UNIVERSITY OF CALIFORNIA
RIVERSIDE

Mass Spectrometric Investigations of Hidden Post-Translational Modifications and the Inevitable Fate of Long-Lived Proteins

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Chemistry

by

Yana Alexandra Lyon

December 2018

Dissertation Committee:
Dr. Ryan Julian, Chairperson
Dr. Joseph Genereux
Dr. Quan Jason Cheng

The Dissertation of Yana Alexandra Lyon is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

Acknowledgments

My life has been filled with wonderful mentors and teachers that through example, encouragement, and support have molded me into the scientist and person I am today.

Firstly, I want to thank my dad, grandparents, brothers, sister-in-laws, nieces, and nephews. The distance was not easy. The trips were always too short and too infrequent. I learned to value every second I did get to be home with you and every phone call we shared. Your love and support helped through this journey, and I am forever grateful for that.

My appreciation and fascination for science began in middle school and continued to grow throughout high school. It was then challenged and tested during my chemical engineering days at Purdue. I eventually switched to chemistry and joined Dr. David McMillin's inorganic chemistry group as an undergraduate researcher. I once again felt like my fourteen-year-old self—eager to learn and excited about conducting experiments. It was in Dave's lab that I realized I wanted to be a scientist. He was patient with me, while instilling in me the passion for learning and equipping me with the fundamentals necessary to excel in grad school.

I would like to thank my PhD advisor, Dr. Ryan Julian, for allowing me to join his lab and being a great mentor to me over the years. Through his guidance I have learned that to excel in graduate school you must be a great critical thinker and problem-solver, and that the most formidable problems to solve make for interesting research and great papers. He also taught me the importance of being an effective communicator, as

both a technical writer and in public speaking. I thank him for always pushing me to be better, not just as an analytical chemist but as a person. I would also like to thank the numerous other faculty for their help and guidance throughout my time in graduate school including Dr. Cindy Larive, Dr. Wenwan Zhong, Dr. Quan Jason Cheng, and Dr. Joseph Genereux.

I would like to thank the many graduate students and post-docs that I had the pleasure to learn from and work with over the years including Dr. Yaunqi Tao, Dr. Vic Zhang, Dr. Thu Huong Pham, Dr. Omar Hamdy, Dr. Nathan Hendricks, Chris Nellesson (G'Day Mate!), James Bonner, Dylan Riggs, Tyler Lambeth, Lance Talbert, Ting Ting Wu, Jacob Silzel, and Jin Tang. Specifically, I would like to thank Omar and Yuanqi for being incredible mentors for me when I first joined the Julian lab. Both taught me how to use the instruments, how to analyze the data, and how to make Ryan happy—I thank you for that. I also thank Yuanqi for contacting me in Spring 2018 about a friend who was looking for an intern at AbbVie for the summer. Having the opportunity to work with Dr. Wenjing Ning for three months taught me not only about antibody-drug conjugates (ADCs) but how the biopharma world works. The internship experience instilled in me the drive to finish up my PhD so I can join the fast-paced, exciting world of drug discovery and development. I also want to thank Dylan Riggs for not only being a collaborator on multiple projects, but also giving me a fresh perspective and ideas about how to tackle problems I was facing with my own projects. I would also like to thank the graduate students from the Hooley group, Calvin Wiley, Dr. Lauren Holloway, Paul Bogie and Tabitha Miller, for allowing me to collaborate with them over the years and

learn how to ionize their metal-ligand cage complexes and analyze them by mass spectrometry. Finally, I would like to thank an organic chemist on the fourth floor for being my post-work, midnight dinner buddy-- and my boyfriend, I love you, Taylor Alexander. Thank you for being a constant source of love, support, and strength.

I thank the following institutions for funding: UC Riverside (RRJ), the National Science Foundation (NSF), National Institutes of Health (NIH), and the UCR Graduate Division. I also owe a tremendous amount of gratitude to Georgette Sabbah, who was an undergraduate researcher in the Julian lab for three years and contributed to the work presented in Chapters 1 and 2. The work in Chapter 3 was done with the help of Dylan Riggs and in collaboration with Dr. Justin Benesch including Dr. Miranda Collier from the University Oxford and Dr. Matteo Degiacomi from Durham University. The calculations performed in Chapter 5 were done by Dr. Gregory Beran from the UC-Riverside.

The text of this dissertation, in part or in full, is a reprint of the materials as they appear in the following publications:

Chapter 2: Lyon, Y.A., Sabbah, G.M., Julian, R.R., "Identification of Sequence Similarities among Isomerization Hotspots in Crystallin Proteins." J. Proteome Res. 2017. 16 (4), 1797- 1805

Chapter 3: Lyon, Y.A., Sabbah, G.M., Julian, R.R. "Isomer Proteomics Reveals Differences in α-Crystallin Degradation and the Role of Protein Isoaspartyl Methyltransferase (PIMT) in the Nucleus and Cortex of Human Lenses." Exp. Eye Res. 2018, 171, 131-141.

Chapter 5: Lyon, Y. A., Beran, G., Julian, R. R. Leveraging Electron Transfer Dissociation for Site-Selective Radical Generation: Applications for Peptide Epimer Analysis. J. Am. Soc. Mass Spectrom. 2017, 7, 1365-1373

Dedication

To my dad,

Bob Lyon,

who has always inspired me to have fun, be good, and learn a lot—

without you none of this would have been possible.

I love you.

ABSTRACT OF THE DISSERTATION

Mass Spectrometric Investigations of Hidden Post-Translational Modifications and the Inevitable Fate of Long-Lived Proteins

by

Yana Alexandra Lyon

Doctor of Philosophy, Graduate Program in Chemistry
University of California, Riverside, December 2018
Dr. Ryan Julian, Chairperson

Protein turnover is essential for the vitality of a functional proteome, as it allows for old and damaged proteins to be destroyed and replaced with newly synthesized copies. However, there is a growing class of proteins that escape this process and are classified as long-lived proteins (LLPs). The intrinsic instability of LLPs makes them susceptible to damage through the accumulation of unwanted post-translational modifications (PTMs), and mounting evidence suggests that this plays a critical role in aging and disease. It is known that crystallin turnover essentially stops in the mature lens fiber cells of the eye which makes it an ideal target for understanding damage due to aging, however mapping out the ensuing degradation has proven to be a formidable challenge. There are a myriad of spontaneous PTMs that accrue in the lens, including isomerization and epimerization, both of which are invisible to many traditional analytical techniques. Such small perturbations to individual amino acids may appear to be innocuous, but in fact, these hidden modifications impact protein structure, function, and solubility, all of which are crucial for lens transparency. The details of protein aging are examined herein by

employing a novel mass spectrometric dissociation technique that utilizes radical chemistry to detect sites of isomerization and epimerization. Studies on young animal eye lenses reveal how the degree of these modifications change as a function of protein solubility and shows that specific structural motifs and amino acid repeats serve as isomerization hotspots. Investigations on human donor lenses expose an increased presence of isoAspartyl residues in the nucleus of the lens compared to the cortex, suggesting decreased activity of a specific repair enzyme. Delving deeper into the structural and functional consequences of these modifications required the use of native mass spectrometry, enzyme assays, and molecular dynamics calculations, all of which help show that proper crystallin oligomerization can be disrupted by these PTMs. An alternative approach to RDD utilizing electron transfer dissociation (ETD) was also developed to detect epimerized residues in peptide models. Collectively, this work serves to better our understanding of how isomerization and epimerization contribute to the inevitable fate of LLPs.

Table of Contents

List of Figures

List of Tables

Chapter 1

Revealing Details about Protein Aging using Mass Spectrometry

1.1 Protein Structure Determines Functionality

Proteins comprise a class of macromolecules that are as complex as they are diverse. The diversity of proteins arises not only by differences in structure but also in function. Proteins serve as catalysts[1], provide structural support[2], transport molecules[3] such as oxygen,[4] control growth and differentiation,[5] protect the body against harmful invaders,[6] and act as chaperones that facilitate protein folding.[7,8] Proteins consist of long chain polypeptides made up of 20 different naturally-occurring amino acids. Protein synthesis occurs during genetic translation where three bases of messenger RNA (mRNA) constitute a codon, and each codon specifies an amino acid. Twenty naturally-occurring amino acids are then joined together through peptide bond formation to grow into a polypeptide chain until the newly-synthesized protein is eventually released by the ribosome.

The ability for proteins to take on an array of functions starts at the structural level. The amino acid sequence comprises the primary structure of the protein and provides the framework that directs the newly-synthesized protein to fold into the correct (native) structure and dictates what structural motifs will be present in the protein. Specific intramolecular interactions between amino acids, such as hydrogen bonding and salt-bridge interactions, are crucial in the formation of secondary structures such as alpha helices and beta sheets. These motifs then comprise the overall three-dimensional protein structure, commonly referred to as the tertiary structure. One example of this is a Greek

key, which consists of four adjacent antiparallel β-sheets and hairpin turns, and is common motif found in the β- and γ-crystallins that serve a structural role in the lens of the eye[9] (Figure 1.1). Finally, the protein quaternary structure is defined as the interaction of protein monomers to make large oligomeric assemblies. The crystallins are an example of proteins that make large oligomeric structures that can range from 15-mers to 50-mers, with the subunits undergoing dynamic exchange.[10] Proteins are susceptible to disruption of their delicate structure and that can lead to aberrant protein function, and in many cases, this involves the degradation of amino acid side chains and accumulation of unwanted post-translational modifications. The structural and functional perturbations induced by these modifications can lead to protein misfolding and irreversible aggregation which is believed to contribute to the pathogenesis of age-related diseases such cataracts, Alzheimer's disease, Parkinson's disease, Amyotrophic Lateral Sclerosis, and Huntington's Disease.[11,12] Understanding and predicting potential degradation sites and protein lifetimes will not only help us reveal the mechanisms by which these diseases occur, but aid in the development of therapeutics that can inhibit aggregation, enhance aggregate clearance, or block aggregation-induced cellular pathways altogether.[13]

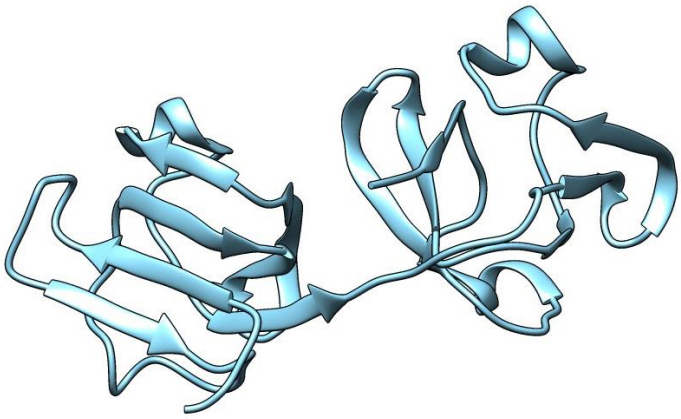


Figure 1.1 Flat ribbons denote β-sheets, and four antiparallel sheets connected by hairpin turns compose a Greek key. X-ray diffraction structure of γS crystallin from human (PDB 1HA4).

1.2 Molecular Aging and Protein Degradation

Proteins are susceptible chemical degradation that leads to changes in the delicate structure and aberrant function. The degradation can be due to environmental stressors such as UV damage[14] and oxidative stress[15,16] or by the accumulation of post-translational modifications such as glycation, phosphorylation, and methylation.[17,18] After proteins become damaged, they can either be refolded into their native conformation or they can be removed by the cell through proteostasis. The former involves the use of chaperone proteins while the latter includes the ubiquitin-proteasome and lysosome-autophagy systems.[19] The ability for cells to proteolytically degrade spent proteins allows for the average half-life of a proteins from the brain, liver, and blood to range from 3 and 9 days, based on turnover studies in mouse models.[20] The quick turnover prohibits the accumulation of unwanted modifications that could lead to protein misfolding and possibly induce aggregation. Protein turnover requires energy and the necessary organelles, which are not existent in every cell type, providing one reason for the

existence of long-lived proteins (LLPs). However, in the case of the protein cohesion which is present in metabolically active cells, the protein can last for decades before being replenished.[21] LLPs are proteins that have half-lives on the order of months to years, including nucleoporins (Nups), collagen, and crystallins.[21,22] The inability of LLPs to undergo normal proteostasis has made them targets for understanding the spontaneous damage that accrues due to aging. It has been shown that proteins undergo non-enzymatic isomerization, in which the connectivity and stereochemistry of the amino acids are changed. Specifically, aspartic acid (Asp) is known to undergo spontaneous isomerization through the formation of a five-membered succinimide ring (Figure 1.2). Asparagine (Asn) can also access the same pathways by loss of ammonia from the sidechain. Once in the ring, the typically unstable intermediate can be hydrolyzed to revert to L-Asp or become L-isoAsp where the peptide backbone has been re-routed through the sidechain. The succinimide intermediate is prone to racemization, and once D-succinimide is formed it can also be hydrolyzed to yield D-Asp or D-isoAsp. Serine (Ser) is also known to readily racemize (or epimerize) to D-Ser. While the exact mechanism remains unknown, it could be speculated that the hydroxyl moiety of the sidechain may participate in non-covalent interactions that pulls the electron density away from the alpha-carbon and, in turn, increases the acidity of the alpha-hydrogen. Isomerization of Asp or epimerization of serine are quite different than traditional post-translational modifications as they are non-enzymatic and unregulated. The accumulation of these spontaneous modifications over time is classified as molecular or protein aging. One report found that over 50% of the Asp residues in amyloid-beta (Aβ) proteins from senile

plaques of the brains of patients with Alzheimer's disease existed in the L-isoAsp form, while nearly 20% of them were D-isoAsp.[23] The brain is not the only place where isoAsp accumulates over time, it has been found that human cartilage and synovial fluid contain significant amounts of isomerized aspartic acid residues that might serve as a biomarker for osteoarthritis.[24]While these modifications may seem inconsequential or benign, in cells where protein turnover does not occur these small changes in the side chain arrangement can alter the higher-order structure of the protein and lead to aberrant protein function and degradation.

The eye lens contains some of the oldest proteins in the body, which can be attributed to the fact that mature lens fiber cells do not undergo protein turnover.[25] This has made the lens a model system for understanding protein degradation due to aging. Decades of research and attention has been focused on identifying the types of modifications that accrue in the lens with time and the list includes PTMs such as oxidation, truncation, disulfide bond formation, deamidation, and isomerization. [26,27,28] However, it has only recently been proposed that isomerization might be the most dominant modification in the human eye lens.[29] It is evident that lens proteins were not meant to last forever, as the accumulation of crystallin damage in the eye lens inevitably leads to cataracts.[30] Cataracts are the leading cause of blindness in the world[31], especially in developing countries, where surgery is not available.[32] The hallmark of this disease is loss of lens transparency due to crystallin aggregation caused by decades of damage that inevitably induces changes in protein structure and leads to loss of function accumulation of large oligomers that inhibit the proper transmission of light onto the retina.[33] However, for how

far we have come in identifying the PTMs that accumulate on crystallins, there still exists a gap in understanding how they contribute to the formation of insoluble aggregates.

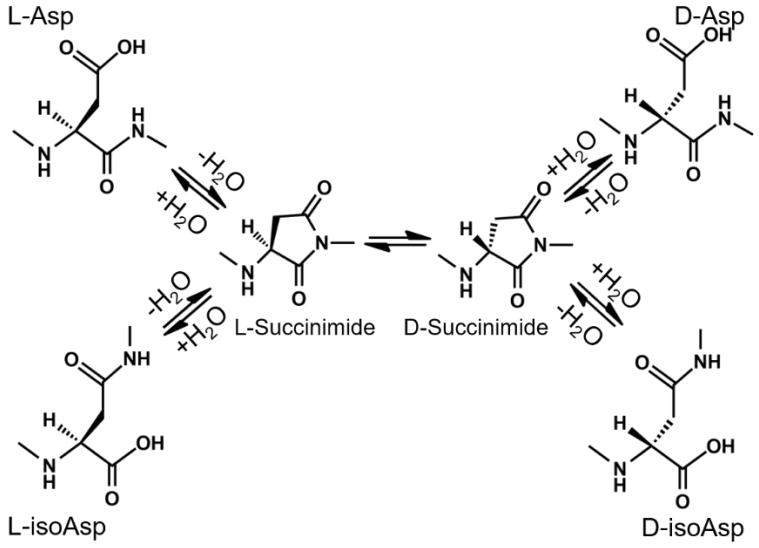


Figure 1.2 Spontaneous isomerization of aspartic acid by a five-membered succinimide intermediate.

1.3 Revealing Changes in Protein Structure and Function Induced by Isomerization

While isomerization of a single amino acid on a protein might seem to be an innocuous modification, there is a growing body of evidence that suggests otherwise. For example, isomerization of a single Asp in the complementarity-determining regions (CDR) of human monoclonal immunoglobulin γ2 (IgG2) antibody is capable of drastically reducing the antigen binding efficacy.[34] The incorporation of isoAspartyl residues in proteins is also known to trigger autoimmune responses, where the B cell response caused by isoAsp fragments of cytochrome *C* has been reported to target both the isoAsp peptides and the native protein in mice.[35] These are all examples of how isoAspartyl residues can effect protein function, but showing exactly how the

incorporation of isomerized residues leads to perturbations in protein structure has proven to be a difficult task. To study the effect different modifications have on protein structure researchers typically induce point mutations in the amino acid sequence and then use techniques such as x-ray crystallography to see how the mutation affected the protein. However, this methodology cannot be used for isomerized or racemized amino acids as they cannot be incorporated during the translation process.

One way to circumvent this is to express the native protein and then selectively modify one site on the protein. For example, it has been shown that the deamidation and isomerization of Asn32 from ribonuclease U2B (PDB 3AGN) causes an approximately 90° bend in the protein backbone and loss of the helical motif in this region (PDB 3AHS Chain B) (Figure 1.3).[36] However, the researchers struggled to get clear electron density in this region for Chain C due to disorder which is why it is represented as a dashed line in Figure 1.3 (far right). One explanation for this is that they may have an impure crystal that consists of two different Asp32 isomers, which would cause the ambiguous sidechain density in this region. Noguchi and co-workers used a similar approach again to show that isoAsp101 in hen egg-white lysozyme, one of the most extensively studied proteins by crystallography, also causes a 90° bend in the polypeptide chain.[37] It is apparent that this technique is capable of revealing the structural consequences that isoAspartyl residues have on protein structure, but in many cases the selective "aging" (and incorporation of isoAsp) at one specific site on a protein is not possible given that most proteins contain more than one Asp residue that will also undergo spontaneous isomerization. New alternatives to x-ray crystallography such as native mass

spectrometry, enzyme assays, and molecular dynamics calculations are used in the work described in Chapter 4 to reveal the ramifications that these modifications have on the structural and functional integrity of proteins.

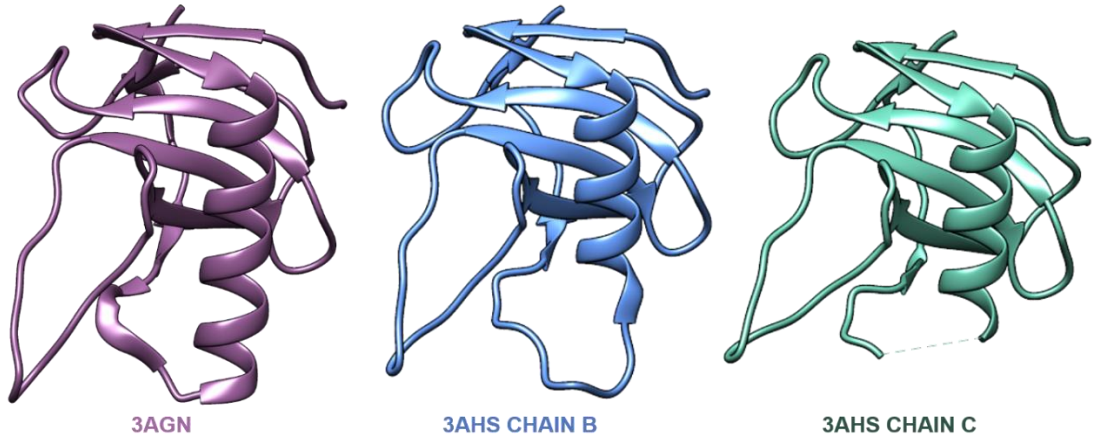


Figure 1.3 Crystal structures of ribonuclease U2B. Purple structure (PDB 3AGN) contains L-Asn 32. Blue structure in middle (PDB 3AHS) represents one of the chains (B) of the deamidated Asn32 containing isoAsp. Chain C of the deamidated version (PDB 3AHS) is shown in green. Ambiguous electron density in this region shown as a dotted line.

1.4 Mass Spectrometry as a Tool for Proteomic Analysis

There exist a variety of analytical techniques capable of characterizing proteins including nuclear magnetic resonance (NMR), UV/Vis and fluorescence spectroscopy, surface plasmon resonance (SPR), and x-ray crystallography. Each of these techniques are unique and powerful on their own but can also be used in conjunction with one another for structural elucidation and characterization. A couple drawbacks of these techniques is that they require high sample concentration as well as pure samples. In the mid-1980s electrospray ionization (ESI) was developed and provided a simple and robust way to desolvate a liquid mixture of biomolecules and analyze their mass to charge (m/z)

ratios in the gas phase. This technique has become one of the most common ionization techniques for the mass spectrometric analysis of proteins, and John Fenn was awarded a share of the Nobel Prize in Chemistry in 2002 for his work in the development of ESI. ESI-MS made it possible to quickly determine the presence of small molecules, peptides, and proteins in high-throughput samples by analysis of the mass to charge ratio (m/z) of the intact species. Now, high resolution MS (HRMS) allows for the determination of the intact mass of an analyst with a mass error sometimes less than one part per million (ppm).[38]

Intact mass analysis can be used for the rapid detection of known analytes, but oftentimes researchers need to gain more information about unknown species. In the 1980s also came the emergence of ion trap mass spectrometers thanks to work of people like Hans Dehmelt and Wolfgang Paul. Instead of the intact mass analysis, ion trap mass spectrometers create the ability to do ion-ion or ion-molecule reactions—chemistry in the gas phase. These instruments allow for tandem mass spectrometry (MS/MS), in which ions can be accelerated into three-dimensional or linear traps where collisions with buffer gases or charged molecules induce fragmentation. Some of the dissociation techniques that are used in mass spectrometry today include collision-induced dissociation (CID), electron capture/transfer dissociation (ECD/ETD), and radical directed dissociation (RDD). To denote which of the three peptide backbone sites are being cleaved during fragmentation, researchers use a, b, and c when the charge is retained on the N-terminal fragment of the peptide and x, y, z when the charge is retained on the C-terminal fragment[39] (Figure 1.2). For example, in CID, a kinetically excited precursor ion

undergoes collisions with the neutral buffer gas that lead to a conversion of the translational energy of the ion to internal energy.[40] The energized ion undergoes bond dissociation, typically of the weakest bonds in the molecule. For peptide and proteins, the amide bond is generally a low energy fragmentation pathway and this leads to the presence of b/y ions. Because peptides and proteins contain many basic residues their fragmentation pathways are also determined by the mobile proton model which states that the initial proton can move to other sites on the ion upon activation and induce dissociation.[41] The variety of dissociation techniques available allow for facile and rapid amino acid sequencing, making it the gold standard for protein sequencing today. Understanding the primary sequence of a protein is not the end of the story, we also know that protein function depends on higher order structure. These dissociation techniques also allow us to probe the gas phase structure of the ions, even if the changes induced in conformation are subtle, such as the stereoinversion of a single amino acid. However, in order to obtain detailed structural information radical techniques such as RDD and ETD are superior.

