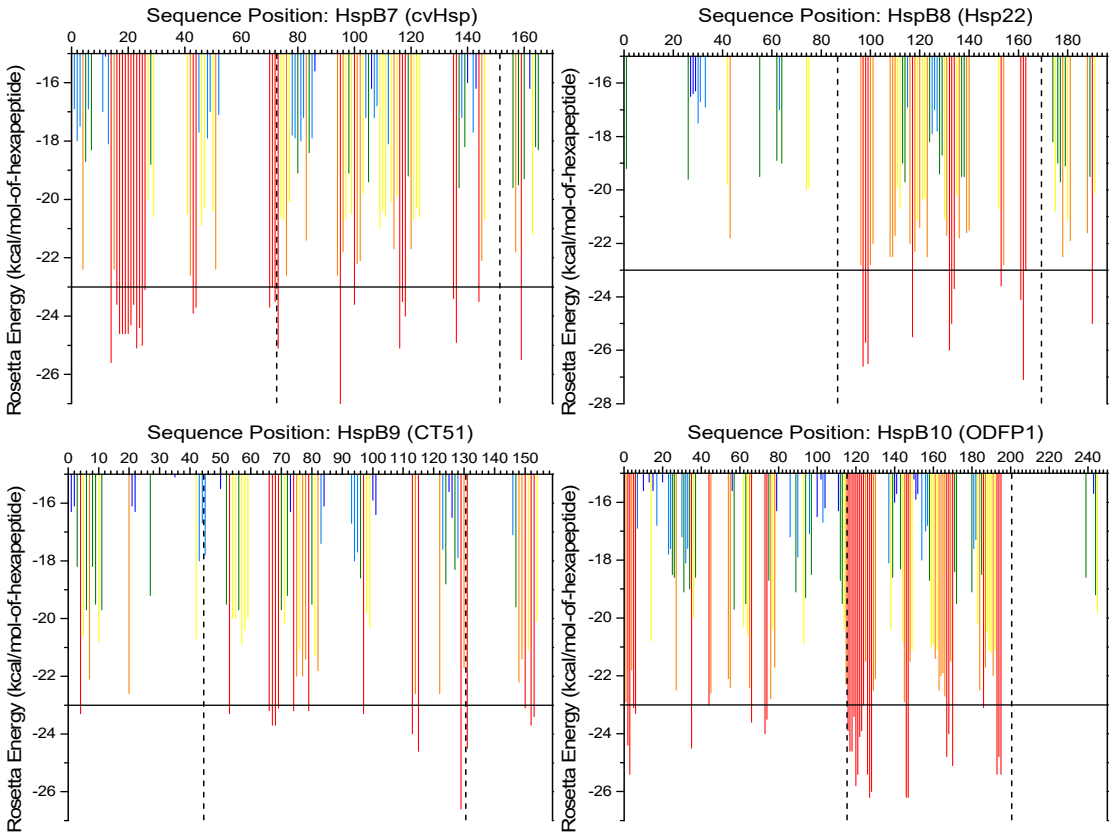


Figure S1. ZipperDB analysis (Goldschmidt et al. 2010) of the amino acid sequences of the ten human sHsps. The N-terminal, ACD, and C-terminal regions are segmented by the black vertical dashed lines in that sequential order. The blue, green, yellow and orange lines are hexapeptide residues that are of increasing Rosetta energy respectively, while the red lines that cross the threshold of -23 kcal/mol (black horizontal line) are those residues associated with a hexapeptide that has a high propensity of forming fibrils.

Supplementary Table 1. Comparison of the predicted fibril-forming or β-aggregation propensity of all ten human sHsps and two *E. coli* sHsps using three different prediction algorithms, i.e. ZipperDB (Goldschmidt et al. 2010), TANGO (Fernandez-Escamilla et al. 2004) and Zyggregator (Tartaglia et al. 2008). Percentages are given as the number of residues classified as having a 'high propensity' to form fibrils within a specific region over all the residues within that specific region of the protein, e.g. N-terminal, ACD or C-terminal regions.