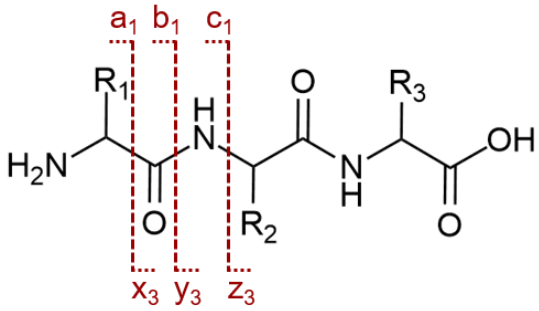


Figure 1.4 Peptide fragmentation nomenclature of a tripeptide proposed by Roepstroff and Fohlman.

1.5 Radical Methods for Peptide and Protein Analysis

1.5.1 Electron Capture/Transfer Dissociation

Electron capture dissociation (ECD) is an example of a radical-based dissociation method and involves the direct capture of an electron, typically on the sidechain of basic residues such as histidine, arginine, and lysine creating a charge reduced radical species.[42] The sidechain radical then abstracts a proton from a sterically proximate amino acid to form an aminoketyl radical that dissociates to form c/z ions.[43] A similar mechanism occurs in electron transfer dissociation (ETD), except the electron is generated by an anionic molecule, typically fluoranthene.[44] Both methods have proven invaluable in studying post-translational modifications in the gas phase as the technique is gentle enough to maintain the labile modifiers during dissociation.[45] One major drawback to ECD/ETD is that while backbone dissociation does occur after the initial capture/transfer event, the product ions are held together by noncovalent interactions leaving very little information obtained about the primary structure. Researchers have overcome this problem using tricks like activated ion electron transfer dissociation (AI-ETD)[46] which involves infrared photoactivation concurrent to ETD reactions or by performing higher energy CID (HCD) after the ETD step (EThcD).[47]

Electron capture and electron transfer dissociation (ECD/ETD) cleaves the N-Cα of the peptide backbone to create c/z ions, and these ions are shifted by +57/-57 when isoAsp is present.[48] This information has been utilized by numerous research groups to establish methods for isoAsp detection by diagnostic ions in the gas phase. However, Asp isomerization creates four isomers in total, and this technique is incapable of

distinguishing between the chiral (L vs. D) partners. The work in Chapter 5 describes a method in which hydrogen deficient peptide radicals can be made using iodo-containing peptides and the electron transfer dissociation capabilities of the Orbitrap Fusion mass spectrometer. Specifically, we show that homolytic cleavage of the C-I bond occurs when the peptide precursors are subjected to ETD, and upon subsequent HCD we can generate the same sidechain and backbone fragments that are made in RDD. Multiple isomer and epimer-containing peptides were tested by this technique and it proved successful at not only distinguishing L-Asp vs. L-isoAsp but could also distinguish the L- vs D- epimers. This novel dissociation technique provides an alternative method to detect and identify these modifications without the need of a laser-equipped mass spectrometer.

1.5.2 Radical Directed Dissociation

Radical directed dissociation (RDD), a technique developed in the Julian lab, involves the irradiation of carbon-iodine bonds, either naturally occurring on the biomolecule or added chemically by direct iodination or through an iodine-containing molecule.[49] Excitation of these chromophores with photons generated by a 266 nm Nd:Yag laser induces homolytic cleavage of the C-I bond and generation of a radical species. The beauty of the photodissociation step is that it induces very little heating of the remaining molecule leading to cleavage of a single bond and creation of a site-specific radical (Figure 1.5A). By gently activating the radical species, the radical can migrate to induce fragmentation of the backbone to yield a/x, c/z and side chain fragments (Figure 1.5B).[50] The migration of the radical is dictated by heavily by the bond dissociation energies (BDEs) for homolytically cleaving any X-H bond on the backbone

or sidechain of the amino acid.[51] If there is a sterically proximate X-H bond that has a lower BDE energy then the radical will continue to fall downhill until it reaches a thermodynamically favored location.[50]

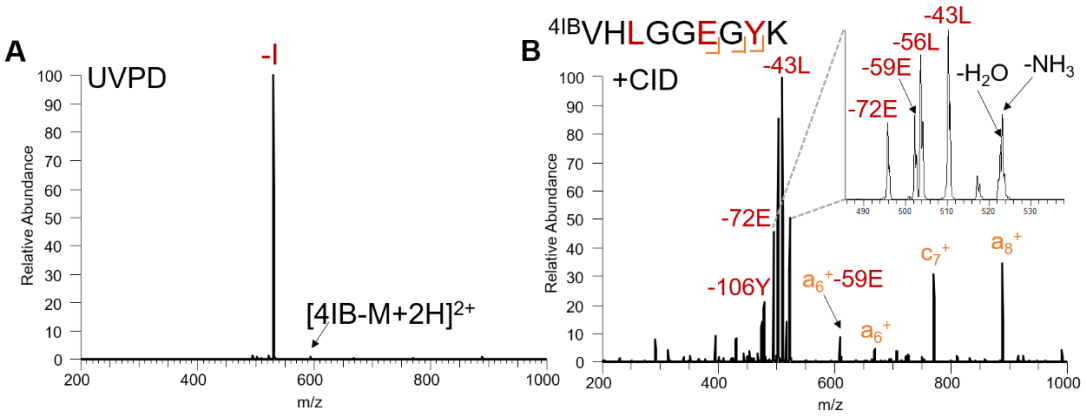


Figure 1.5 A) UVPD spectrum of doubly protonated ^{4IB}VHLGGEGYK (where 4IB denotes the *para*-iodobenzoyl chromophore and the (-I) denotes the loss of iodine due to homolytic cleavage of the C-I bond. B) Subsequent CID on the isolated radical species induces a variety of backbone fragments (orange) and side chain losses (red). The inset is zoomed in on the side chain loss m/z range.

RDD has proven to be capable of discerning isomerization of Asp[52] and epimerization (L→D) of Ser and other amino acids.[53,54] The generated radical species can probe the structural landscape of the peptide through migration to sidechains and the backbone via hydrogen-abstraction events. The differences in fragment ion intensities between isomers can be attributed to the structural perturbations that are teased out by radical. In order to perform these RDD experiments, a *para*-iodobenzoyl chromophore is often covalently attached to primary amines on the peptide through an N-hydroxysuccimide ester reaction[55] (Figure 1.6A). The C-I bond of the chromophore is homolytically cleaved, resulting in the radical species and upon collisional activation, the radical migrates and produces the fragmentation seen in Figure 1.6B. The major fragment in the L-Asp

version of the HFSPEDLTVK peptide is a γ-hydrogen abstraction from the leucine side chain to produce a -56 Da loss from the doubly protonated precursor. This fragment decreases significantly in abundance in the L-isoAsp version, and examination of low energy configurations from conformational searching with MacroModel[56] puts the leucine sidechain further away from the radical generation site. Another fragment that changes significantly between the L-Asp and L-isoAsp versions is the c7+ backbone fragment which is generated via an aminoketyl radical through the threonine sidechain. This is just one example of using RDD to distinguish isomers, and the dominant focus of this dissertation is to apply this technique to delve deeper into the isomerization tendencies of long-lived proteins to understand better understand the ramifications of these hidden modifications. The work described in Chapter 2 describes a bottom-up LC-MS/MS approach where tryptic peptides generated from the crystallins of animal lenses are probed for isomerization. While some peptide isomers separated chromatographically, others were hidden as coelutions. RDD was then employed to analyze the front and back end of each chromatographic peak to detect these coeluting isomers. While Asp isomers seemed to be more easily separated chromatographically, Ser epimers oftentimes could not be resolved. By employing RDD we mapped out the prevalence of both Asp isomerization and Ser epimerization from the crystallins of young animal lenses. RDD is also utilized in Chapter 3 where it is used to better understand the differences in isomerization tendencies between the nucleus and cortex of aged human lenses.

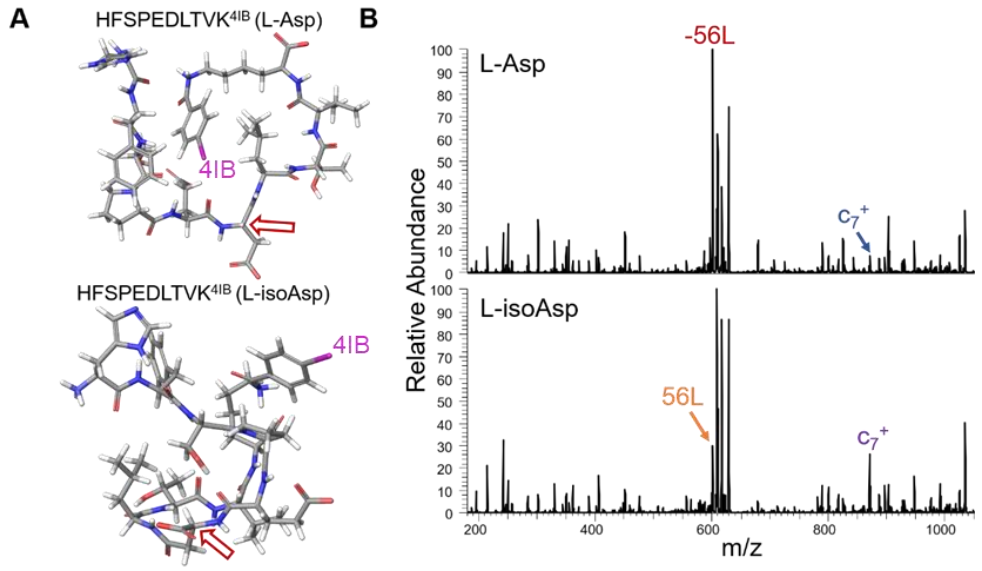


Figure 1.6 A) Structures of HFSPEDLTVK^{4IB} generated by conformation searching. Red arrows pointing to sites of L-Asp (top) and L-isoAsp (bottom). B) RDD of HFSPEDLTVK^{4IB} L-Asp and L-isoAsp yield fragment ions with significantly different intensities. The -56 loss from the leucine sidechain and the c7 backbone fragment shows the greatest changes between the two isomers.

1.6 Utility of Mass Spectrometry in Understanding Protein Aging

Isomerization remains an invisible post-translational modification to traditional analytical techniques, and it might seem counterintuitive to believe that mass spectrometry can detect a modification that leads to no change in the mass or charge of the peptide. While the full mass of the peptide is in fact useless in detecting these isobaric species, the fragmentation patterns can yield ions that are unique to a specific isomer or fragments that have significantly different intensities.[57] Collectively, the work described in this dissertation aims to provide new details into the isomerization tendencies of long-lived proteins using a variety of mass spectrometric techniques, while also developing new methodologies to dig deeper into the consequences these modifications have on the structure and function of a protein.

1. G. M. Cooper, *The Central Role of Enzymes as Biological Catalysts*, **2000**.
2. F. O. Schmitt, *Adv. Protein Chem.*, **1944**, *1*, 25.
3. H. Lodish, A. Berk and S. Zipursky, in *Molecular Cell Biology*, *2000*, 15.
4. P. Scholander, *Science*, **1960**, *131*, 585.
5. A. J. Levine, *Cell*, **1997**, *88*, 323.
6. L. L. Lu, T. J. Suscovich, S. M. Fortune and G. Alter, *Nat. Rev. Immunol.*, **2018**, *18*, 46.
7. H. Saibil, *Nat. Rev. Mol. Cell Biol.*, **2013**, *14*, 630.
8. M. Beissinger and J. Buchner, *Biol. Chem.*, **1998**, *379*, 245.
9. T. Blundell, P. Lindley, L. Miller, D. Moss, C. Slingsby, I. Tickle, B. Turnell and G. Wistow, *Nature*, **1981**, *289*, 771.
10. A. J. Baldwin, H. Lioe, C. V. Robinson, L. E. Kay and J. L. P. Benesch, *J. Mol. Biol.*, **2011**, *413*, 297.
11. E. Serebryany and J. A. King, *Prog. Biophys. Mol. Biol.*, 2014, **115**, 32.
12. P. J. Thomas, B. H. Qu and P. L. Pedersen, *Trends Biochem. Sci.*, 1995, 20, 456.
13. A. Aguzzi and T. O'Connor, *Nat. Rev. Drug Discov.*, 2010, 9, 237.
14. N. Fujii, H. Uchida and T. Saito, *Mol. Vis.*, **2004**, *10*, 814.
15. P. M. Masters, J. L. Bada and J. Samuel Zigler, *Nature*, 1977, **268**, 71–73.
16. G. M. Martin, S. N. Austad and T. E. Johnson, *Nat. Genet.*, **1996**, *13*, 25.
17. Stadtman and R. Earl, *Ann. N. Y. Acad. Sci.*, 2001, **928**, 22.
18. S. I. S. Rattan, *Adv. Exp. Med. Biol.*, 2010, **694**, 1.
19. S. Kaushik and A. M. Cuervo, *Nat. Med.* **2015**, *21* (12), 1406.
20. J. C. Price, S. Guan, A. Burlingame, S. B. Prusiner and S. Ghaemmaghami, *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (32), 14508.
21. B. H. Toyama, J. N. Savas, S. K. Park, M. S. Harris, N. T. Ingolia, J. R. Yates, and M. W. Hetzer, *Cell* **2013**, *154* (5), 971.
22. B. H. Toyama and M. W. Hetzer, *Nat. Rev. Mol. Cell Biol.*, **2013**, *14*, 55.
23. A. E. Roher, J. D. Lowenson, S. Clarke, C. Wolkow, R. Wang, R. J. Cotter, I. M. Reardon, H. A. Zurcher-Neely, R. L. Heinrikson, M. J. Ball and B. D. Greenberg, *J. Biol. Chem.*, **1993**, *268*, 3072.
24. J. B. Catterall, D. Barr, M. Bolognesi, R. D. Zura and V. B. Kraus, *Arthritis Res. Ther.*, **2009**, *11*, R55.
25 M. A. Wride, *Philos. Trans. R. Soc. B Biol. Sci.*, 2011, **366**, 1219.
26. S. R. A. Hanson, A. Hasan, D. L. Smith and J. B. Smith, *Exp. Eye Res.*, **2000**, *71*, 195.
27. A. L. Lund, J. B. Smith and D. L. Smith, *Exp. Eye Res.*, **1996**, *63*, 661.
28. K. J. Lampi, Z. Ma, S. R. A. Hanson, M. Azuma, M. Shih, T. R. Shearer, D. L. Smith, J. B. Smith and L. L. David, *Exp. Eye Res.*, **1998**, *67*, 31.
29. R. J. W. Truscott and M. G. Friedrich, *Biochim. Biophys. Acta*, 2016, 1860, 192..
30. P. M. Masters, J. L. Bada and J. Samuel Zigler, *Nature*, 1977, **268**, 71.
31. D. Pascolini and S. P. Mariotti, *Br. J. Ophthalmol.*, 2012, **96**, 614.
32. G. Tabin, M. Chen and L. Espandar, *Curr. Opin. Ophthalmol.*, 2008, 19, 55.
33. K. L. Moreau and J. A. King, *Trends Mol. Med.*, 2012, 18, 273–282.

34. D. S. Rehder, D. Chelius, A. McAuley, T. M. Dillon, G. Xiao, J. Crouse-Zeineddini, L. Vardanyan, N. Perico, V. Mukku, D. N. Brems, M. Matsumura and P. V. Bondarenko, *Biochemistry*, **2008,** *47*, 2518.
35. M. J. Mamula, R. J. Gee, J. I. Elliott, A. Sette, S. Southwood, P. J. Jones and P. R. Blier, *J. Biol. Chem.*, **1999**, *274*, 22321.
36. S. Noguchi, *Biopolymers*, **2010**, *93*, 1003.
37. S. Noguchi, K. Miyawaki and Y. Satow, *J. Mol. Biol.*, **1998**, *278*, 231.
38. J. V. Olsen, L. M. F. de Godoy, G. Li, B. Macek, P. Mortensen, R. Pesch, A. Makarov, O. Lange, S. Horning and M. Mann, *Mol. Cell. Proteomics*, 2005, **4**, 2010.
39. P. Roepstorff and J. Fohlman, *Biol. Mass Spectrom.*, 1984, **11**, 601.
40. J. Mitchell Wells and S. A. McLuckey, *Methods Enzymol.*, 2005, **402**, 148.
41. A. R. Dongré, J. L. Jones, Á. Somogyi and V. H. Wysocki, *J. Am. Chem. Soc.*, 1996, **118**, 8365.
42. R. A. Zubarev, *Mass Spectrom. Hyph. Tech. Neuropept. Res.*, 2002, 277.
43. E. A. Syrstad and F. Tureček, *J. Am. Soc. Mass Spectrom.*, 2005, 16, 208.
44. A. Chi, C. Huttenhower, L. Y. Geer, J. J. Coon, J. E. P. Syka, D. L. Bai, J. Shabanowitz, D. J. Burke, O. G. Troyanskaya and D. F. Hunt, *Proc. Natl. Acad. Sci.*, 2007, **104**, 2193.
45. Chi, A. *et al.* Analysis of phosphorylation sites on proteins from Saccharomyces cerevisiae by electron transfer dissociation (ETD) mass spectrometry. *Proc. Natl. Acad. Sci.* **104**, 2193–2198 (2007).
46. Riley, N. M., Westphall, M. S. & Coon, J. J. Activated Ion-Electron Transfer Dissociation Enables Comprehensive Top-Down Protein Fragmentation. *J. Proteome Res.* **16**, 2653–2659 (2017).
47. Yu, Q., Shi, X., Feng, Y., Kent, K. C. & Li, L. Improving data quality and preserving HCD-generated reporter ions with EThcD for isobaric tag-based quantitative proteomics and proteome-wide PTM studies. *Anal. Chim. Acta* **968**, 40–49 (2017).
48 P. B. O'Connor, J. J. Cournoyer, S. J. Pitteri, P. A. Chrisman and S. A. McLuckey, *J. Am. Soc. Mass Spectrom.*, 2006, **17**, 15.
49. T. Ly and R. R. Julian, *J. Am. Chem. Soc.*, 2008, **130**, 351–358.
50. Q. Sun, H. Nelson, T. Ly, B. M. Stoltz and R. R. Julian, *J. Proteome Res.*, 2009, **8**, 958.
51. Moore, B. N. & Julian, R. R. Dissociation energies of X-H bonds in amino acids. *Phys. Chem. Chem. Phys.* **14**, 3148–3154 (2012).
52. D. L. Riggs, S. V. Gomez and R. R. Julian, *ACS Chem. Biol.*, **2017**, *12*, 2875.
53. Y. Tao and R. R. Julian, *Anal. Chem.*, 2014, **86**, 9733.
54. Y. Tao, N. R. Quebbemann and R. R. Julian, *Anal. Chem.*, 2012, **84**, 6814.
55 T. Ly, X. Zhang, Q. Sun, B. Moore, Y. Tao and R. R. Julian, *Chem. Commun.*, 2011, **47**, 2835.
56. Maestro, ver. 9.7. Schrödinger, LLC: New York, NY (2012)
57 N. DeGraan-Weber, J. Zhang and J. P. Reilly, *J. Am. Soc. Mass Spectrom.*, 2016, **27**, 2041.

Chapter 2

Identification of Sequence Similarities among Isomerization Hotspots in Crystallin Proteins

2.1 Introduction

In most cells proteostasis involves continuous regulation of protein synthesis and degradation, but the eye lens is a notable exception.[1] During the final stages of maturation, lens cells eject all organelles, preventing normal protein turnover.[2] The lens therefore provides a unique opportunity to observe the effects of aging in relation to damage accumulation in long-lived proteins. The α-crystallins are important multimeric chaperones that help prevent protein aggregation. They constitute approximately 35-40% of the total soluble mass of the lens and are comprised of a 3:1 ratio of αA to αB.[3] While αA is found almost exclusively in the lens, αB has been found throughout the body including heart, glia, muscle, kidney, lung and skin cells.[4] Gene knockout studies revealed that a mouse lens can develop normally without αB but not without αA, which led to prompt cataract formation. [5,6] Nevertheless, the chaperone activity of heteropolymers formed from a 3:1 ratio of αA to αB is higher than either homopolymer, suggesting the combination of both proteins is optimal.[7] Beta and gamma crystallins are the two other major subgroups of crystallins in the lens. They act in conjunction with the α-crystallins to maintain lens structure, transparency, and to achieve a suitable refractive index.[8,9]

In order to function as chaperones, α-crystallin monomers must interact with each other to form polydisperse oligomers. αA and αB form dimers which then assemble into larger complexes ranging from 15 to 50 monomers, with the subunits dynamically intermixing and exchanging.[10] Consequently, obtaining crystal structures of the α-crystallins has been daunting. Removing the N- and C-terminal extensions allows for determination of partial crystal structures.[11,12] When crystallins do not interact with each other properly, they become water-insoluble and begin to aggregate. The influence of many post-translational modifications (PTMs) on this process has been investigated, including deamidation, oxidation, disulfide formation and truncation.[13,14] Importantly, all of these PTMs lead to mass shifts that are easily detectable by mass spectrometry (MS).

More subtle modifications that also increase with age, such as isomerization and epimerization, do not lead to easily detectable mass shifts and have been significantly less studied.[15,16] Methods for detecting isomerization in peptides (meaning either inversion of the chiral center of an amino acid to produce an epimer, or conversion of aspartic acid to isoaspartic acid) include stereoselective enzymatic digestion[17], ion mobility,[18] and mass spectrometry.[19] Previous studies have shown that aspartic acid and serine are the two amino acids most susceptible to spontaneous isomerization.[20] Deamidation of asparagine residues generates aspartic acid, and this transformation can also be accompanied by isomerization. The mechanism by which L-Asp converts into L-isoAsp, D-isoAsp and D-Asp has been studied previously is known to proceed via non-enzymatic formation of a succinimide ring which can yield four isomers shown in Figure 2.1. L-succinimide can be converted to D-Succinimide via keto-enol tautomerism, and then hydrolyzed to form D-

Asp or D-isoAsp.[21] The mechanism by which L-Ser epimerizes to D-Ser in proteins is not well-established, although proposals have been made.[22,23]

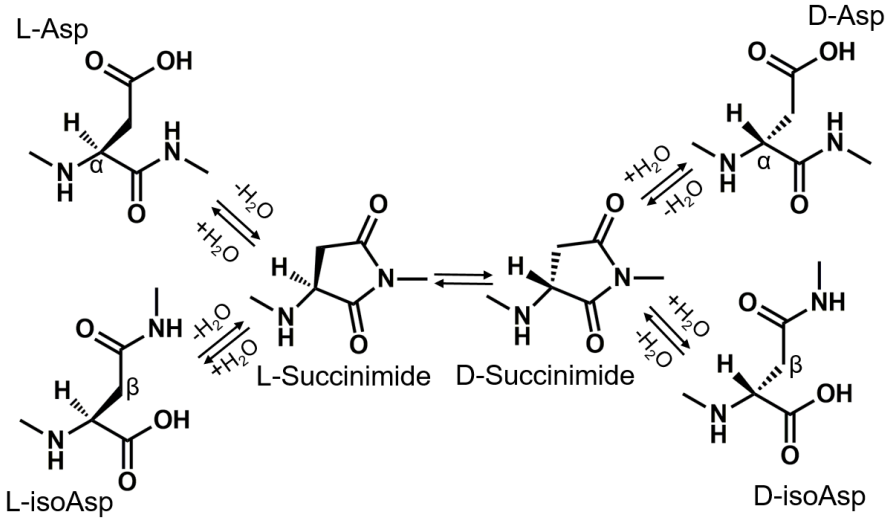


Figure 2.1 Mechanism of spontaneous aspartic acid isomerization via a succinimide intermediate. Alpha and beta carbons are labeled to highlight difference between L-asp and L-isoasp, respectively.

Though isomerization might appear to be a "benign" PTM, studies have shown it can cause major perturbations in protein structure. Noguchi and co-workers successfully crystallized Hen egg-white lysozyme with an isoaspartic acid substitution at Asp101 that caused backbone deflection of nearly 90° relative to the native structure.[24] Crystal structures of a modified ribonuclease revealed that isoAsp-32 induces conversion of an α-helix to a U-shaped loop.[25] Isomerization not only perturbs three-dimensional structure, but can also affect physical properties such as solubility and bioactivity. For example, isomerization of Asp92 in immunoglobulin γ2 (IgG2) leads to deactivation of the antigen-binding region.[26]

Large scale (i.e. proteomic) identification of single amino acid isomerism within peptides is challenging because there is no change in mass. However, differences in fragment intensities following MS/MS analysis can be used to detect isomers. This method was first applied to stereoisomers by Tao and co-workers, who reported differences as a ratio of relative abundances of a pair of fragment ions differing most between L- and D- enantiomers ($R_{chiral} = R_D/R_L$).[27] It has been shown that radical-directed dissociation[28] (RDD) yields the best chiral discrimination for analysis of peptide epimers.[29] Previous work on the detection of isomerization and epimerization in the sheep lens using tandem LC-MS revealed novel sites of isomerization in the crystallins.[30]

The present study focuses on changes in isomerization and epimerization for αA, αB and βB3 crystallin between water-soluble (WS) and water-insoluble (WI) protein fractions from cow, sheep, and pig eye lenses. To determine which regions of these proteins are most susceptible to isomerization, enzymatic digestion into peptides was followed by LC-MS/MS analysis using both collision-induced dissociation (CID) and RDD. Importantly, we found several regions in the well-ordered crystallin domain that are disproportionately isomerized in the WI fractions. Specific isomerization "hotspots" were identified, and correspond to regions with serine or aspartic acid repeats. Potential explanations for the isomerization of these sites are offered, and comparison with other long-lived and problematic proteins reveals these sequence motifs are common.

2.2 Experimental Methods

2.2.1 Protein Extraction and Digestion

Cow, pig and sheep eye lenses were acquired as discarded tissue from Corona Cattle Inc. (Corona, CA). The approximate ages for each of the animals were 12-18 months for the cow, 5-6 months for the pig and 6-8 months for the sheep. The lenses were separated and washed with distilled H2O. Whole lenses were then homogenized in 50mM Tris-HCl pH=7.8. The supernatant was separated from the precipitate following centrifugation at 15,100 g for 20 minutes at 4°C. The supernatant (water-soluble) was purified by dialysis against water. The precipitate (water-insoluble) was solubilized in 6 M urea and purified by dialysis against 6 M urea. For water-soluble digestion, the protein was dissolved in 50 mM NH4HCO3 buffer, pH =7.8, disulfide bonds were then reduced with 100 mM DTT at 95°C for 5 minutes. After returning to room temperature, reduced cysteine residues were capped using 100 mM iodoacetamide in the dark for 20 minutes. Finally, the proteins were digested with trypsin overnight at 37°C using a 50:1 protein to enzyme ratio. For the water-insoluble digestion, the proteins were dissolved using 6 M in 50 mM Tris-HCl, pH= 8.0. Disulfide bonds were cleaved using 200 mM DTT in Tris-HCl, pH=8.0 at 37°C for 20 minutes. Following this, 200 mM iodoacetamide in Tris-HCl, pH= 8.0 was added and the mixture was incubated in the dark for one hour. To consume unreacted iodoacetamide, 200 mM DTT was added and incubated for one hour in the dark. Next, the urea concentration was diluted to < 0.6 M using 50 mM Tris-HCl, 1 mM CaCl2, pH= 7.6. The proteins were digested using trypsin with a 50:1 protein to enzyme ratio for 16 hours at 37°C. For the iodobenzoic acid modification, the digested

peptides were desalted and cleaned using a peptide trap (Michrom Bioresource Inc). Approximately 5 nmoles of the digestion mixture, 15 μL of 15 mM 4-iodobenzoic acid NHS-activated ester in dioxane and 5 μL of borate buffer (pH =8.6) were combined and incubated for 1 hour at 37°C. Important: Dimethyl sulfoxide should not be substituted for dioxane in this step because it can cause aspartic acid isomerization. The modification side products at arginine and tyrosine side chains were removed by incubating the reaction mixture in 1 M hydroxylamine, pH= 8.5. The same procedure was used for the synthetic peptide standards. These procedures have been determined previously not to yield any detectable isomerization/epimerization in control experiments.30

2.2.2 Calculation of R Values

To quantify isomer identification, an R_{chiral} value approach, originally reported by Tao et. al., was utilized.27 In this paper, R_{isomer} represents the ratios of the relative intensities of a pair of fragments that varies the most between two isomers (RA/RB). If $R_{isomer} = 1$ then there the two tandem MS spectra are indistinguishable and the species are likely not isomers. If $R_{isomer} >1$, a larger number indicates a higher probability that two unique molecules are represented. To confidently identify each of these isomers by MS/MS we use a threshold that was determined by performing a t test on the R_{isomer} values obtained by performing CID and RDD on a mixture of synthetic peptides separated by LCMS. Using 99% confidence intervals, the R_{isomer} threshold for CID is >1.9 and for RDD it is >2.4.30

2.2.3 Materials

Organic solvents were purchased from Sigma Aldrich (St. Louis, MO) or Fisher Scientific and used without further purification. Water was purified to 18.2 MΩ by a Millipore 147 (Billerica, MA) Direct Q system. Amino acids and resins were purchased from Anaspec (Fremont, CA). Trypsin and urea were purchased from Sigma Aldrich (St. Louis, MO).

2.2.4 Peptide and Radical Precursor Synthesis

All synthetic peptides were synthesized manually using standard FMOC procedures with Wang Resins as the solid support.[31] N-hydroxysuccinimide (NHS) activated iodobenzoyl esters were synthesized by following a previous procedure.[32]

2.2.5 Mass Spectrometry and Radical Directed Dissociation

Solutions were analyzed using an LTQ linear ion trap mass spectrometer (Thermo Fisher Scientific, San Jose, CA) with a standard electrospray ionization source. The posterior plate of the ion trap was modified with a quartz window to allow fourth-harmonic (266 nm) laser pulses from a flash-pumped Nd:YAG laser (Continuum, Santa Clara, CA). A digital delay generator (Berkeley Nucleonics, San Rafael, CA) was used to synchronize laser pulses. Photodissociation of the *para*-iodobenzoate labeled peptide causes homolytic dissociation of the carbon-iodine bond producing a radical peptide. Further collision-induced dissociation (MS^3) was performed on the radical by re-isolation of the largest peak in the photodissociation spectrum.

2.2.6 LC-MS Data Acquisition and Analysis

An Agilent 1100 series HPLC system (Agilent, Santa Clara, CA) with a BetaBasic C-18 column (150 mm x 2.1 mm, particle size 3 μm) was coupled to the LTQ mass spectrometer. Peptides were separated using a 0.1% formic acid in water (mobile phase A) and a 0.1% formic acid in acetonitrile (mobile phase B) binary system at a flow rate of 0.2 mL/min. The digestion mixtures were loaded onto the column and separated using the following gradient: 5% B to 20% B over 25 minutes, 20% B to 30% B over the next 35 minutes, 30% B to 50% B over the next 15 minutes and 50% B to 95% B over the final 10 minutes. Another gradient of 3%B to 20%B over 80 minutes was also used to help separate a few co-eluting peptides with very short retention times. The LTQ was operated in data-dependent mode using the Xcalibur program (Thermo Fisher Scientific). Specifically, in the CID-only LC-MS run, the first scan event was a full MS from m/z 400-2000 Da, followed by an ultrazoom (MS^2) and then CID (MS^3). In the RDD LC-MS experiments, the laser pulses were triggered during the MS^2 step and CID was performed as a pseudo-MS^3 step. Due to the high photodissociation yield of the 4-iodobenzoic acid chromophore, the major peak during this step was the loss of iodine, and it is the subsequent precursor for MS^3. The exclusion time was set to 60 seconds for the identification of peptides. For isomer identification, an inclusion mass list was added and the exclusion time was reduced to 16 seconds.

MS data was acquired with Xcalibur software. The .raw files were converted to .mgf files using MSConvert software. The .mgf files were searched with X!Tandem PILEDRIVER edition (version 2015.04.01.1) against the *Ovis aries, Sus scrofa* and *Bos*

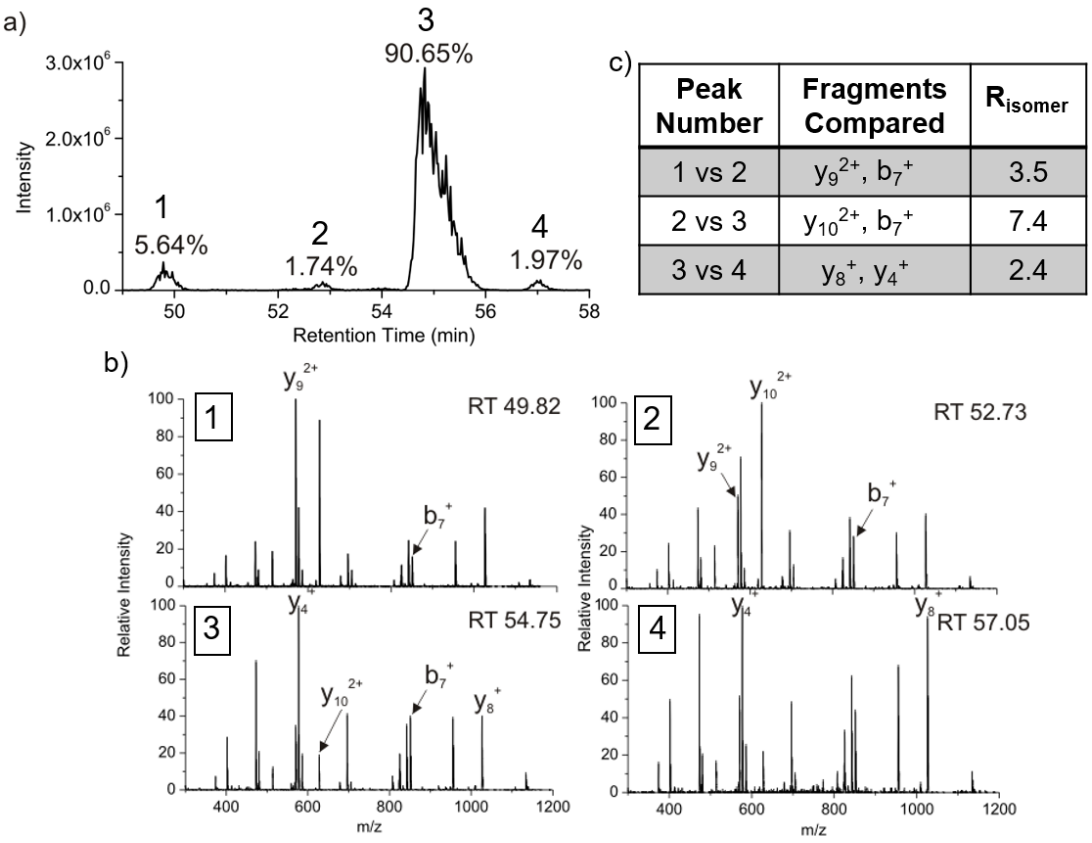
taurus UniprotKB databases. The cleavage sites were set as lysine and arginine (semi cleavage was turned on), allowing up to two missed cleavages and one point mutation). Carbomylation (+57.02 Da at Cys) was set as a fixed modification, and N-acetylation, phosphorylation and oxidation were all considered as variable modifications. For the iodobenzoate-modified digestion mixture, the modifier (+230.01 Da) was considered as a possible modification at either the N-terminus or lysine side chain. The parent monoisotopic mass error was set to ±1 Da, the minimum parent mass was set to 400 Da and the fragment mass error was set to ±0.4 Da. The criteria for accepting peptides identified by X! Tandem was $e < 0.005$. The false discovery rate was set to 0.1% and was calculated by searching the data against the reversed database.

Clustal Omega was used to determine sequence homology between the crystallins of different animals.[33] The models of crystal structures in Figures 2, 3, 4 and 7 were made using UCSF Chimera.[34] *In silico* mutation of Asp109 was performed using Maestro MacroModel (Schrodinger Inc., Portland, OR, USA).[35]

2.3 Results and Discussion

General Observations. We have previously described a modified bottom-up strategy for detecting sites of isomerization, including epimerization, in proteins.[30] Briefly, proteolytically digested peptides are separated by reverse-phase column chromatography and analyzed by MS with minimal time exclusion windows to favor repeated analysis the same m/z. For example, the N-terminal fragment of αA crystallin from the water-insoluble (WI) protein fraction of sheep lens, Acetyl-MDIAIQHPWFK, elutes at four different times, each separated by about two minutes as shown in Figure 2.2a. In order to

establish that these peaks represent isomers, repeated MS/MS analysis of each eluting peak is required (both CID and RDD are used in separate runs).[30] Figure 2.2b shows CID spectra from each of the LC peaks in Figure 2.2a. Although the spectra are similar, suggesting they originate from the same peptide sequence, the relative intensities of certain peaks change noticeably. For example, the intensities of y_9^{2+}, b_7^{+} and y_{10}^{2+} ions vary considerably between LC peaks 1 and 2. These differences can be quantified into R_{isomer} values,[30] which are provided in Figure 2.2c. For CID analysis, R_{isomer} values above 1.9 represent statistically significant differences, meaning that all four peaks in Figure 2.2a are different isomers of Acetyl-MDIAIQHPWFK. Mass spectra alone do not provide information about which isomer corresponds to the native L-Asp peptide, but in this case, the largest peak constitutes over 90% of the total relative abundance and likely represents the unmodified L-Asp isomer. If the WI fraction is highly isomerized, in which case identification of the L-Asp isomer becomes ambiguous, comparison with the less isomerized WS digest or an authentic standard is used to identify the L-Asp isomer.



Peak Number	Fragments Compared	R_{isomer}
1 vs 2	y_9^{2+}, b_7^+	3.5
2 vs 3	y_{10}^{2+}, b_7^+	7.4
3 vs 4	y_8^+, y_4^+	2.4

Figure 2.2. a) LC chromatogram for the separation of the four isomers of Ac-MDIAIQHPWFK in the WI sheep lens digest. The percentages represent the calculated peak area. b) CID spectra from each of the LC peaks. Labeled fragments are those used to determine R_{isomer} values. c) Table of the R_{isomer} values including which fragments were compared.

The degree of peptide isomerization from the WS and WI protein fractions of sheep αA crystallin is summarized in Figure 2.3. All isomers were confirmed by comparison of either CID or RDD MS/MS spectra as described above. For example, the two isomers of ^{158}AIPVSR^{163} were indistinguishable by CID, but could be confidently detected using RDD (see Figure 2.4). Additionally, some isomers co-elute, preventing quantitation of the relative abundance of each form. Isomerization of these peptides is denoted with stars. The first downward bar on the left side of Figure 2.3a (1-11) represents the data

from Figure 2.2. This peptide, Acetyl-MDIAIQHPWK, was found to be 9% isomerized, with error bars representing standard deviations from three technical replicates. The color coding of the bars corresponds to the three distinct structural regions present in α-crystallins, the N-terminal disordered region, the crystallin ordered region, and the disordered C-terminal extension.[36] These regions are illustrated relative to a partial crystal structure on the right side of Figure 2.3.[12] The full protein sequence is provided at the bottom of Figure 2.3, color coded and with residues of interest marked for easy location. The trends in Figure 2.3 illustrate that the average amount of isomerization is significantly higher in the N- and C- termini than in the well-ordered crystallin domain for both the WS and WI fractions. These dynamic regions play important roles in the assembly of α-crystallins into higher-order structures,[37] but this flexibility may also enable isomerization by facilitating more frequent access to the succinimide intermediate.[38]

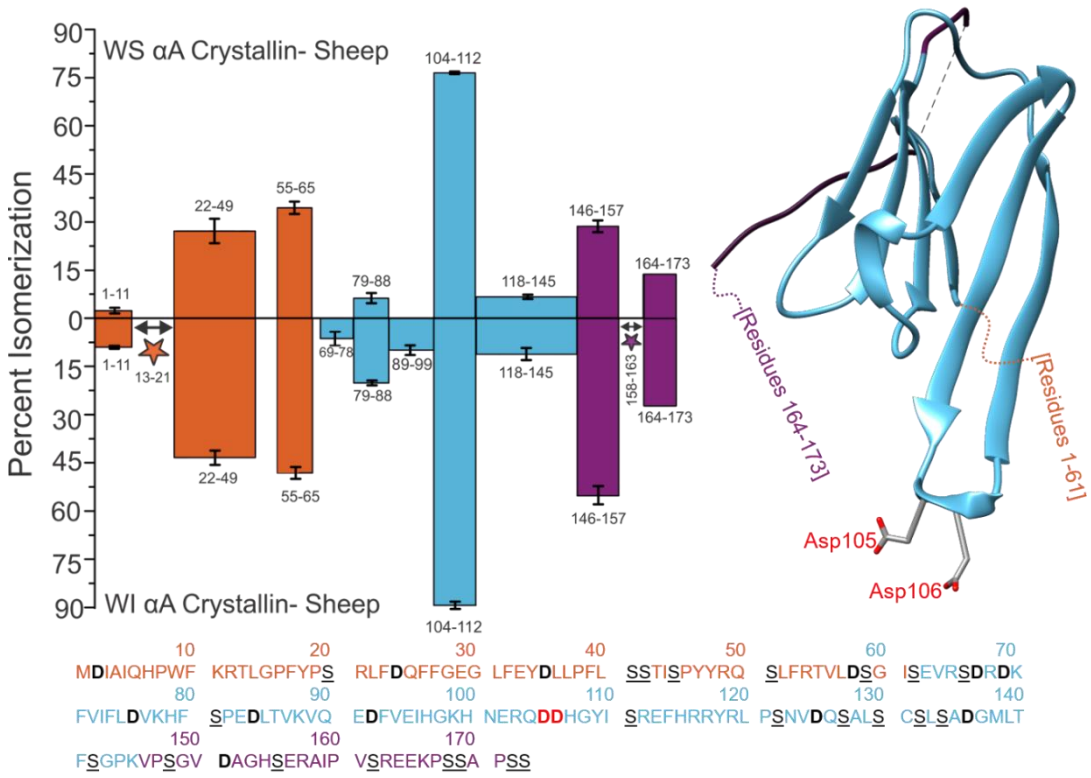


Figure 2.3 Percent isomerization of water-soluble (WS) αA Sheep versus water-insoluble (WI) αA Sheep. Orange= disordered N-terminus, blue= structured α-crystallin domain and purple= disordered C-terminus. Three separate digests were performed; error bars represent standard deviations. Number ranges represent peptide sequences. Peptide 164-173 does not contain error bars because it only appeared baseline resolved in one digest. The full protein sequence is given below the plot, with aspartic acid residues in bolded/black and serine residues in underlined/black. Asp105 and Asp106 are in bold, red text in the amino acid sequence and are shown explicitly in the crystal structure (PDB 3L1F) to highlight an important region of isomerization. Stars indicate isomerized regions where isomerization was identified, but quantitation was not possible due to incomplete chromatographic separation.

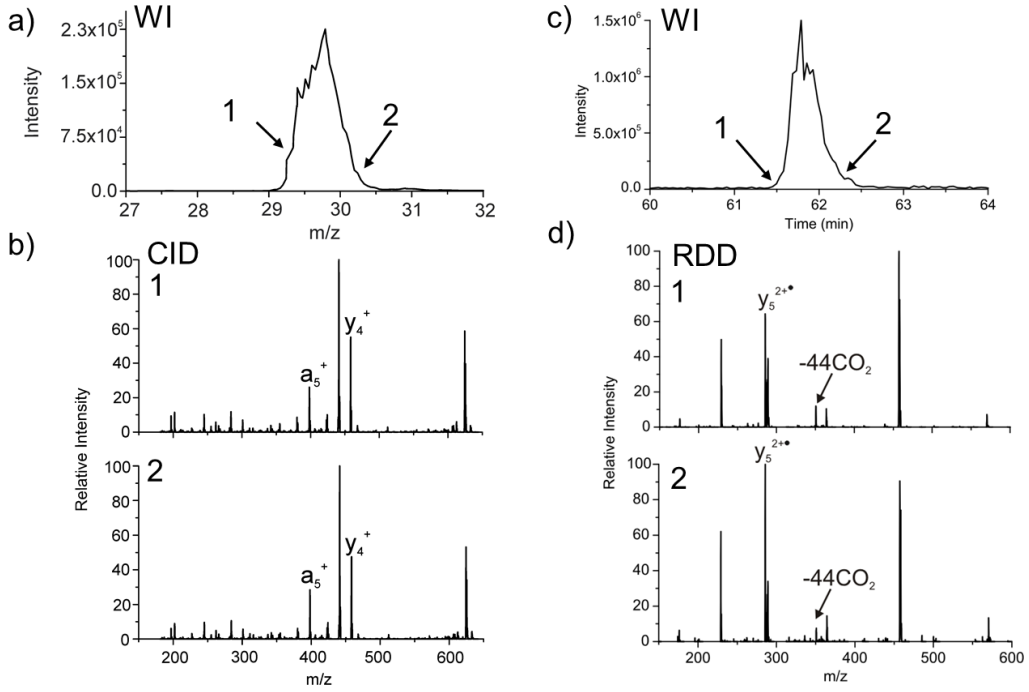


Figure 2.4 Chromatogram of AIPVSR peptide from WI sheep. **b)** CID is performed on the front of the peak and tail of it to see if the isomers were co-eluting. The R_{isomer} value between the two resulting spectra (1 and 2) is 1.4 which is below our threshold for confident detection. **c)** The WI digest is then modified with the iodobenzoyl chromophore, and again only one 4IB-AIPVSR peak appears in the chromatogram. **d)** Performing RDD on the edges of this peak (1 and 2) yields an R_{isomer} value of 3.3 which is well above our threshold of 2.4, confirming that they are isomers

Interestingly, the degree of isomerization in the disordered regions, although slightly more abundant in the WI fraction, is similar in both the WS and WI fractions. This suggests these modifications do not significantly drive aggregation and loss of solubility, as reported previously.[39] Differences in degree of isomerization are more notable within the crystallin region, where isomerization is signficantly more abundant in the WI fraction (excepting the dramatically modified peptide 104-112 that will be discussed further below). Cumulatively, these results imply that modifications to flexible regions may be more facile, but also less consequential in terms of altered functionality.