Name and Abbreviation	Protein Details and Program Used	N-terminal Region	Alpha-crystallin Domain	C-terminal Region
Heat shock protein beta-1 (HspB1)/ Heat shock 27 kDa protein (Hsp27)	Amino Acid Range	1-85	86-168	169-205
	No. of residues	85	83	37
	ZipperDB (≤ -23)	2.35%	12.05%	8.11%
	TANGO (> 0)	0.00%	0.00%	0.00%
	Zyggregator (≥ 1)	5.88%	20.48%	2.70%
Heat shock protein beta-2 (HspB2)/ MDPK-binding protein (MKBP)	Amino Acid Range	1-64	65-147	148-182
	No. of residues	64	83	35
	ZipperDB (≤ -23)	1.56%	6.02%	11.43%
	TANGO (> 0)	12.50%	12.05%	20.00%
	Zyggregator (≥ 1)	4.69%	10.84%	0.00%
Heat shock protein beta-3 (HspB3)/ Heat shock 17 kDa protein (Hsp17)	Amino Acid Range	1-61	62-144	145-150
	No. of residues	61	83	6
	ZipperDB (≤ -23)	1.64%	20.48%	0.00%
	TANGO (> 0)	0.00%	42.17%	0.00%
	Zyggregator (≥ 1)	0.00%	21.69%	0.00%
Heat shock protein beta-4 (HspB4)/ Alpha-crystallin A chain, αA-crystallin, (αAc)	Amino Acid Range	1-61	62-144	145-173
	No. of residues	61	83	29
	ZipperDB (≤ -23)	8.20%	18.07%	0.00%
	TANGO (> 0)	9.84%	21.69%	0.00%
	Zyggregator (≥ 1)	21.31%	18.07%	10.34%
Heat shock protein beta-5 (HspB5)/ Alpha-crystallin B chain, αB-crystallin, (αBc)	Amino Acid Range	1-65	66-149	150-175
	No. of residues	65	84	26
	ZipperDB (≤ -23)	4.62%	20.24%	0.00%
	TANGO (> 0)	0.00%	22.62%	0.00%
	Zyggregator (≥ 1)	4.62%	16.67%	0.00%
Heat shock protein beta-6 (HspB6)/ Heat shock 20 kDa protein (Hsp20)	Amino Acid Range	1-64	65-147	148-160
	No. of residues	64	83	13
	ZipperDB (≤ -23)	1.56%	21.69%	0.00%
	TANGO (> 0)	0.00%	32.53%	0.00%
	Zyggregator (≥ 1)	0.00%	2.41%	0.00%

Heat shock protein beta-7 (HspB7)/ Cardiovascular heat shock protein (cvHsp)	Amino Acid Range	1-72	73-151	152-170
	No. of residues	72	79	19
	ZipperDB (≤ -23)	23.61%	11.39%	5.26%
	TANGO (> 0)	13.89%	22.78%	0.00%
	Zyggregator (≥ 1)	6.94%	18.99%	5.26%
Heat shock protein beta-8 (HspB8)/ Heat shock 22 kDa protein (Hsp22), (H11)	Amino Acid Range	1-86	87-169	170-196
	No. of residues	86	83	27
	ZipperDB (≤ -23)	0.00%	13.25%	3.70%
	TANGO (> 0)	0.00%	21.69%	0.00%
	Zyggregator (≥ 1)	0.00%	12.05%	18.52%
Heat shock protein beta-9 (HspB9)/ Cancer/testis antigen 51 (CT51)	Amino Acid Range	1-44	45-130	131-159
	No. of residues	44	86	29
	ZipperDB (≤ -23)	2.27%	12.79%	13.79%
	TANGO (> 0)	0.00%	13.95%	17.24%
	Zyggregator (≥ 1)	9.09%	15.12%	10.34%
Heat shock protein beta-10 (HspB10)/ Outer dense fiber protein 1 (ODFP1)	Amino Acid Range	1-115	116-200	201-250
	No. of residues	115	85	50
	ZipperDB (≤ -23)	7.83%	24.71%	0.00%
	TANGO (> 0)	6.09%	22.35%	0.00%
	Zyggregator (≥ 1)	33.91%	35.29%	4.00%
Small heat shock protein IbpA (IbpA) (*E. coli*)	Amino Acid Range	1-40	41-122	123-137
	No. of residues	40	82	15
	ZipperDB (≤ -23)	2.50%	14.63%	0.00%
	TANGO (> 0)	0.00%	40.24%	0.00%
	Zyggregator (≥ 1)	7.50%	8.54%	13.33%
Small heat shock protein IbpB (IbpB) (*E. coli*)	Amino Acid Range	1-39	40-121	122-142
	No. of residues	39	82	21
	ZipperDB (≤ -23)	7.69%	13.41%	0.00%
	TANGO (> 0)	25.64%	32.93%	0.00%
	Zyggregator (≥ 1)	10.26%	26.83%	0.00%