Results from an identical analysis of αB are shown in Figure 2.5. Fewer peptides are isomerized in αB and the average degree of isomerization is less in both the WS and WI fractions compared to αA. The flexible N-terminal domain is highly isomerized, similar to αA. Interestingly, the C-terminal extension is not isomerized, which contrasts with αA and is largely due to the fact that the C-terminal extension of αB lacks any aspartic acid residues and contains only a single serine. The degree of isomerization observed in the WS versus WI fractions also varies more dramatically compared to αA, with greater isomerization being observed in the WI fraction for all peptides. αA and αB share similar functionality and freely intermix to form higher order structures, yet comparison of Figures 2.3 and 2.5 illustrates significant differences in propensity and effects of isomerization. The sequence alignment of the two proteins is less than 60% (Figure 2.6), but the tertiary structures of αA and αB are very similar as illustrated by the crystal structure of truncated bovine αB on the right side of Figure 2.5.[12] Despite the similarity, sequence variation appears to have a significant effect on the ability of αB to accommodate isomerization and retain solubility, in agreement with previous observations.[40] Another interesting difference between the two α-crystallins is the overall abundance of acidic amino acids. Aspartic acid in αA comprises 8.8% of residues and glutamic acid contributes another 5.8%, whereas αB contains 6.3% aspartic acid and 8.0% glutamic acid. Although isomerization of glutamic acid is possible, formation of the glutarimide intermediate is much slower compared to the succinimide equivalent in aspartic acid.[41] Therefore, even though both proteins contain a similar percentage of

acidic residues, αA contains more residues prone to isomerization, and these residues often reside in regions where isomerization is facile.

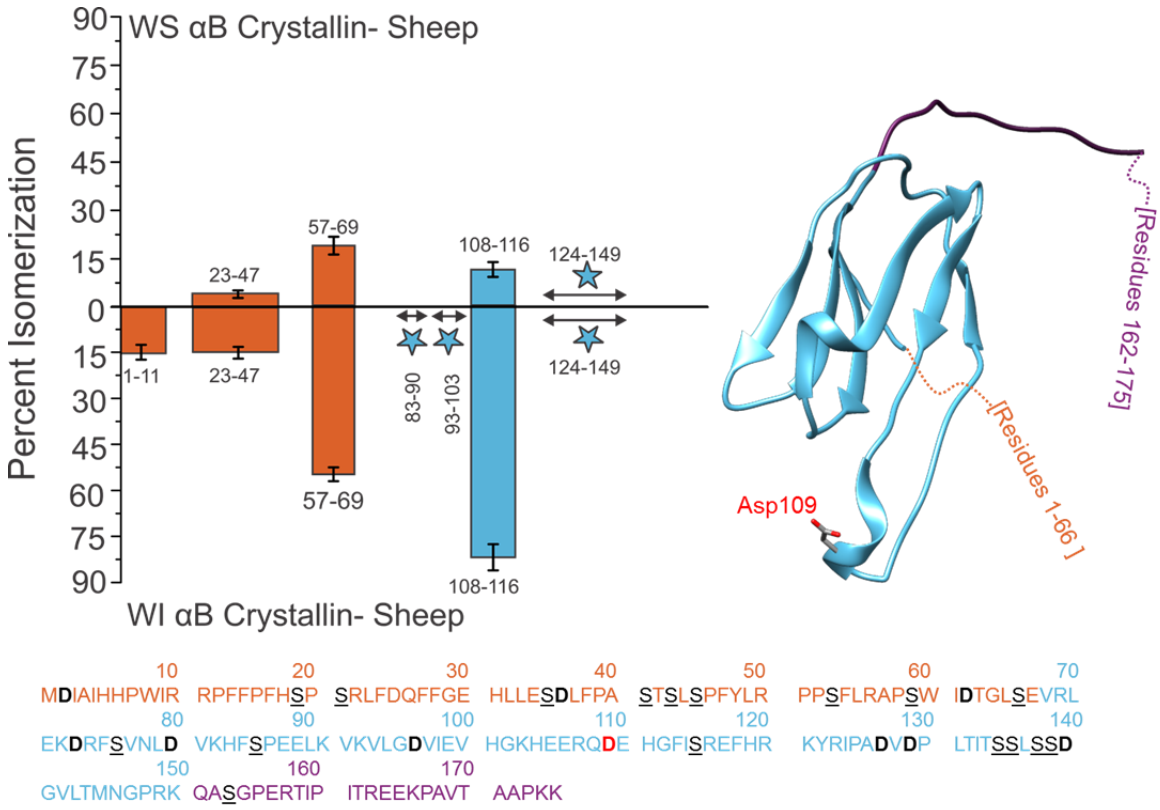


Figure 2.5 Isomerization of αB from sheep, which differs significantly from what was observed for αA (see Figure 2a, formatting is identical). Asp109 in the amino acid sequence is in red, bold text and shown explicitly in the crystal structure (PDB 3L1G) to highlight an important site of isomerization.

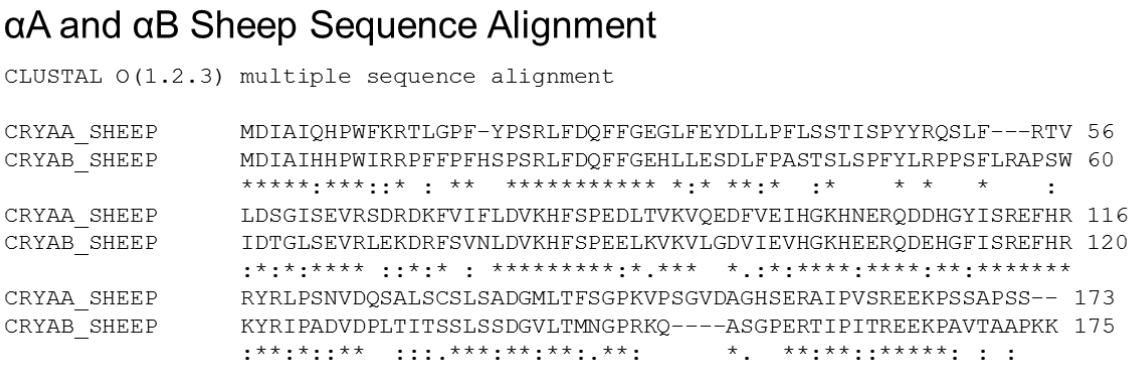


Figure 2.6 Sequence alignment between αA and αB crystallin from sheep using the Clustal Omega matrix. The two sequences share a 56% homology.

The same analysis that was performed on the alpha crystallins was also conducted on one of the beta crystallins. Results for βB3 crystallin are shown in Figure 2.7. βB3 is comprised of 211 amino acids with two large structured domains connected by a flexible linker. The disordered N-terminal and C-terminal regions are much smaller than observed in the α-crystallins. The total amount of isomerization detected in the WS and WI fractions of βB3 is lower than in αA and αB, which is likely due to the greater fraction of highly structured regions. Similar to αB, βB3 is characterized by significant disparity between the degree of isomerization in the WS versus WI fractions, with greater isomerization being observed in the latter. This suggests again that structural perturbations in βB3 are consequential and lead to rapid loss of solubility. The isomerized versions of several peptides, including ^{103}SLRPLHIDGPDHK^{115} and ^{128}KMEIVDDDVPSLW^{140} are only detectable in the WI fractions (see Figure 2.8 and Table 2.1). ^{103}SLRPLHIDGPDHK^{115} is part of the structurally critical connecting peptide that joins the two domains of the monomer, denoted by the green bar in the Figure 2.7. ^{128}KMEIVDDDVPSLW^{140} contains aspartic acid residues that form ion pairs during

assembly into higher-order oligomers.[42] Sequence alignment for human, pig, cow and sheep βB3 show that each contain this motif along (Figure 2.9). The results in Figure 2.7 suggest isomerization of either of these peptides leads to dramatic loss of solubility and, by extension, function.

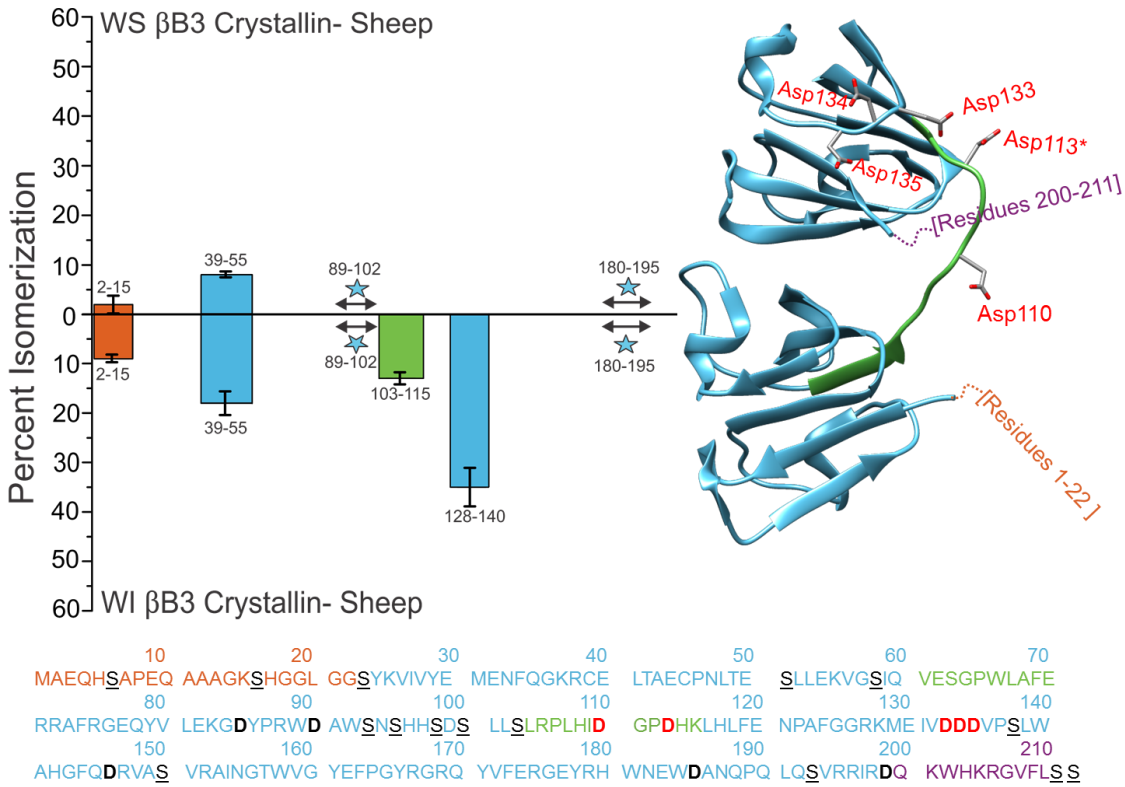


Figure 2.7 Percent isomerization of WS versus WI βB3 Sheep. Formatting is identical to Figure 2, except for the linker of the structured domains, shown in green. Asp110, Asp113, Asp133, Asp134 and Asp135 are all in red, bold text in the amino acid sequence and are shown explicitly in the human crystal structure (PDB 3QK3) to highlight important regions of isomerization. His113 from human βB3 was mutated to Asp113*.

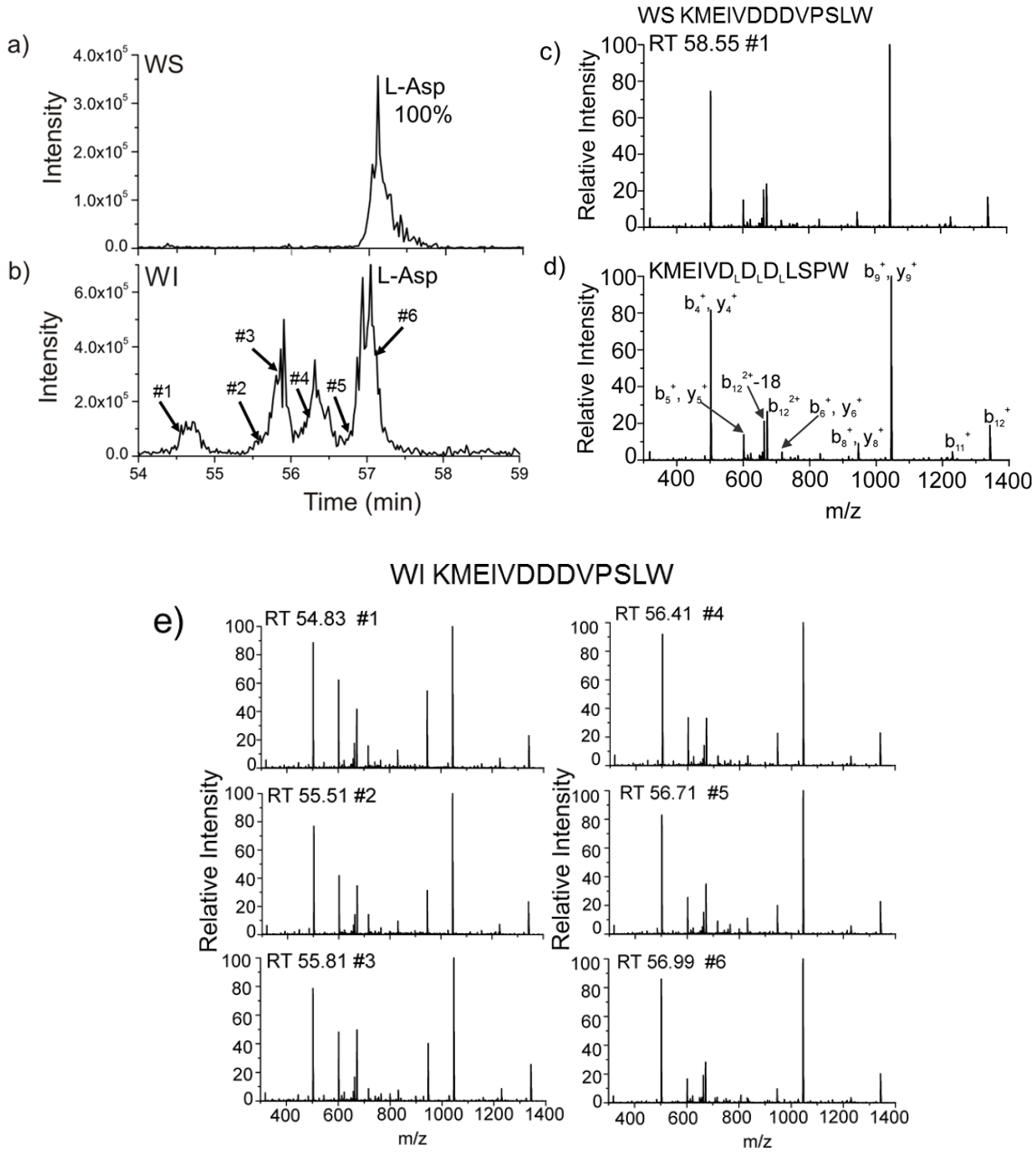


Figure 2.8 a) Chromatogram of the KMEIVDDDVPSLW peptide from the WS pig digest and **b)** WI pig digest. Arrows represent when CID was performed on the peak. **c)** The resulting CID spectrum from the peak in the WS chromatogram. **d)** CID spectrum from the synthetic all L-Asp version of the peptide. The R_{isomer} value between the c) and d) is 1.2 which is well below our threshold, confirming that this peak matches the all L-Asp version. **e)** The resulting CID spectra from each of the time points indicated by the arrows in b.

Table 2.1 R_{isomer} values for the L-Asp synthetic version of KMEIVDDDVPSLW versus each peak from the WI pig digest. Only one of the CID spectra yielded an R_{isomer} score that was below the threshold (shown in bold), and could confidently assigned as the all L-Asp version of the peptide.

Spectrum 1	Spectrum 2	R_{isomer}
L-Asp Synthetic	Peak 1	6.1
L-Asp Synthetic	Peak 2	3.7
L-Asp Synthetic	Peak 3	4.6
L-Asp Synthetic	Peak 4	2.5
L-Asp Synthetic	Peak 5	2.3
L-Asp Synthetic	**Peak 6**	**1.2**

βB3 Sequence Alignment

```
CLUSTAL O(1.2.3) multiple sequence alignment

CRYBB3_HUMAN       MAEQHGAPEQAAAGKSHGDLGGSYKVILYELENFQGKRCELSAECPSLTDSLLEKVGSIQ 60
CRYBB3_PIG         MAEQHGTPEQAAAGKSHGGLGGGYKVIVYEMENFQGKRCELSAECPNLTESLLEKVGSIQ 60
CRYBB3_BOVINE      MAEQHSTPEQAAAGKSHGGLGGSYKVIVYEMENFQGKRCELTAECPNLTESLLEKVGSIQ 60
CRYBB3_SHEEP       MAEQHSAPEQAAAGKSHGGLGGSYKVIVYEMENFQGKRCELTAECPNLTESLLEKVGSIQ 60
                   *****.:***********.***.****:**:**********:****.**:**********
CRYBB3_HUMAN       VESGPWLAFESRAFRGEQFVLEKGDYPRWDAWSNSRDSDSLLSLRPLNIDSPHHKLHLFE 120
CRYBB3_PIG         VESGPWLAFERRAFRGEQFVLEKGDYPRWDAWSNSHHSDSLLSLRPLHIDGPDHKLHLFE 120
CRYBB3_BOVINE      VESGPWLAFERRAFRGEQYVLEKGDYPRWDAWSNSHHSDSLLSLRPLHIDGPDHKLHLFE 120
CRYBB3_SHEEP       VESGPWLAFERRAFRGEQYVLEKGDYPRWDAWSNSHHSDSLLSLRPLHIDGPDHKLHLFE 120
                   ********** *******:****************:.**********:**.*.*******
CRYBB3_HUMAN       NPAFSGRKMEIVDDDVPSLWAHGFQDRVASVRAINGTWVGYEFPGYRGRQYVFERGEYRH 180
CRYBB3_PIG         NPAFGGRKMEIVDDDVPSLWAHGFQDRVASIRAINGTWVGYEFPGYRGRQYVFERGEYRH 180
CRYBB3_BOVINE      NPAFGGRKMEIVDDDVPSLWAHGFQDRVASVRAINGTWVGYEFPGYRGRQYVFERGEYRH 180
CRYBB3_SHEEP       NPAFGGRKMEIVDDDVPSLWAHGFQDRVASVRAINGTWVGYEFPGYRGRQYVFERGEYRH 180
                   ****.************************:****************************
CRYBB3_HUMAN       WNEWDASQPQLQSVRRIRDQKWHKRGRFPSS 211
CRYBB3_PIG         WNEWDANQPQLQSVRRIRDQKWHKRGCFLSS 211
CRYBB3_BOVINE      WNEWDANQPQLQSVRRIRDQKWHKRGVFLSS 211
CRYBB3_SHEEP       WNEWDANQPQLQSVRRIRDQKWHKRGVFLSS 211
                   ******.****************** * **
```

Figure 2.9 Sequence alignment of human, pig, cow and sheep βB3 proteins using Clustal Omega. Aspartic acid residues in red are conserved in each species and were found to be isomerized in only the WI fraction of the pig, cow and sheep lenses. His113 and Arg196 (also in red) are conserved amongst each of the species as well. Sequence homology is over 90% between pig/sheep/cow compared to human.

The same analysis was carried out on pig and cow lenses and the results are summarized with the sheep results in Figure 2.10. The percent isomerization per protein was determined by taking the sum of the percent of isomerization per peptide and diving it by the total

number of peptides, including a value of zero for those which were not isomerized. Peptides where the degree of isomerization could not be determined were omitted for all samples of the same protein. For example, tryptic digestion of 173 residue αA yielded twelve peptides that cover 92% of the sequence. Contributions from these same peptides were used to calculate the numbers for each of the WS and WI αA digests for each animal. The data was then normalized to the αA digest that was most isomerized, which was the water-insoluble sheep digest. The same approach was applied to αB and βB3. Although there is some variation between species, in general similar results were obtained for all three organisms. The degree of isomerization is on average significantly greater in the WI fractions for all species. Specific data from the WI sheep fraction is shown in Table 2.2.

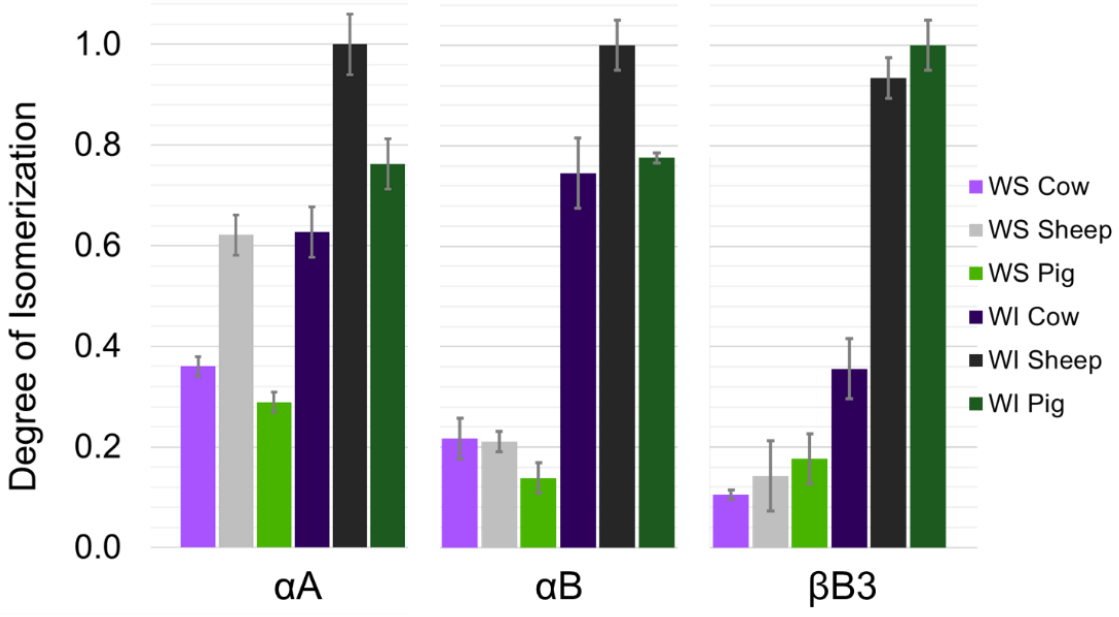


Figure 2.10 Total normalized degree of isomerization amongst the three major crystallin proteins detected.

Table 2.2 Tryptic peptides found to be isomerized in the three major proteins from water-insoluble sheep digest. Single letter codes are used for amino acids. Underlined residues represent the most likely sites of epimerization. Bold residues correspond to most likely sites of isomerization. pQ represents N-terminal cyclization of glutamine to form pyroglutamic acid which is a common product of terminal glutamines during digestion. Ac- is N-terminal acetylation.

Peptide Sequence	Crystallin	Number of peaks in LC separation	Relative abundance (%)	Number of peptide isomers confirmed by MS/MS
Ac- ^{1}MDIAIQHPWFK^{11}	αA	4	5.9%, 1.6%, 90.9%, 1.6%	4
^{13}TLGPFYPSR^{21}	αA	1		2
^{22}LFDQFFGEGLFEYDLLPFLSSTISPYYR^{49}	αA	4	1.7%, 10.3%, 31.4%, 56.6%	4
^{55}TVLDSGISEVR^{65}	αA	4	42.4, 0.8, 53.6%, 3.2%	4
^{69}DKFVIFLDVK^{78}	αA	2	6.2%, 93.8%	2
^{79}HFSPEDLTVK^{88}	αA	2	20.1%, 79.9%	3
^{89}VQEDFVEIHGK^{99}	αA	2	10%, 90%	2
^{104}pQDDHGYISR^{112}	αA	6	0.6%, 8.1%, 20.9%, 14.0%, 3.4%, 9.6%, 43.4%	5
^{118}YRLPSNVDQSALSCSLSADGMLTFSGPK^{145}	αA	2	11.2%, 88.8%	3
^{146}VPSGVDAGHSER^{157}	αA	3	12.4%, 44.8%, 42.8%	3
^{158}AIPVSR^{163}	αA	1		2
^{164}EEKPSSAPSS^{173}	αA	6	13.4%, 0.6%, 72.7%, 3.4%, 1.2%, 8.7%	6
Ac- ^{1}MDIAIHHPWIR^{11}	αB	2	15.0%, 85.0%	2
^{23}LFDQFFGEHLLESDLFPASTSLSPF^{47}	αB	2	15.1%, 84.9%	2
^{57}APSWIDTGLSEMR^{69}	αB	6	3.2%, 41.8%, 4.2%, 45.3%, 3.7%, 1.8%	6
^{83}HFSPEELK^{90}	αB	1		2
^{93}VLGDVIEVHGK^{103}	αB	1		2
^{108}pQDEHGFISR^{116}	αB	3	2.1%, 16.0%, 81.9%	3
^{124}IPADVDPLTITSSLSSDGVLTVNGPR^{149}	αB	1		2
Ac- ^{1}AEQHSAPEQAAAGK^{15}	βB3	2	9.0%, 91.0%	2
^{39}CELTAECPNLTESLLEK^{55}	βB3	2	18.2%, 81.8%	2
^{89}WDAWSNSHHSDSLL^{102}	βB3	1		2
^{103}SLRPLHIDGPDHK^{115}	βB3	2	12.5%, 87.5%	3
^{128}KMEIVDDDVPSLW^{140}	βB3	4	5.8%, 15.2%, 13.7%, 65.3%	4
^{180}HWNEWDANQPQLQSVR^{195}	βB3	1		2

2.3.1 Multiple Aspartic Acid Effect.

Several of the peptides in Figures 2.3, 2.5, and 2.7 appear to exhibit an unusual degree of isomerization relative to their peers. For example, ^{104}QDDHGYISR^{112} from αA is not only the most isomerized site in the protein, but it is also significantly more isomerized than any other peptide in the structured α-crystallin domain. Quantitatively, the average amount of isomerization per peptide in the structured domain of αA is 17.9% for the WS digest and 27.3% in the WI digest. At 76.7% and 89.3% isomerization in the WS and WI fractions, ^{104}QDDHGYISR^{112} is isomerized ~4*x* the average rate in the ordered crystallin region. The analogous sequence in αB, ^{108}QDEHGFISR^{116}, is also the most isomerized peptide extracted from the WI fraction, although full quantitative comparisons cannot be made due to lack of separation of some isomers for other peptides in the ordered region (see Figure 2.11 and Table 2.3 for details). Finally, ^{128}KMEIVDDDVPSLW^{140} is the most isomerized peptide in the WI fraction from βB3, which is explained. Each of these peptides share a common feature, sequential repeats of acidic residues. In two cases, multiple sequential aspartic acid residues are present. These observations suggest that sequential acidic residues represent sites of greater propensity for isomerization.

There are at least two possible explanations that could account for increased isomerization at acid residue repeats. First, if there are multiple aspartic acid residues, then isomerization at multiple sites is possible. Structural perturbation engendered by the first modification could easily lead to increased local backbone flexibility, facilitating isomerization at additional sites. Indeed, examination of the elution profiles for these peptides reveals an abundance of isomers, indicating modification at more than one

residue. For example, the elution profiles for the WS and WI fractions from pig for ^{104}pQDDHGYSIR^{112} reveal the presence of seven isomers, see Figure 2.12. The lowercase "p" at the N-terminus of the sequence denotes pyroglutamate, which is a common product during trypsin digestion. MS^2 experiments confirm that each peak represents a different isomer. Comparison with synthetically prepared all L-Asp isomer was used to confirm the peak corresponding to the canonical peptide. R_{isomer} values (4.8, 5.5, 3.7, 4.6, 4.5, 1.2 and 3.4) relative to the synthetic version confirm that the sixth peak in the digest is the all L-Asp isomer (see Figure 2.13 and Table 2.4 for details). The WS/WI degree of isomerization for ^{104}pQDDHGYSIR^{112} from sheep is shown in Figures 2.12c and 2.12d. A similar elution profile is seen between WI pig and sheep, however the amount of L-Asp in the WS pig lens is nearly twice as great as it is in WS sheep.

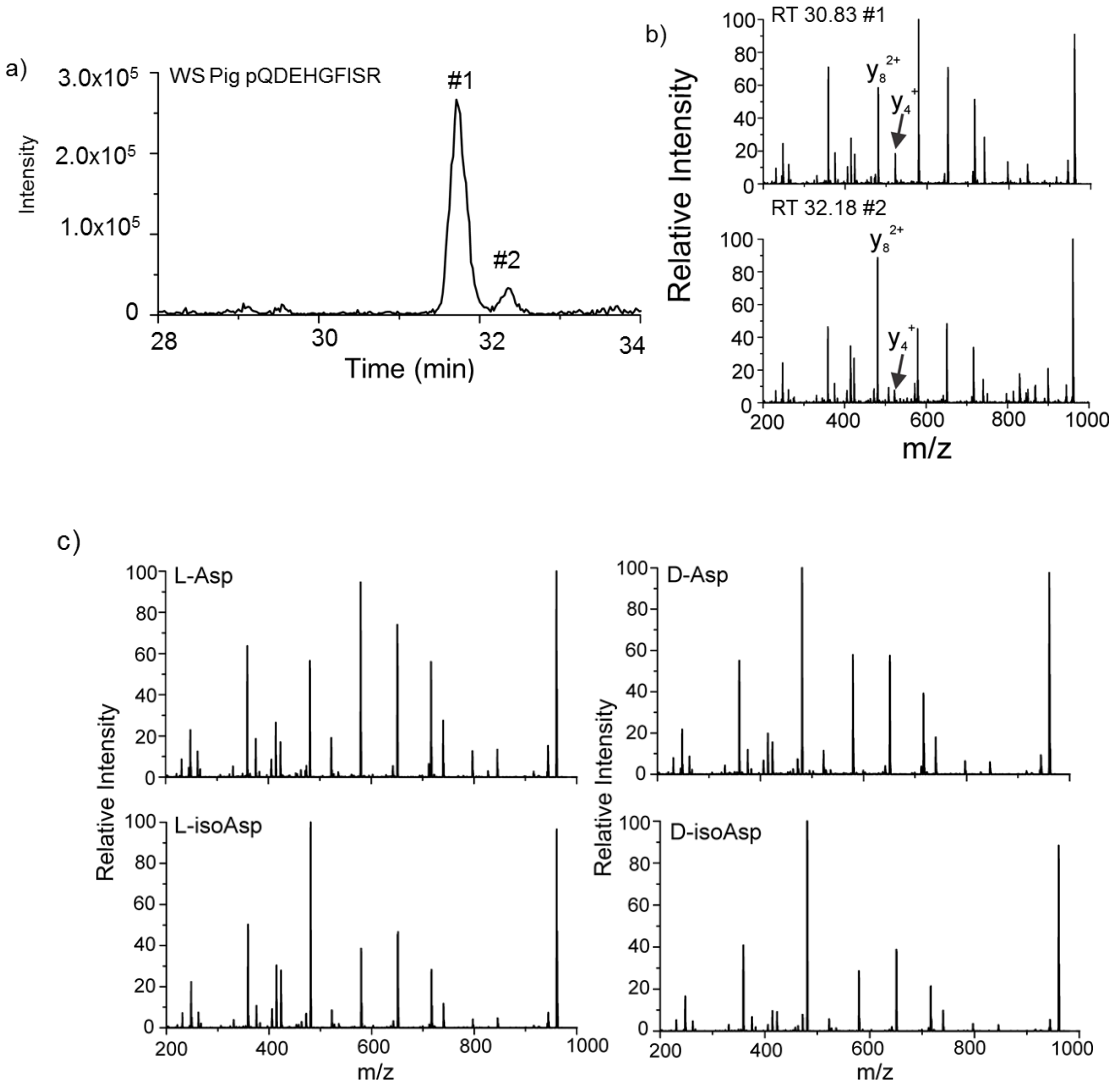


Figure 2.11. **a)** Chromatogram of the pQDEHGFISR peptide from the WS pig digest. **b)** Comparison of the CID fragmentation spectra of the two peaks shows notable differences between the y_8^{2+} and y_4^{+} fragments. The R_{isomer} score is 3.6, which verifies that they are isomers. **c)** All four aspartic acid isomers of this peptide were synthesized and CID analysis was performed to determine which forms were present in the digest.

Table 2.3 R_{isomer} values for each of the synthetic pQDEHGFISR versus the digest peaks from the WS pig digest. For each peak, only one of the synthetic isomers matched, meaning the R_{isomer} score was below the 1.9 threshold (shown in bold).

Spectrum 1	Spectrum 2	R_{isomer}
L-Asp Synthetic	**Peak 1**	**1.5**
D-Asp Synthetic	Peak 1	3.6
L-isoAsp Synthetic	Peak 1	4.4
D-isoAsp Synthetic	Peak 1	5.8
Spectrum 1	**Spectrum 2**	**R_{isomer}**
L-Asp Synthetic	Peak 2	3.6
D-Asp Synthetic	Peak 2	2.2
L-isoAsp Synthetic	**Peak 2**	**1.7**
D-isoAsp Synthetic	Peak 2	3.8

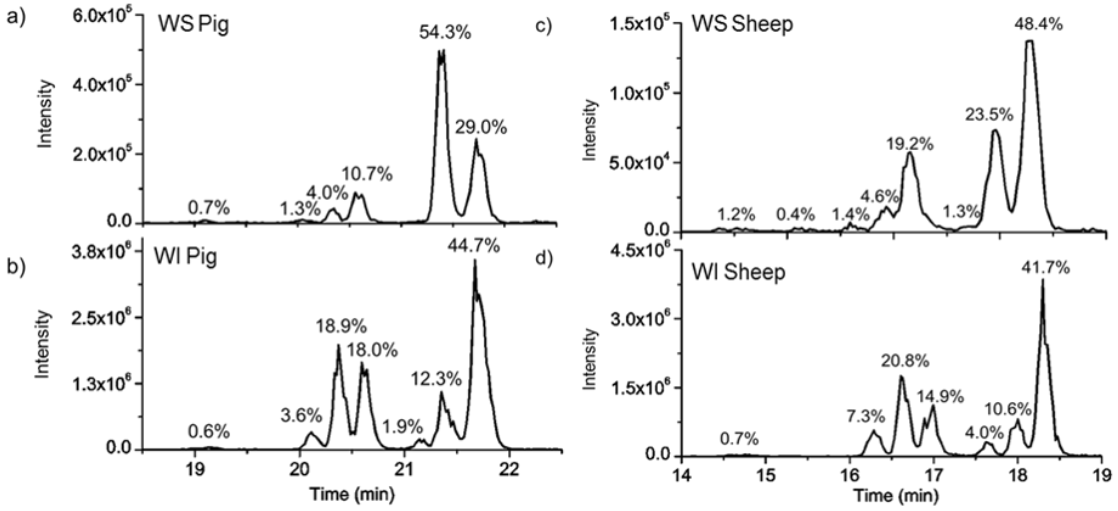


Figure 2.12 a) WS chromatogram and b) WI chromatogram for ^{104}pQDDHGYSIR^{112} from pig. c) WS Sheep. d) WI Sheep. Lower case "p" indicates the pyroglutamate that forms during tryptic digestion of N-terminal glutamine residues. Isomers were confirmed by MS/MS analysis and are labeled with percent abundance.

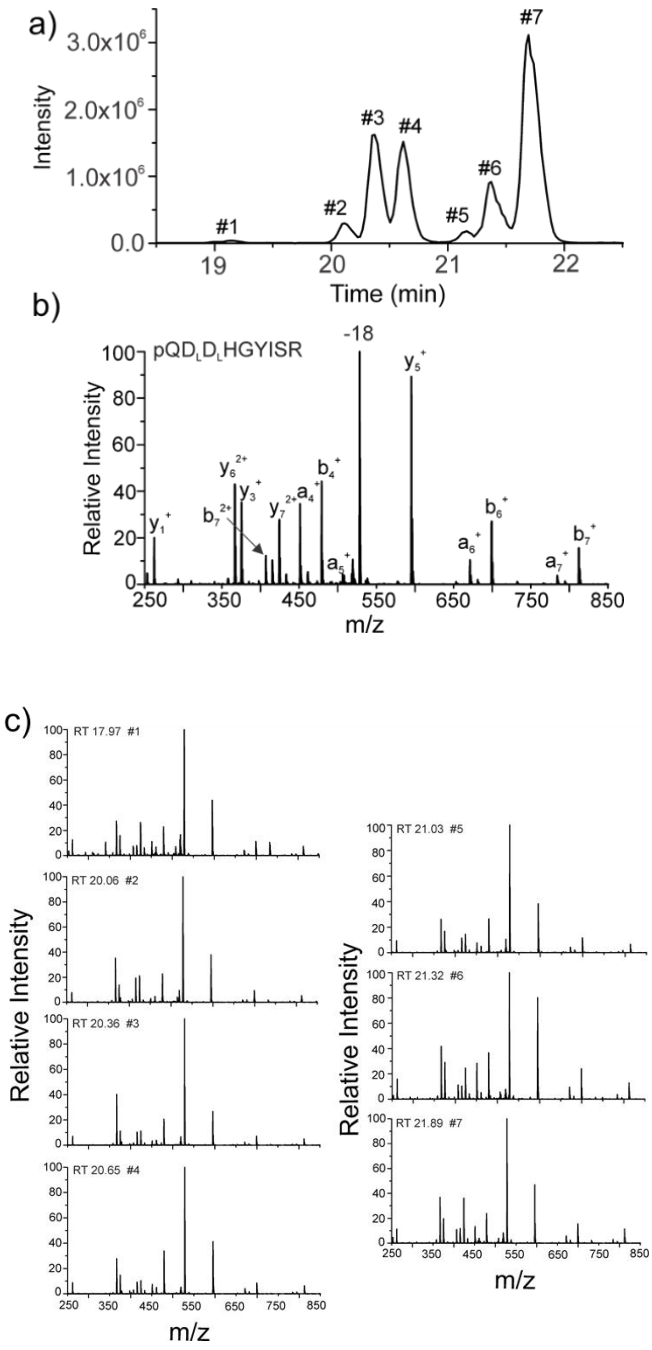


Figure 2.13 a) WI Pig LC chromatogram of pQDDHGYISR peptide, where numbers represent detected isomers. **b)** Resulting CID spectrum from the all L-Asp synthetic of version of the peptide. **c)** CID spectra from each of the peaks in the WI digest chromatogram. Notable differences can be seen between each, especially #6 which has an unusually large y_5^+ peak.

Table 2.4 R_{isomer} values for the L-Asp synthetic version of pQDDHGYSIR versus each of the peaks in the LC. Only one of the resulting values is lower than the 1.9 threshold (shown in bold) which helps us distinguish this one as the normal form of the peptide. This is also the peak that decreases substantially in abundance from the WS to the WI fractions

Spectrum 1	Spectrum 2	R_{isomer}
L-Asp Synthetic	Peak 1	4.8
L-Asp Synthetic	Peak 2	5.5
L-Asp Synthetic	Peak 3	3.7
L-Asp Synthetic	Peak 4	4.6
L-Asp Synthetic	Peak 5	4.5
L-Asp Synthetic	**Peak 6**	**1.2**
L-Asp Synthetic	Peak 7	3.4

Another factor contributing to isomerization at sites with multiple acidic residues, which may be more important than increased backbone flexibility, is inhibited repair by protein isoaspartyl methyltransferase (PIMT). PIMT is the only known repair enzyme that targets protein damage caused by aging. It is present and active in bovine lenses and reverts L-isoAsp back to L-Asp.[43,44] D-Asp is also repaired, though much less efficiently, but D-isoAsp is not a substrate.[45] Importantly, sequence effects for substrate recognition have been identified and revealed that PIMT has substantially lower affinity for isoAsp if the n+1, n+2 or n+3 residues are negatively charged or if the n+1 site is proline.[46] Therefore, the sequence regions with multiple acidic resides that we have observed to be highly modified are unlikely to be repaired by PIMT, and our results confirm the importance of this enzyme in the repair of long-lived proteins.

Acidic residue repeat sites may therefore be prone to sequential isomerization and difficult to repair, a potent combination with important consequences for protein aging. Indeed, examination of other long-lived proteins associated with age-related diseases

reveals the presence of numerous acidic acid repeat sites that would similarly be susceptible to isomerization. For example, amyloid precursor protein associated with Alzheimer's disease contains 27 acidic residue repeat sites that would be poor substrates for PIMT. Table 2.5 lists long-lived proteins that contain aspartic acid repeats, are aggregation-prone, and are associated with diseases. It is clear that the results obtained here may be relevant on a much broader scale for any system comprised of long-lived proteins.

Table 2.5 Aggregation-prone proteins associated with various diseases.

Protein	tau-protein P10636-1	Amyloid Precursor Protein (APP) P05067-1	Amyloid-Beta (1-42)	Alpha-synuclein P37840-1	TAR TDP-43 Q13148-1	Huntingtin P42858-1	Alpha-A crystallin P02489	Alpha-B crystallin P02511	Beta-B3 Crystallin P26998-1
Disease	PD[a]	AD[b]	AD	PD/AD	ALS[c]	HD[d]	Cat[e]	Cat	Cat
Length	758	770	42	140	414	3,142	173	175	211
% Asp	5.8	6.8	7.1	4.3	5.3	4.4	8.8	6.3	6.2
"DD" or "DDD"	"DD" x1	"DD" x4 & "DDD" x1	None	None	"DD" x2	"DD" x5 & "DDD" x1	"DD" x2	None	"DDD" x1
Acidic repeats[f]	16	27	1	1	9	44	4	3	2
% Ser	10.6	4.5	4.8	2.9	9.9	9.7	10.4	9.7	8.1

a) Parkinson's Disease, b) Alzheimer's Disease, c) Amyotrophic Lateral Sclerosis, d) Huntington's Disease, e) cataracts, f) number of asp residues with an acidic residue in the n+1, n+2 or n+3 position

We have also identified isomerized regions with multiple serine residues, such as ^{164}EEKPSSAPSS^{173}. In the WS and WI fraction of the sheep lens this peptide elutes in six different peaks (Figure 2.14), and comparison of resulting CID spectra confirms they represent different isomers. It is interesting to note that the total amount of isomerization in the WI sheep fraction for ^{164}EEKPSSAPSS^{173} is much lower than what was observed for ^{104}QDDHGYISR^{112}, 27.3% in compared to 90.4%, respectively. It is possible that isomerization of aspartic acid may have a larger influence on local structure than epimerization of serine, inducing isomerization of nearby residues more efficiently. Alternatively, there is no known repair enzyme for epimerization of serine, which may suggest that the failure of PIMT to repair isomerization may be the most important factor influencing isomerization at acidic residue repeats. However, the number of peptides with unambiguous epimerization at serine, i.e. peptides that are isomerized and do not contain aspartic acid, is relatively small. Therefore, more data will need to be acquired before strong conclusions about serine epimerization can be drawn.

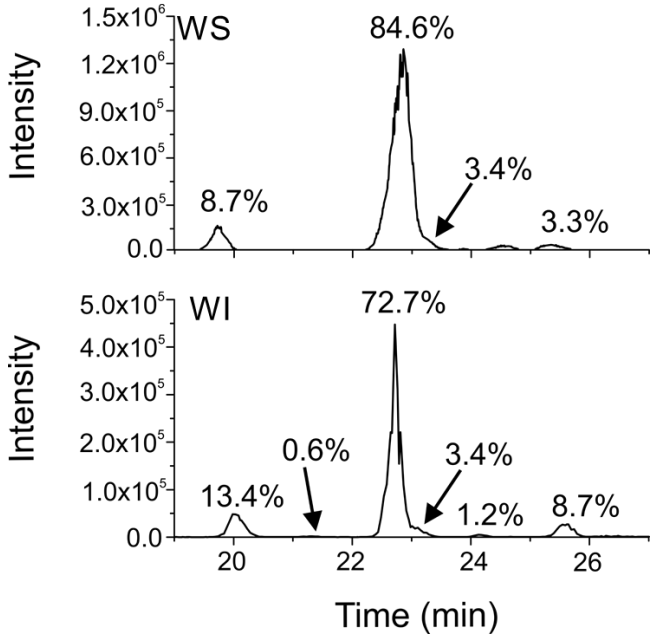


Figure 2.14 Comparison of the degree of isomerization in ^{164}EEKPSSAPSS^{173} between WS and WI sheep LC chromatograms.

2.4 Conclusions

Peptide epimerization and isomerization are difficult to detect and remain among the least studied PTMs despite their potentially important role in numerous diseases. Herein, we have evaluated these modifications in crystallin proteins from the eye lenses of several organisms, and our results suggest that isomerization and epimerization lead to reduced protein solubility in many cases. Although there is variation between species, several common themes emerge from analysis of the data. For example, isomerization occurs more readily in regions of proteins that are disordered. This finding suggests that many other long-lived proteins with known disordered regions, such as α-synuclein, tau, and β-amyloid, may also exhibit pathology related to isomerization or epimerization. Although less prone to modification, isomerization within well-structured regions of proteins leads to more drastic changes in behavior, including unchecked aggregation. It

was previously known that serine and aspartic acid are most easily isomerized among the natural amino acids, but we have further demonstrated that when multiple acidic residues are close in sequence, the propensity for modification is enhanced further. We postulate that increased flexibility following an initial modification and failure of PIMT to repair damaged residues contribute to the dramatic isomerization of sites with multiple acidic residues. The degree of isomerization at these sites also confirms the importance of PIMT for long-term maintenance of protein structure. Proteins associated with Parkinson's disease, Alzheimer's disease, Amyotrophic lateral sclerosis and Huntington's disease all contain regions with multiple aspartic acid and/or serine residues in close proximity, suggesting these proteins may also exhibit hotspots of isomerization.

Both αA and αB exhibit similar structural and functional behavior, yet αA is only found in the eye lens. The isomerization behavior of αA and αB are quite distinct, with αA being more prone to isomerization, but with αB suffering greater loss of solubility following isomerization. The ability of αA to sustain isomerization without loss of solubility may make it uniquely suited to the zero-turnover environment of the eye lens. It is clear that further, detailed, examination of isomerization and epimerization in other tissues is needed and will expand our understanding of the role of long-lived proteins in aging-related diseases.

---

(1) Toyama, B. H.; Hetzer, M. W. Protein homeostasis: live long, won't prosper. *Nat. Rev. Mol. Cell Biol.* **2013**, *14* (1), 55–61.
(2) Bassnett, S. Lens organelle degradation. *Exp. Eye Res.* **2002**, *74* (1), 1–6.
(3) Horwitz, J.; Bova, M. P.; Ding, L. L.; Haley, D. a; Stewart, P. L. Lens alpha-crystallin: function and structure. *Eye (Lond).* **1999**, *13 (Pt 3b*, 403–408.

(4) Oertel, M. F.; May, C. A.; Bloemendal, H.; Lutjen-Drecoll, E. Alpha-B-crystallin expression in tissues derived from different species in different age groups. *Ophthalmologica* **2000**, *214* (1), 13–23.
(5) Brady, J. P.; Garland, D.; Duglas-Tabor, Y.; Robison, W. G.; Groome, A.; Wawrousek, E. F. Targeted disruption of the mouse alpha A-crystallin gene induces cataract and cytoplasmic inclusion bodies containing the small heat shock protein alpha B-crystallin. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94* (3), 884–889.
(6) Brady, J. P.; Garland, D. L.; Green, D. E.; Tamm, E. R.; Giblin, F. J.; Wawrousek, E. F. alphaB-crystallin in lens development and muscle integrity: A gene knockout approach. *Investig. Ophthalmol. Vis. Sci.* **2001**, *42* (12), 2924–2934.
(7) Srinivas, P.; Narahari, A.; Petrash, J. M.; Swamy, M. J.; Reddy, G. B. 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.
(8) Berbers, G.A..; Hoekman, W.A.; Bloemendal, H.; de Jong, W.W.; Kleinschmidt, T.; Braunitzer, G. Homology between the primary structures of the major bovine beta-crystallin chains. *Eur. J. Biochem.* **1984**, *139* (3), 467–479..
(9) Mahendiran, K.; Elie, C.; Nebel, J.-C.; Ryan, A.; Pierscionek, B. K. Primary sequence contribution to the optical function of the eye lens. *Sci. Rep.* **2014**, *4*, 5195.
(10) Baldwin, A. J.; Lioe, H.; Robinson, C. V.; Kay, L. E.; Benesch, J. L. P. ??B-Crystallin Polydispersity Is a Consequence of Unbiased Quaternary Dynamics. *J. Mol. Biol.* **2011**, *413* (2), 297–309.
(11) Bagnéris, C.; Bateman, O. A.; Naylor, C. E.; Cronin, N.; Boelens, W. C.; Keep, N. H.; Slingsby, C. Crystal Structures of α-Crystallin Domain Dimers of αB-Crystallin and Hsp20. *J. Mol. Biol.* **2009**, *392* (5), 1242–1252.
(12) Laganowsky, A.; Benesch, J. L. P.; Landau, M.; Ding, L.; Sawaya, M. R.; Cascio, D.; Huang, Q.; Robinson, C. V; Horwitz, J.; Eisenberg, D. Crystal structures of truncated alphaA and alphaB crystallins reveal structural mechanisms of polydispersity important for eye lens function. *Protein Sci.* **2010**, *19*, 1031–1043.
(13) Hanson, S. R.; Hasan, A.; Smith, D. L.; Smith, J. B. 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.
(14) Lund, A. L.; Smith, J. B.; Smith, D. L. Modifications of the water-insoluble human lens alpha-crystallins. *Exp. Eye Res.* **1996**, *63* (6), 661–672.
(15) Hooi, M. Y. S.; Raftery, M. J.; Truscott, R. J. W. Accelerated aging of Asp 58 in alphaA crystallin and human cataract formation. *Exp. Eye Res.* **2013**, *106*, 34–39.
(16) Masters, P. M.; Bada, J. L.; Zigler, J. S. Aspartic acid racemization in heavy molecular weight crystallins and water insoluble protein from normal human lenses and cataracts. *Proc. Natl. Acad. Sci. U. S. A.* **1978**, *75* (3), 1204–1208.
(17) Fujii, N.; Takata, T.; Fujii, N.; Aki, K. Isomerization of aspartyl residues in crystallins and its influence upon cataract. *BBA - Gen. Subj.* **2016**, *1860* (1), 183–191.
(18) Jia, C.; Lietz, C. B.; Yu, Q.; Li, L. Site-Specific Characterization of <scp>d</scp> -Amino Acid Containing Peptide Epimers by Ion Mobility Spectrometry. *Anal. Chem.* **2014**, *86* (6), 2972–2981.

(19) O'Connor, P. B.; Cournoyer, J. J.; Pitteri, S. J.; Chrisman, P. A.; McLuckey, S. A. Differentiation of aspartic and isoaspartic acids using electron transfer dissociation. *J. Am. Soc. Mass Spectrom.* **2006**, *17* (1), 15–19.
(20) Hooi, M. Y. S.; Truscott, R. J. W. 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 (Omaha).* **2011**, *33* (2), 131–141.
(21) Takahashi, O.; Kirikoshi, R.; Manabe, N. Racemization of the succinimide intermediate formed in proteins and peptides: A computational study of the mechanism catalyzed by dihydrogen phosphate ion. *Int. J. Mol. Sci.* **2016**, *17* (10), 10–15.
(22) Demarchi, B.; Collins, M.; Bergström, E.; Dowle, A.; Penkman, K.; Thomas-Oates, J.; Wilson, J. New Experimental Evidence for In-Chain Amino Acid Racemization of Serine in a Model Peptide. *Anal. Chem.* **2013**, *85* (12), 5835–5842.
(23) Takahashi, O. Two-water-assisted racemization of the succinimide intermediate formed in proteins. A computational model study. *Health (Irvine. Calif).* **2013**, *5* (12), 2018–2021.
(24) Noguchi, S.; Miyawaki, K.; Satow, Y. Succinimide and isoaspartate residues in the crystal structures of hen egg-white lysozyme complexed with tri-N-acetylchitotriose. *J. Mol. Biol.* **1998**, *278* (1), 231–238.
(25) Noguchi, S. Structural changes induced by the deamidation and isomerization of asparagine revealed by the crystal Structure of Ustilago sphaerogena ribonuclease U2B. *Biopolymers* **2010**, *93* (11), 1003–1010.
(26) Rehder, D. S.; Chelius, D.; McAuley, A.; Dillon, T. M.; Xiao, G.; Crouse-Zeineddini, J.; Vardanyan, L.; Perico, N.; Mukku, V.; Brems, D. N.; et al. Isomerization of a Single Aspartyl Residue of Anti-Epidermal Growth Factor Receptor Immunoglobulin γ2 Antibody Highlights the Role Avidity Plays in Antibody Activity. *Biochemistry* **2008**, *47* (8), 2518–2530.
(27) Tao, W. A.; Zhang, D.; Nikolaev, E. N.; Cooks, R. G. Copper(II)-assisted enantiomeric analysis of D,L-amino acids using the kinetic method: Chiral recognition and quantification in the gas phase. *J. Am. Chem. Soc.* **2000**, *122* (43), 10598–10609.
(28) Sun, Q. Y.; Nelson, H.; Ly, T.; Stoltz, B. M.; Julian, R. R. Side Chain Chemistry Mediates Backbone Fragmentation in Hydrogen Deficient Peptide Radicals. *J. Proteome Res.* **2009**, *8* (2), 958–966.
(29) Tao, Y.; Quebbemann, N. R.; Julian, R. R. Discriminating d-amino acid-containing peptide epimers by radical-directed dissociation mass spectrometry. *Anal. Chem.* **2012**, *84* (15), 6814–6820.
(30) Tao, Y.; Julian, R. R. Identification of amino acid epimerization and isomerization in crystallin proteins by tandem LC-MS. *Anal. Chem.* **2014**, *86* (19), 9733–9741.
(31) Hood, C. A.; Fuentes, G.; Patel, H.; Page, K.; Menakuru, M.; Park, J. H. Fast conventional Fmoc solid-phase peptide synthesis with HCTU. *J. Pept. Sci.* **2008**, *14* (1), 97–101.
(32) Ly, T.; Zhang, X.; Sun, Q. Y.; Moore, B.; Tao, Y. Q.; Julian, R. R. Rapid, quantitative, and site specific synthesis of biomolecular radicals from a simple photocaged precursor. *Chem. Commun.* **2011**, *47* (10), 2835–2837.

(33) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7* (1), 539.
(34) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—A Visualization System for Exploratory Research and Analysis. *J Comput Chem* **2004**, *25*, 1605–1612.
(35) Maestro, ver. 9.7. Schrödinger, LLC: New York, NY (2012)
(36) Kim, K. K.; Kim, R.; Kim, S. H. Crystal structure of a small heat-shock protein. *Nature* **1998**, *394* (6693), 595–599.
(37) Kumar, L. V. S.; Rao, C. M. Domain swapping in human alphaA and alphaB crystallins affects oligomerization and enhances chaperone-like activity. *J. Biol. Chem.* **2000**, *275* (29), 22009–22013.
(38) Clarke, S. Propensity for spontaneous succinimide formation from aspartyl and asparaginyl residues in cellular proteins. *Int. J. Pept. Protein Res.* **1987**, *30* (6), 808–821.
(39) Takata, T.; Fujii, N. Isomerization of Asp residues plays an important role in αA-crystallin dissociation. *FEBS J.* **2016**, *283* (5), 850–859.
(40) Sakaue, H.; Takata, T.; Fujii, N.; Sasaki, H.; Fujii, N. Alpha B- and βA3-Crystallins Containing D-Aspartic Acids Exist in a Monomeric State. *Biochim. Biophys. Acta - Proteins Proteomics* **2015**, *1854* (1), 1–9.
(41) Robinson, N. E.; Robinson, Z. W.; Robinson, B. R.; Robinson, A. L.; Robinson, J. A.; Robinson, M. L.; Robinson, A. B. Structure-dependent nonenzymatic deamidation of glutaminyl and asparaginyl pentapeptides. *J. Pept. Res.* **2004**, *63* (5), 426–436.
(42) Slingsby, C.; Wistow, G. J.; Clark, A. R. Evolution of crystallins for a role in the vertebrate eye lens. *Protein Sci.* **2013**, *22* (4), 367–380.
(43) McFadden, P. N.; Horwitz, J.; Clarke, S. Protein carboxyl methyltransferase from cow eye lens. *Biochem. Biophys. Res. Commun.* **1983**, *113* (2), 418–424.
(44) McFadden, P. N.; Clarke, S. Chemical conversion of aspartyl peptides to isoaspartyl peptides. A method for generating new methyl-accepting substrates for the erythrocyte D-aspartyl/L-isoaspartyl protein methyltransferase. *J. Biol. Chem.* **1986**, *261* (25), 11503–11511.
(45) Lowenson, J. D.; Clarke, S. Recognition of D-aspartyl residues in polypeptides by the erythrocyte L- isoaspartyl/D-aspartyl protein methyltransferase. Implications for the repair hypothesis. *J. Biol. Chem.* **1992**, *267* (9), 5985–5995.
(46) Lowenson, J. D.; Clarke, S. Structural elements affecting the recognition of L-isoaspartyl residues by the L-isoaspartyl/D-aspartyl protein methyltransferase. Implications for the repair hypothesis*. *J. Biol. Chem.* **1991**, *266* (29), 19396–19406.

Chapter 3

Differences in α-Crystallin Isomerization Reveal the Activity of Protein Isoaspartyl Methyltransferase (PIMT) in the Nucleus and Cortex of Human Lenses

3.1 Introduction

Proteins within the human lens help provide the requisite refractive index needed for sight and inhibit molecular aggregation to particle sizes capable of scattering light (Horwitz et al., 1998; Rao et al., 1995). Given the lack of organelles within mature lens fiber cells, these functions must be performed in the absence of significant protein renewal or turnover (Samuel Zigler and Goosey, 1981). The proteins in the lens must therefore remain functional over significant lengths of time. The human lens also continues to grow throughout life, with newly synthesized cells continuously adding to the periphery (Augusteyn, 2007). Consequently, the approximate age of proteins within the lens is determined by their spatial location. The oldest proteins will be located in the central nucleus, while younger proteins will be found in the peripheral cortex. Crystallins constitute nearly 90% of the total water-soluble protein content in the lens fiber cells (Groenen et al., 1994). Protein degradation due to aging is thought to lead to loss of function in diseases such as cataracts and presbyopia, but complete characterization of all age-related protein post-translational modifications (PTMs) that might contribute to these maladies has proven elusive.

Among the major protein decay pathways, the mass-shifting PTMs in crystallins have been studied extensively, including disulfide bond formation, oxidation, phosphorylation,

deamidation, and truncation (Ma et al., 1998). More subtle modifications such as epimerization and isomerization, which do not yield easily detectable mass-shifts, have received less attention. Epimerization is a specific type of isomerization that occurs when any single residue in a peptide converts from the L to D configuration at the alpha carbon position. Isomerization can also occur without affecting chirality, which in peptides occurs primarily through the formation of iso-aspartic acid (isoAsp). Importantly, recent work has suggested that, collectively, isomerization is the most prevalent class of PTM in the lens(Truscott and Friedrich, 2016). Although difficult to detect, isomerization significantly impacts protein integrity, contributing to loss of solubility and function by altering structure (Fujii et al., 2012, 2016; Lund et al., 1996; Masters et al., 1978) A study of α-crystallins in sheep lenses revealed that isomerization was highest in the disordered termini, which serve as intermolecular bridges in the formation of highly dynamic oligomers (Tao and Julian, 2014). It has also been shown that isomerization in the structured crystallin domain leads to reduced solubility (Lyon et al., 2017)

Aspartic acid is most susceptible to isomerization among the canonical amino acids, and asparagine can produce the same products via deamidation. The vulnerability of these residues is attributed to the susceptibility of the side-chain γ-carbonyl carbon to undergo nucleophilic attack from the backbone nitrogen of the C-terminal amino acid (Fig. 3.1). The metastable L-succinimide ring thus formed can be hydrolyzed at one of two positions to form L-Asp or L-isoAsp. The increased acidity of the alpha-H in the succinimide ring also facilitates racemization to form D-succinimide and subsequent hydrolysis yields either D-Asp or D-isoAsp. Over time, sequences containing L-Asp will

spontaneously convert in four isomeric forms. Both of the isoAsp products reroute the protein backbone through the β-carbonyl of the sidechain, which induces major perturbations in protein structure (Noguchi, 2010). Protein L-isoaspartyl methyltransferase (PIMT) is the only known repair enzyme for age-related Asp isomerization. PIMT is an *S*-adenosyl-L-methionine (SAM)- dependent enzyme that methylates L-isoAsp residues and D-Asp (700-10,000-fold lower affinity), allowing partial reformation of the L-Asp form (Lowenson and Clarke, 1992; McFadden and Clarke, 1987). Gene knock-out studies have shown that PIMT-deficient mice exhibit a nine-fold increase of L-isoAsp in the brain and die from epileptic seizures between 4 and 12 weeks (Qin et al., 2015; Yamamoto et al., 1998), while overexpression of PIMT in *Drosophila* increased lifespan by nearly 30% (Chavous et al., 2001) .

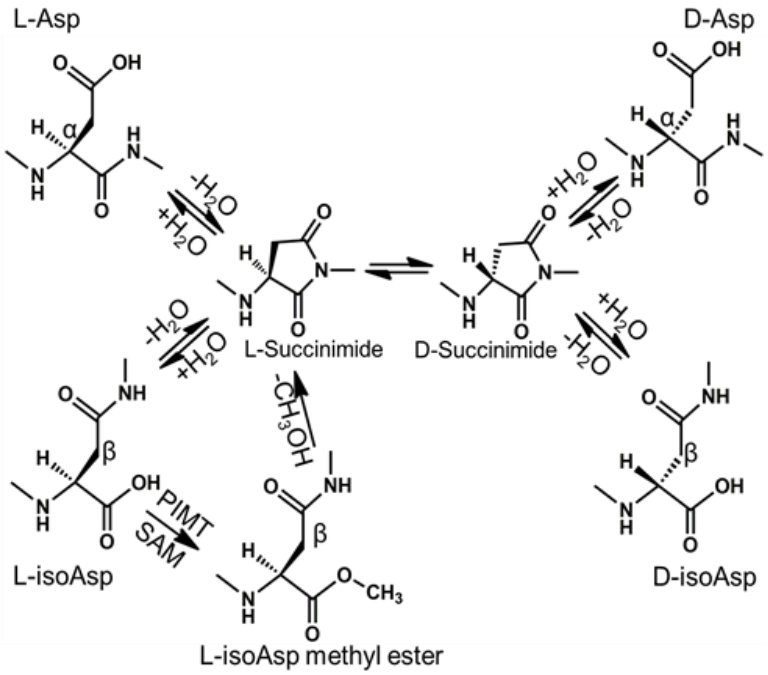


Figure 3.1 Spontaneous aspartic acid isomerization via a succinimide intermediate. Alpha and beta carbons are labeled. PIMT conversion of L-isoAsp back to the L-succinimide intermediate via L-isoAsp methyl ester.

Isomerization is an inherently difficult PTM to detect, as isomers have similar chemical properties and identical mass. Nevertheless, new gas-phase techniques involving both ion mobility spectrometry (IMS) and tandem mass spectrometry have been developed to tackle this problem. For example, D-amino acid containing peptides (DAACP) can be site-specifically detected by changes in drift time using ion mobility spectrometry coupled to mass spectrometry IMS-MS (Jia et al., 2014). Ultrahigh resolution IMS-MS can separate amyloid beta peptides containing all four aspartyl isomers (Zheng et al., 2017). Mass spectrometry-based techniques focus on discerning isomers by creating diagnostic fragments or comparing changes in fragmentation intensities. Electron capture dissociation (ECD) and electron transfer dissociation (ETD)

can distinguish Asp from isoAsp through observation of unique mass c and z-type ions (O'Connor et al., 2006). In particular, radical-directed dissociation (RDD) has proven to be well-suited for isomer detection because it can be implemented in the analysis of semi-complex samples, such as the protein complement found in the lens. In RDD experiments, a radical is created site-specifically and is then activated to cause dissociation. The radical will migrate away from the initial site by pathways dictated by the three-dimensional peptide structure. Various side-chain and backbone fragments will be generated with isomer specific intensities, allowing for isomer identification in a high-throughput fashion (Tao et al., 2012). Thus, RDD enables full characterization of isomers and epimers, which are typically invisible to traditional proteomics.

Herein, the differences in isomerization from the cortex and nucleus of aged, human lenses are detailed. Prior to examination, the two regions of each lens were separated, and then further divided based on solubility. Enzymatic digestion was performed on each of the fractions and followed by tandem LC-MS using both collision-induced dissociation (CID) and RDD to distinguish isomers. Crystallin isomerization was studied as a function of age, and the results are analyzed in terms of location within the lens, sequence, tertiary structure, and PIMT activity. Additional experiments were carried out *in vitro* to establish the intrinsic isomerization propensities and influence of PIMT for comparison with the results obtained from the lens. The most important factors influencing age-induced isomerization are revealed and discussed in relation to maintenance of lens function.

3.2 Experimental Section

3.2.1 Protein Extraction and Digestion

Human lenses were acquired from the National Disease Research Interchange (NDRI) (Philadelphia, Pennsylvania). Each lens was snap frozen and sent on dry ice overnight. Upon arrival, the samples were stored at -20ºC until lens extraction was performed. The nuclei and cortices of the thawed lenses were separated using a 4mm trephine. The endcaps of each nucleus were removed by gently scraping off the outer tissue until only the dense, nuclear portion remained. The nucleus and cortex were then homogenized in 50mM Tris-HCl pH =7.8. The lens fractions were then centrifuged at 15,100g for 20 min at 4ºC to separate the supernatant from the precipitate. The supernatant (water-soluble) was purified by dialysis against water. The precipitate (water-insoluble) was solubilized in 6 M urea and purified by dialysis against 6 M urea. For water-soluble digestion, 50 µg of protein was dissolved in 50 mM NH_4HCO_3 buffer, pH =7.8, disulfide bonds were then reduced with 1.5 µL of 100 mM DTT at 70°C for 10 minutes. After returning to room temperature, reduced cysteines were capped using 3 µL of 100 mM iodoacetamide in the dark for 20 minutes. Finally, the proteins were digested with trypsin for 12 hours at 37°C using a 50:1 protein to enzyme ratio. For the water-insoluble digestion, 100 µg of protein was dissolved using 6 M in 50 mM Tris-HCl, pH= 8.0. Disulfide bonds were reduced using 5 µL of 200 mM DTT in Tris-HCl, pH=8.0 at 37°C for 20 minutes. Following this, 20 µL of 200 mM iodoacetamide in Tris-HCl, pH= 8.0 was added, and the mixture was incubated in the dark for one hour. To consume unreacted iodoacetamide, 20 µL of 200 mM DTT was added and incubated for one hour in the dark. Next, the urea concentration

was diluted to < 0.6 M using 50 mM Tris-HCl, 1 mM $CaCl_2$, pH= 7.6. The proteins were digested using trypsin with a 50:1 protein to enzyme ratio for 16 hours at 37°C. For the iodobenzoate modification, the digested peptides were desalted using a Michrom peptide trap (3 x 8 mm, C8) (Michrom Bioresource Inc). Approximately 5 nmoles of the digestion mixture, 10-100x excess of 4-iodobenzoic acid N-hydroxysuccinimide-activated ester in dioxane and borate buffer (pH =8.6) were combined and incubated for 1 hour at 37°C. Important: Dimethyl sulfoxide should not be substituted for dioxane in this step because it can cause aspartic acid isomerization. The modification side products at arginine and tyrosine side chains were removed by incubating the reaction mixture in 1 M hydroxylamine, pH= 8.5. The same procedure was used for the synthetic peptide standards. These procedures have been determined previously not to yield any detectable isomerization in control experiments (Tao and Julian, 2014).

3.2.2 Peptide and Radical Precursor Synthesis

All synthetic peptides were synthesized manually using standard FMOC SPPS procedures with Wang Resins as the solid support (Hood et al., 2008). N-hydroxysuccinimide (NHS) activated iodobenzoyl esters were synthesized by a previous procedure (Ly et al., 2011).

3.2.3 IQTGLDATHAER and IQTGLNATHAER Incubations

Synthetic model peptides were first purified using a Jupiter Proteo column (250 mm x 4.6 mm, 4 μm, 90 Å, C12) (Phenomenex, Torrence, CA). Following purification, 100 μL of 0.1mM model peptides were incubated in Tris-HCl, pH= 7.4 at 37ºC for 22-43 days in parafilm-wrapped centrifuge tubes.

3.2.4 PIMT Experiments

Enzymatic methylation reactions were carried out in Tris-HCl pH= 7.4 at 37ºC in a final volume of 50 µL containing 50 µM peptide substrate, 3.9 µM Protein isoaspartyl methyltransferase (Human, His tag, *E. coli*) (ATGen, Seongnam-si, South Korea), 400 µM S-5'-adenosyl-L-methionine (SAM) hydrochloride (Cayman Chemical, Ann Arbor, MI). Aliquots were sampled at 0, 0.45, 3, 6, and 24 hr. Reactions were quenched using 0.2% trifluoroacetic acid and stored at -20ºC.

3.2.5 Mass Spectrometry and Radical Directed Dissociation

Solutions were analyzed using an LTQ linear ion trap mass spectrometer (Thermo Fisher Scientific, San Jose, CA) with a standard electrospray ionization source. The posterior plate of the ion trap was modified with a quartz window to allow fourth-harmonic (266 nm) laser pulses from a flash-pumped Nd:YAG laser (Continuum, Santa Clara, CA). Photodissociation (MS^2) of the *para*-iodobenzoate labeled peptide induces homolytic dissociation of the carbon-iodine bond producing a radical peptide. This is done using a digital delay generator (Berkeley Nucleonics, San Rafael, CA) which synchronizes laser pulses to irradiate the trapped ion cloud without supplemental collision energy. Further collision-induced dissociation (MS^3) was performed on the radical by re-isolation of the largest peak in the photodissociation spectrum.

3.2.6 LC-MS Data Acquisition and Analysis

An Agilent 1100 series HPLC system (Agilent, Santa Clara, CA) with a BetaBasic column (150 mm x 2.1 mm, 3 µm, 150 Å, C18) was coupled to an LTQ mass spectrometer. Peptides were separated using a 0.1% formic acid in water (mobile phase

A) and a 0.1% formic acid in acetonitrile (mobile phase B) binary system at a flow rate of 0.2 mL/min. The digestion mixtures were loaded onto the column and separated using the following gradient: 5% B to 20% B over 60 minutes, 20% B to 30% B over the next 45 minutes, 30% B to 50% B over the next 15 minutes and 50% B to 95% B over the final 10 minutes. Another gradient of 3% B to 20% B over 80 minutes was also used to help separate a few coeluting peptides with very short retention times. The LTQ was operated in data-dependent mode using the Xcalibur program (Thermo Fisher Scientific). Specifically, in the CID-only LC-MS run, the first scan event was a full MS from m/z 300-2000 Da, followed by an ultrazoom (MS^2) and then CID (MS^3). In the RDD LC-MS experiments, the laser pulses were triggered during the MS^2 step and CID was performed as a pseudo-MS^3 step. Due to the high photodissociation yield of the 4-iodobenzoic acid chromophore, the major peak during this step is the loss of iodine, and it is the subsequent precursor for MS^3. The exclusion time was set to 60 seconds for the identification of peptides. For isomer identification, an inclusion mass list was added and the exclusion time was reduced to 16 seconds to enable repeated analysis of isomers. For both peptide and isomer identification, a rejection mass list of common mass spectrometry contaminants and interferences was applied to each experiment. Relative peptide isomer abundances were determined using the ICIS Peak Integration algorithm provided in Xcalibur and reported as the area of the single isomer peak over the cumulative area of each isomer.

MS data was acquired with Xcalibur software. The .raw files were converted to .mgf files using MSConvert software. The .mgf files were searched with X!Tandem

PILEDRIVER edition (version 2015.04.01.1) against the *Homo sapiens* UniprotKB database. The cleavage sites were set as lysine and arginine (semi cleavage was turned on), allowing up to two missed cleavages and one point mutation). Carbamidomethylation (+57.02 Da at Cys) was set as a fixed modification, and N-acetylation, phosphorylation and oxidation were all considered as variable modifications. For the iodobenzoate-modified digestion mixture, the modifier (+230.01 Da) was considered as a possible modification at either the N-terminus or lysine side chain. The parent monoisotopic mass error was set to ±1 Da, the minimum parent mass was set to 300 Da and the fragment mass error was set to ±0.4 Da. The criteria for accepting peptides identified by X! Tandem was $e < 0.005$. The false discovery rate is, on average 1.9%, calculated using decoy sequences created by reverse database searching.

3.2.7 Calculation of R Values

To quantify isomer identification, R values were calculated, as originally described by Tao and co-workers (Tao et al., 2000). In this paper, R_{isomer} represents the ratios of the relative intensities of a pair of fragments that vary most between two isomers (R_A/R_B). Following acquisition of a tandem mass spectrum, R_{isomer} values are calculated for all pairs of peaks to reveal fragments that yield the best differentiation. If $R_{isomer} = 1$ then the two MS^n spectra are indistinguishable and the species are likely not isomers. If $R_{isomer} > 1$, a larger number indicates a higher probability that the two species are isomers. The relative intensity of the peaks used for the R_{isomer} score must be higher than 5% of the base peak in at least one of the spectra. To confidently identify each of these isomers by MS^n, we use a threshold that was determined by performing a *t*-Test on the R_{isomer} values

obtained by performing CID and RDD on a mixture of synthetic peptides separated by LCMS. Using 99% confidence intervals, the R_{isomer} threshold for CID is >1.9 and for RDD it is >2.4 (Tao and Julian, 2014).

3.2.8 Calibration Curve for TVLDSGISEVR L-Asp/D-isoAsp Coelution

A calibration curve generated from the L-Asp and D-isoAsp standards was needed to quantify the relative abundance of these coeluting isomers from TVLDSGISEVR. Mixtures containing 0%, 25%, 50%, 75% and 100% of D-isoAsp were prepared and analyzed using the same LC-MS parameters for the digests. The y-axis represents the difference over the sum of the two peaks that had the largest differences in relative intensities in the fragmentation spectra of these two isomers.

3.2.9 Protein Modeling

UCSF Chimera version 1.9 was used to generate the models used in Figs. 5 and 6 (Pettersen et al., 2004).

3.2.10 E91D Mutation of αA

The crystal structure of αA was obtained from PDB 3L1F, which reports the structure for the bovine protein where the residue in position 91 is a glutamic acid. In the human αA-crystallin, there is as conserved mutation placing an aspartic acid at this position. To show the explicit location of the Asp91-Asp92 motif in αA from human, an E91D mutation was performed in silico. The αA mutant was then minimized using the OPLS3 force field in Maestro MacroModel version 10.5.014 (Schrodinger Inc., Portland, OR, USA)

3.3 Results

3.3.1 Changes in αA-crystallin isomerization profiles in aged human lenses

Figure 3.2a illustrates LC chromatograms for Ac-MDVTIQHPWFK, the acetylated N-terminal peptide obtained by tryptic digestion of αA. 39-, 55- and 72-year-old lenses were sampled, and the results obtained from the water-insoluble (WI) fraction of the nucleus (the oldest proteins in the lens) are shown. To confirm the identities of the isomers, synthetic versions of the peptide containing L-Asp, L-isoAsp, D-Asp and D-isoAsp were synthesized. The fragmentation pattern of the peptides from the digest was compared to the standards, then R_{isomer} scores were used to confirm the exact identity of each isomer (Figures 3.3 and 3.4). The results indicate that the first Asp isomer to elute is L-isoAsp, followed by D-isoAsp, L-Asp and finally D-Asp. The relative amount of L-Asp decreases steadily as a function of age from 78 ± 4% in the 39-year-old lens to 45 ± 3% in the 72-year-old lens. D-isoAsp is the most abundant degradation isomer for all ages, though the proportion relative to L-isoAsp becomes nearly equivalent over time. Another product identified in the chromatograms corresponds to glutamine deamidation, denoted as Q→E. The three deamidated glutamine peaks may represent three different isomers of Glu or a combination of coeluting, isomerized Asp and Glu.

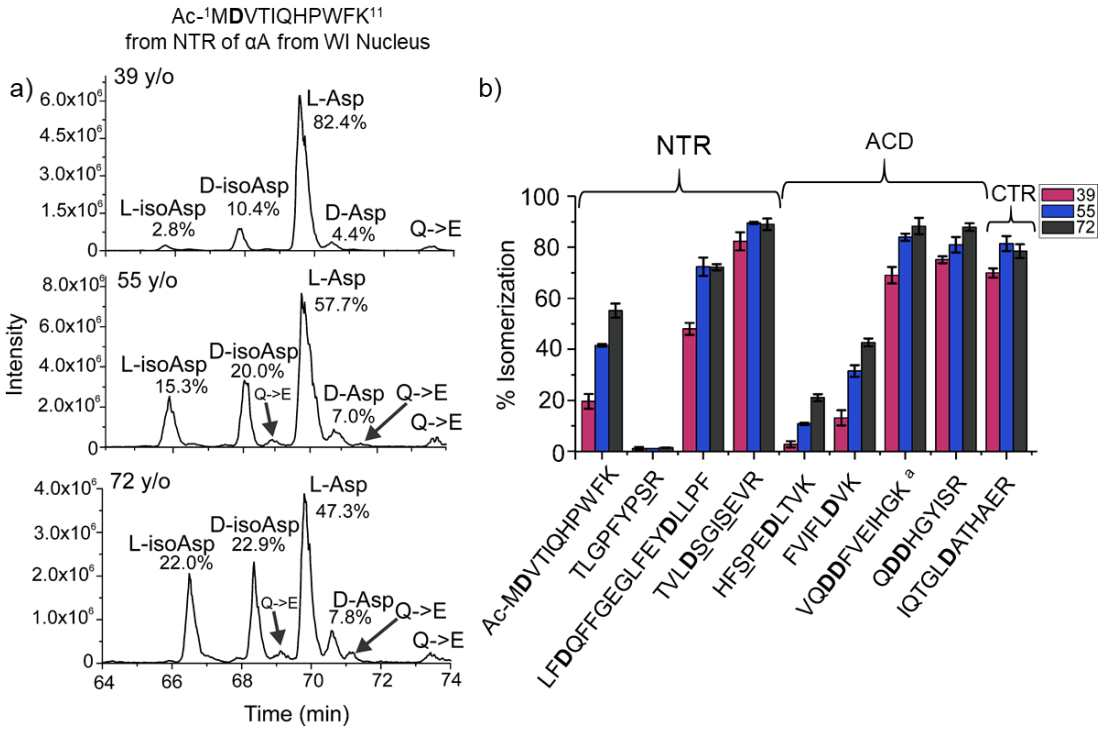
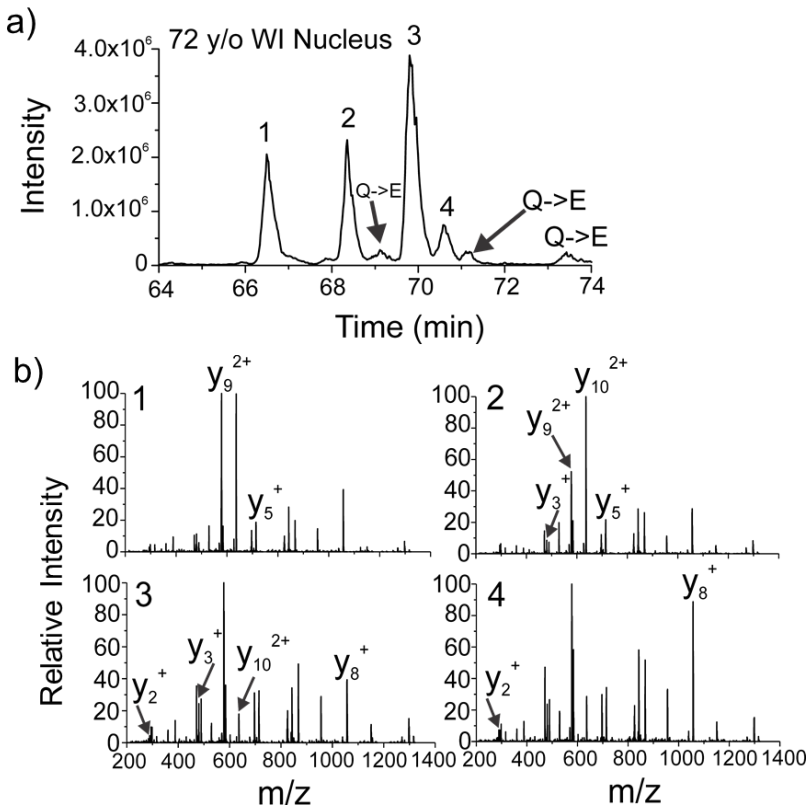


Figure 3.2 Results for the changes in isomerization from the WI nuclei of the 39, 55 and 72-year-old lenses. a) LC chromatograms of Ac-MDVTIQHPWFK, from the WI nucleus of each age. Glutamine deamidation is labeled Q→E. b) Compilation of the percent isomerization per peptide. Aspartic acid and serine residues are bold and underlined, respectively. NTR= N-terminal region, ACD= alpha-crystallin domain and CTR= C-terminal region. [a]Minimum percent isomerization for peptides where L-Asp coelutes with another isomer. Error bars represent standard deviation of technical triplicates.



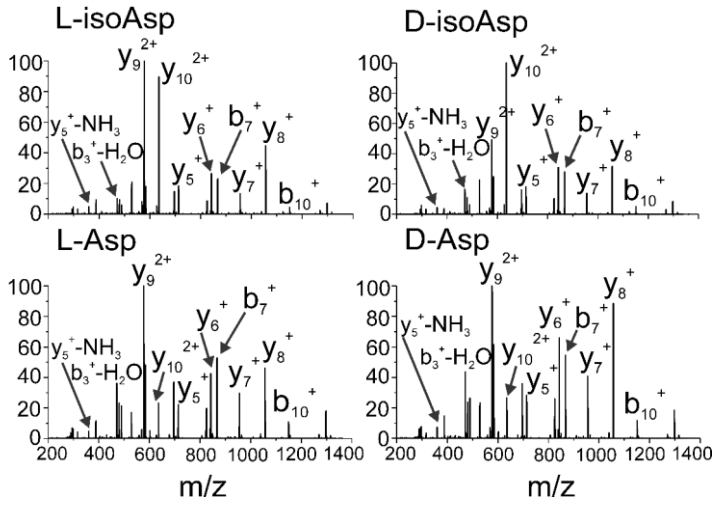
Spectrum 1	Spectrum 2	R_{isomer}	Fragments Compared
Peak 1	Peak 2	3.6	y_9^{2+}, y_5^{+}
Peak 2	Peak 3	18.3	y_{10}^{2+}, y_3^{+}
Peak 3	Peak 4	3.1	y_8^{+}, y_2^{+}

$$R_{isomer} = \frac{R_{Peak\,1}}{R_{Peak\,2}} = \frac{\frac{I_A}{I_B}}{\frac{I_A}{I_B}} = \frac{\frac{100.00}{13.46}}{\frac{44.20}{21.62}} = \frac{7.43}{2.04} = 3.6;$$

where I_A = Relative intensity of y_9^{2+}, I_B = Relative Intensity of y_5^{+}

Figure 3.3 Detection of four isomeric species from Ac-MDVTIQHPWFK from the 72 y/o WI Nucleus. (A) LC chromatogram of the N-terminal peptide in the 72 y/o lens. The numbers on top of the peaks represent the four isomeric spcies based on elution. Q→E indicates glutamine deamidation. (B) Collision-induced dissociation (CID) of the [Ac-MDVTIQHPWFK+2H]$^{2+}$ precursor in each of the peaks. Number of the upper left-hand side of each spectrum signifies the corresponding LC peak. The specific fragments labeled in each represent those that are used for R_{isomer} analysis. For mass spectra, base peak is set to 100% relative intensity. (C) Comparing the y_9^{2+} and y_5^{+} fragment intensities from peak 1 and peak 2 yields an R_{isomer} value of 3.6 (above the threshold of 1.9), indicating that these two peaks are different isomers. The same analysis was performed on peak 2 vs. 3 and peak 3 vs. 4 and the R_{isomer} scores indicate these peaks each contain unique isomers. (D) R_{isomer} calculation for Peak 1 vs Peak 2.

a)



b)

Spectrum 1	Spectrum 2	R_{isomer}	Fragments Compared
Peak 1	**L-isoAsp**	**1.6**	**$[M+2H]^{2+}$ - H_2O, y_3^+**
Peak 2	L-isoAsp	5.2	y_9^{2+}, y_6^+
Peak 3	L-isoAsp	14.2	b_3^+-H_2O, y_{10}^{2+}
Peak 4	L-isoAsp	19.2	b_3^+-H_2O, y_{10}^{2+}

Spectrum 1	Spectrum 2	R_{isomer}	Fragments Compared
Peak 1	L-Asp	10.2	y_{10}^{2+}, b_7^+
Peak 2	L-Asp	11.3	y_{10}^{2+}, b_{10}^+
Peak 3	**L-Asp**	**1.5**	**b_7^+, b_{10}^+**
Peak 4	L-Asp	2.3	y_8^+, b_{10}^+

Spectrum 1	Spectrum 2	R_{isomer}	Fragments Compared
Peak 1	D-isoAsp	3.3	$[M+2H]^{2+}$ - H_2O, y_3^+
Peak 2	**D-isoAsp**	**1.6**	**y_{10}^{2+}, y_3^+**
Peak 3	D-isoAsp	14.3	y_8^+, y_2^+
Peak 4	D-isoAsp	12.4	y_{10}^{2+}, y_7^+

Spectrum 1	Spectrum 2	R_{isomer}	Fragments Compared
Peak 1	D-Asp	11.0	y_{10}^{2+}, b_{10}^+
Peak 2	D-Asp	11.1	y_{10}^{2+}, b_{10}^+
Peak 3	DAsp	3.1	$[M+2H]_2^+$ - H_2O, y_5^+ -NH_3
Peak 4	**D-Asp**	**1.8**	**y_7^+, y_5^+**

Figure 3.4 Identification of the four isomeric species from Ac-MDVTIQHPWFK from the 72 y/o WI Nucleus. (A) CID spectra from the synthetic $[Ac\text{-}MDVTIQHPWFK+2H]^{2+}$ precursor for each of the Asp isomers of this peptide. (B) Comparison of each CID spectrum from the digest to each of the synthetics. To identify the isomeric species, R_{isomer} scores against each synthetic isomer peptide are calculated, and scores that fall below the threshold of 1.9 are allow for confident determination. For example, the CID spectrum for the first peak in the digest in Fig. S1 compared each of the synthetics only yields one R_{isomer} value below the threshold, (1.6 for L-isoAsp), signifying that this peak is the L-isoAsp species. Similar analysis was performed for the other three peaks and it was found that peak 2 was D-isoAsp, peak 3 was L-Asp and peak 4 was D-Asp.

The results from Ac-MDVTIQHPWFK and other αA peptides are summarized in Figure 3.5b by reporting the total percent isomerization. Additionally, the protein regions are labeled according to their respective structural domains. Although the overall trend illustrates isomerization increases with age, there are some regions where isomerization is nearly equivalent for all ages. ^{13}TLGPFYPSR^{21} from the N-terminal region (NTR) undergoes very little isomerization, while ^{146}IQTGLDATHAER^{157} from the C-terminal region (CTR) is already extensively isomerized in the youngest sample.

3.3.2 Changes in αA-crystallin isomerization between the cortex and nucleus

Figure 3.5a shows LC chromatograms from Ac-MDVTIQHPWFK, comparing results obtained from the WS cortex, WI cortex, and WI nucleus of the 72-year-old lens. Although the fraction corresponding to the WS nucleus was also examined, it contained insufficient concentration of this peptide for analysis. The L-Asp isomer is dominant in the WS cortex. D-isomers are prevalent in the WI cortex, where L-isoAsp represents the lowest abundance isomer. For the WI nucleus, L-isoAsp increases dramatically, accounting for nearly 3x the abundance of D-Asp. The results from similar analyses of other peptides are summarized in Figure 3.5b, where the total percent isomerization is reported. In general, a consistent trend is noted where the degree of isomerization increases from WS cortex to WI cortex to WI nucleus. For peptides where data from the WS nucleus was obtained, the degree of isomerization is either comparable to, or slightly below the level observed in the WI nucleus.

Although the total percent isomerization in Figure 3.5b is useful for examining overall trends, detailed information about individual isomers is lost. More comprehensive data is

provided in Figure 3.5c, where the fractional percent of each isomer from the 72-year-old lens is illustrated in a stacked bar chart. The fractional percent isomerization for each isomer was determined by the peak areas from the LC chromatogram, e.g. the data shown in Figure 3.5a. Isomers with confirmed identities are illustrated in gray (L-Asp), pink (L-isoAsp), teal (D-isoAsp), or green (D-Asp). Unidentified isomers are shown in muted pastel colors, and all isomers are ordered by elution times. Co-eluting isomers were quantified, when necessary, using a calibration curve (Figure 3.6). Results for the N-terminal peptide, Ac-^{1}MDVTIQHPWFK^{11}, are shown in the lowest dataset grouping on the y-axis. The fractional percentages match the results shown in Figure 3.5a. For many of the peptides, the L-Asp isomer is most abundant. However, if the peptide contains multiple Asp residues, it is more likely that an isomerized peptide will be most abundant. L-isoAsp is typically found in small quantities in the WS cortex but occupies a greater fraction of the total in the nucleus. Significant amounts of D-isomers are observed, particularly in the insoluble fractions.

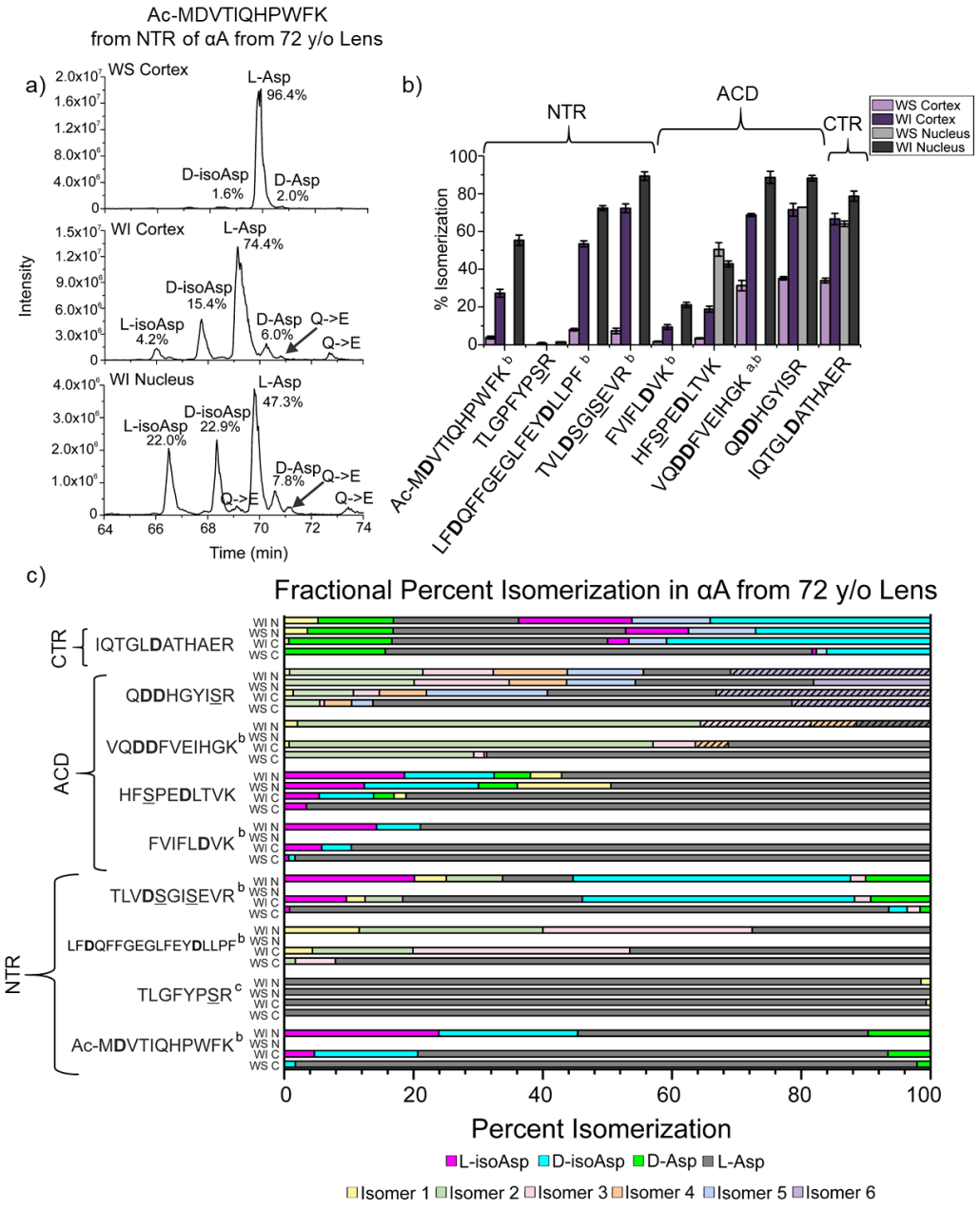


Figure 3.5 Results for αA from the 72-year-old lens. a) LC chromatograms for Ac-MDVTIQHPWFK. b) Compilation of the percent isomerization per peptide. Format is similar to Fig. 2b. [a]Minimum percent isomerization for peptides where L-Asp coelutes with another isomer. [b]Insufficient ion count for analysis of the peptide from the WS nucleus. [c]Peptide containing serine epimers, gray bar represents L-Ser. c) Fractional percent isomerization for each peptide. The first line in the legend contains isomers with known identities, and the second line unknown isomers. Red asterisk indicates serine isomerization. Black diagonal stripes represent co-eluting species.

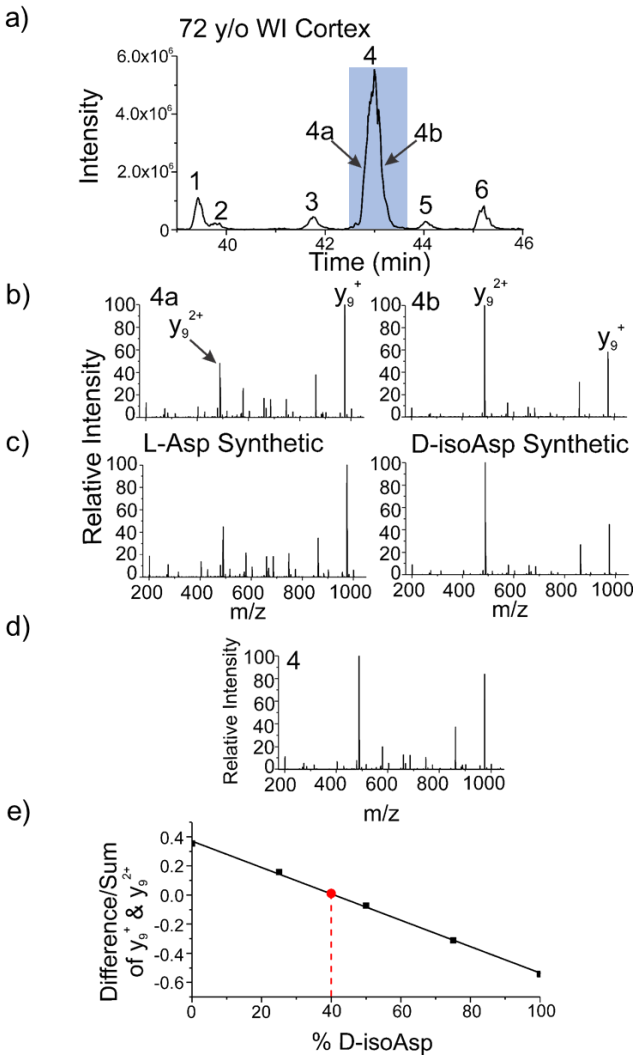


Figure 3.6 Determination of the relative amounts of L-Asp and D-isoAsp from TVLDSGISEVR using a calibration curve. (A) LC chromatogram of TVLDSGISEVR from the 72 y/o WI cortex. Peak 4 is shown in blue and the front end and back end of the peak are labeled 4a and 4b, respectively. (B) CID was performed on the $[TVLDSGISEVR+2H]^{2+}$ precursor from the front end (4a) and back end (4b) to produce an R_{isomer} value of 4.6 indicating that they are unique isomers. (C) Comparison of 4a and 4b to each Asp isomer (L-Asp, L-isoAsp, D-Asp and D-isoAsp) of this peptide indicates that the L-Asp and D-isoAsp isomers are the two coeluting species in peak 4. (D) Averaging the fragmentation abundances of the 33 scans in peak 4 yields a spectrum that contains a mixture of D-isoAsp and L-Asp. (E) Calibration curve generated by taking the difference over the sum of the two peaks that had the largest differences in relative intensities between these two isomers. The average fragmentation abundance for peak 4 was then used to determine that this peak contains 40.1% D-isoAsp with the remainder (59.9%) being L-Asp.

3.3.3 In vitro versus in vivo isomerization of a model peptide from αA-Crystallin

The *in vivo* isomerization products for ^{146}IQTGLDATHAER^{157} as determined by LC are shown in Figures 3.7a-d. L-isoAsp represents a very modest fraction of the total in both the WS and WI cortex fractions. In contrast, L-isoAsp is abundant in the nucleus, shifting from the least abundant isomer identified in the WI cortex to the second largest peak in the WI nucleus. Although the amount of D-isomers also increases in the nucleus, the rise of L-isoAsp is disproportionately large. The intrinsic isomerization tendencies of IQTGLDATHAER were evaluated by incubating synthetic L-Asp peptide at 37ºC in Tris-HCl, pH 7.4. After 43 days, the L-Asp form still accounted for >80% of the peptide, with L-isoAsp representing the largest degradation product (see Figure 3.7e). It is known that formation of the crucial succinimide intermediate is 13-35 times faster by replacement of Asp for Asn (Stephenson and Clarke, 1989). This substitution also allows us to determine the propensity for ring-opening to the L-Asp isomer. After 22 days, 49% ±3% of the L-Asn peptide deamidated, yielding the results shown in Figure 3.7f. L-isoAsp is the dominant product, in agreement with previous observations (Riggs et al., 2017), and only ~30% of the peptide returns to the L-Asp form. The deamidated product shown in Fig. 3.7f was then incubated with PIMT and SAM in Tris-HCl pH= 7.4 at 37ºC. Fig. 3.7g shows the chromatogram after 45 minutes of incubation. The full mass spectrum on the bottom shows that the L-isoAsp peak contains three coeluting species: L-isoAsp, L-isoAsp methyl ester and the succinimide ring. Results obtained after 6 hours of incubation are shown in Fig. 4h, where the total relative amount of L-isoAsp has

decreased to 2.5%. The L-Asp product increases from 32% to 77%, which is also accompanied by an increase in the amount of D-Asp and D-isoAsp.

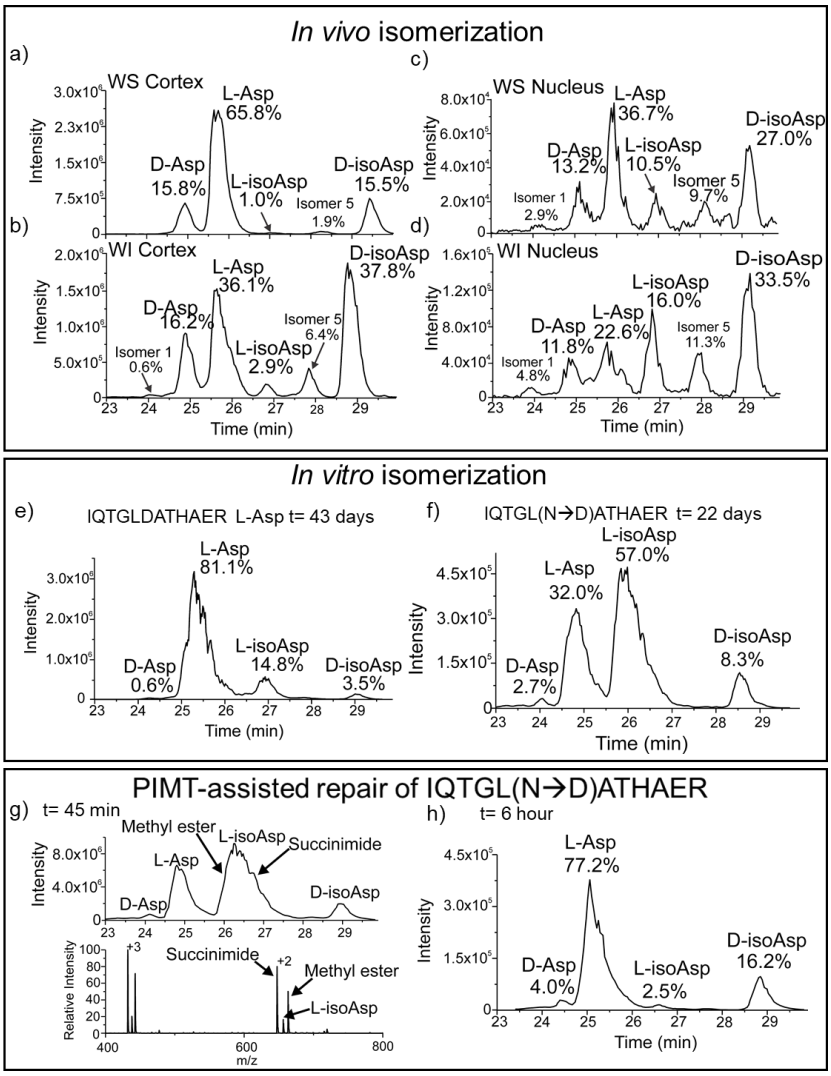


Figure 3.7 Differences between *in vivo* versus *in vitro* IQTGLDATHAER isomerization products. a-d) IQTGLDATHAER chromatograms from the WS cortex, WI cortex, WS nucleus and WI nucleus from the 72 -year-old lens. e) *in vitro* isomerization of the IQTGLDATHAER L-Asp synthetic peptide. f) *in vitro* isomerization products emerging from the deamidation of a synthetic peptide IQTGL(N→D)ATHAER after 22 days. g) Products from the PIMT treatment of IQTGLDATHAER after 45 minutes. Mass spectrum on bottom shows that the succinimide and methyl ester intermediates coelute with L-isoAsp. h) Extracted ion chromatogram of IQTGLDATHAER after 6 hours of incubation with PIMT**.**

3.3.4 Isomerization patterns for structurally similar α-Crystallins

LC chromatograms for ^{90}VQDDFVEIHGK^{100} from αA and ^{93}VLGDVIEVHGK^{103} from αB are shown in Figure 3.8 along with relevant portions of the crystal structures containing these peptides. The native L-Asp isomers are labeled in each spectrum, but the remaining isomers are not identified because it is unfeasible to synthesize all possible standards for peptides with multiple Asp residues. It is also unlikely that all isomers will chromatographically resolve. To verify that the L-Asp peak does not contain multiple isomers, the CID spectra at the front and the back end of the peak were compared using an R_{isomer} score. The R_{isomer} scores are 1.5 (WS cortex), 1.7 (WI cortex) and 2.3 (WI nucleus). The R_{isomer} score for the WI nucleus is above the threshold (>1.9), suggesting this peak contains coeluting isomers. The abundance of the L-Asp isomer decreases from 70.4% in the WS cortex to 31.4% in the WI cortex. Since the L-Asp peak of the WI nucleus contains another isomer that cannot be quantified, the amount of L-Asp cannot be specifically determined, but the entire peak comprises only 9.6% of the total, which represents an upper limit for the amount of L-Asp present. The crystal structure of αA shows that Asp91 and Asp92 are both located within a beta-hairpin (Fig. 3.8b). The higher order structure in αB is identical to αA in this region, with Asp96 of ^{93}VLGDVIEVHGK^{103} also residing in a beta-hairpin (Fig. 3.8c). Despite the structural similarities, the LC chromatograms show that ^{93}VLGDVIEVHGK^{103}, which contains a single Asp residue, is significantly less prone to isomerization (compare Fig. 3.8a and

5d).

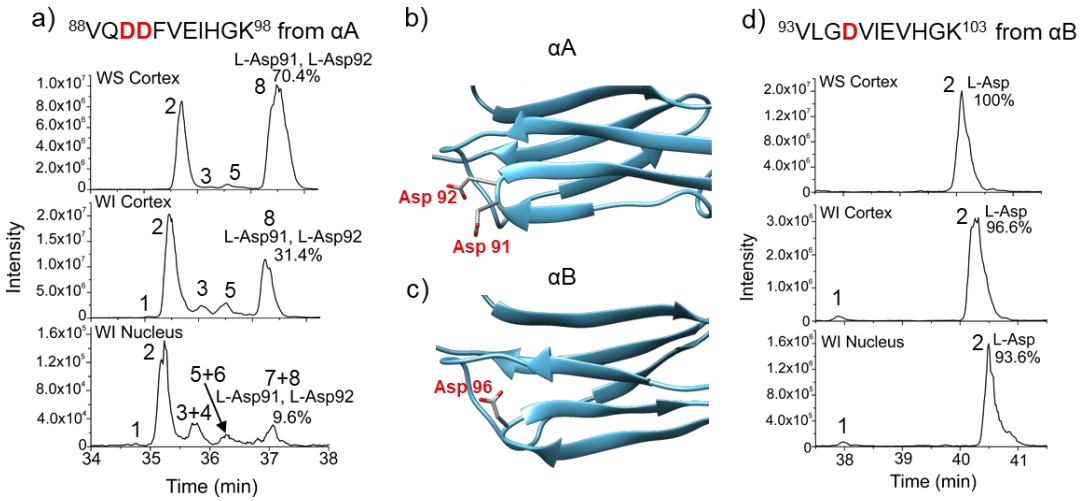


Figure 3.8. Asp-Asp isomerization hotspot in αA Crystallin. a) LC chromatogram of ^{88}VQDDFVEIHGK^{98} from the WS cortex, WI cortex and WI nucleus of the 72 -year-old lens. Peaks are labeled 1-8, where peak #8 was identified as L-Asp91,L-Asp92. b,c) Highlighted region of αA and αB where this motif is present. Crystal structure of bovine αA (PDB ID: 3LIF) (34). Crystal structure of human αB (PDB ID: 3L1G) (34). d) LC chromatogram of ^{93}VLGDVIEVHGK^{103} from the WS cortex, WI cortex and WI nucleus of the 72-year-old lens. Both peaks are labeled, and peak #2 was identified as L-Asp.

Figures 3.9a, 6b, 6f, and 6g compare changes in the total % isomerization between αA and αB for the WI cortex and WI nucleus of a 72-year-old lens. Fig. 3.9c illustrates the sequences of the two proteins, color-coordinated by structural motif, with Asp and Ser residues denoted in black and acidic repeats in red. In Figures 3.9d and 3.9e, representative crystal structures for αA and αB are shown. Both proteins have similar structural motifs, but their primary sequences differ significantly. Human αA and αB share modest 57% sequence homology (Horwitz et al., 1999). Noticeable differences in the primary sequences include deviations in the aspartic acid content and location. In αA, there are 15 aspartic acids compared to 11 in αB. Unlike αA, αB does not contain any aspartic acids in the C-terminal tail and lacks any aspartic acid repeats. The isomerization

profiles also differ substantially between proteins. In general, αB is less isomerized than αA, but isomerized peptides are more likely to remain soluble in αA than αB. For example, there are five isomerized peptides in αB that are only found in the WI fraction, including significantly degraded peptides such as ^{108}QDEHGFISR^{116}. In the WI fractions, αA is more degraded, particularly in the ordered alpha-crystallin domain (ACD).

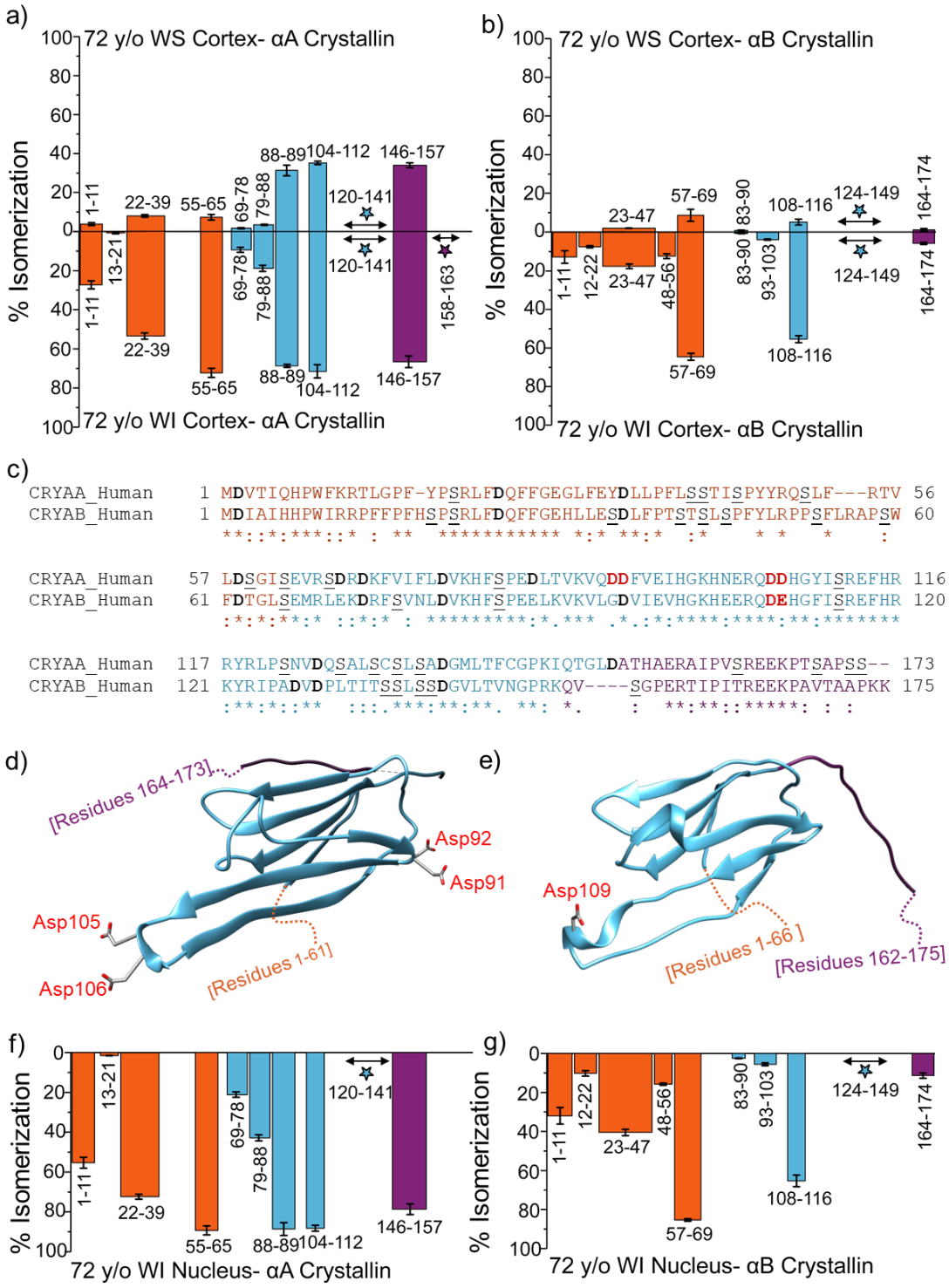

a)
72 y/o WS Cortex- αA Crystallin
% Isomerization
1-11
13-21
22-39
55-65
69-78
79-88
88-89
104-112
120-141
146-157
158-163
72 y/o WI Cortex- αA Crystallin
b)
72 y/o WS Cortex- αB Crystallin
% Isomerization
1-11
12-22
23-47
48-56
57-69
83-90
93-103
108-116
124-149
164-174
72 y/o WI Cortex- αB Crystallin
c)
CRYAA_Human 1 MDVTIQHPWFKRTLGPF-YPSRLFDQFFGEGLFEYDLLPFLSSTISPYYRQSLF---RTV 56
CRYAB_Human 1 MDIAIHHPWIRRPFFPFHSPSRLFDQFFGEHLLESDLFPTSTSLSPFYLRPPSFLRAPSW 60
CRYAA_Human 57 LDSGISEVRSDRDKFVIFLDVKHFSPEDLTVKVQDDFVEIHGKHNERQDDHGYISREFHR 116
CRYAB_Human 61 FDTGLSEMRLEKDRFSVNLDVKHFSPEELKVKVLGDVIEVHGKHEERQDEHGFISREFHR 120
CRYAA_Human 117 RYRLPSNVDQSALSCSLSADGMLTFCGPKIQTGLDATHAERAIPVSREEKPTSAPSS-- 173
CRYAB_Human 121 KYRIPADVDPLTITSSLSSDGVLTVNGPRKQV----SGPERTIPITREEKPAVTAAPKK 175
d)
[Residues 164-173]
Asp92
Asp91
Asp105
Asp106
[Residues 1-61]
e)
Asp109
[Residues 1-66]
[Residues 162-175]
f)
% Isomerization
1-11
13-21
22-39
55-65
69-78
79-88
88-89
104-112
120-141
146-157
72 y/o WI Nucleus- αA Crystallin
g)
% Isomerization
1-11
12-22
23-47
48-56
57-69
83-90
93-103
108-116
124-149
164-174
72 y/o WI Nucleus- αB Crystallin

Figure 3.9. Changes in isomerization between αA and αB in the 72 -year-old lens. a,b) Degree of isomerization in αA and αB from 72 -year-old WS cortex. Each tryptic peptide is numbered according to position in the protein, and colored based on structural motif. Orange= N-terminal region (NTR), blue= alpha-crystallin domain (ACD) and purple= C-terminal region (ACD). Sites, where isomerization was detected but could not be identified, are represented as stars. c) Amino acid sequences of αA and αB color coded based on structural motif. Aspartyl residues are in bold, black text. Serine residues are in underlined, black text. Acidic amino acid repeats are in bold, red text. d) Crystal structure of bovine αA (PDB ID: 3LIF). e) Crystal structure of human αB (PDB ID: 3L1G). f,g) Degree of isomerization in αA and αB from 72 -year-old WI nucleus. Identical formatting to Fig. 6a,b.

3.4 Discussion

In general, our results agree with the well-known trend that epimerization and isomerization increase as a function of protein age (Fujii, 2005; Fujii et al., 2010), but, interestingly, many individual peptides exhibit a weak change in isomerization over time. For example, many of the peptides in Figure 3.2 are approximately 80% isomerized in all samples. Given that the succinimide intermediate always produces some of the native L-isomer, 80% isomerization likely is close to the expected equilibrium value. These highly isomerized peptides are found in both the disordered N-terminal and C-terminal regions and in the structured alpha-crystallin domain, suggesting that tertiary structure alone does not dictate isomerization rate. The most highly modified peptides contain multiple sites prone to isomerization (i.e. multiple Asp or Ser residues), which likely increases the probability for modification by enhancing local structural flexibility and the rate of succinimide formation (for Asp containing peptides). Interestingly, both of the highly degraded peptides observed in the ordered domain contain sequential Asp residues, which have previously been identified as isomerization hotspots (Lyon et al., 2017; Yi et al., 2013; Zhang et al., 2011). This phenomenon is discussed in further detail below. In contrast, peptides lacking sequential Asp residues in the alpha-crystallin domain show steady increases in isomerization as a function of age, suggesting that tertiary structure can influence the rate of isomerization (even if it is not a completely dominant factor).

More consistent trends are found when the degree of isomerization is examined as a function of location within the lens, as illustrated in Figure 3.5. Crystallins are typically water-soluble, but degradation can lead to loss of solubility, which becomes dominant in

the nucleus around 40 years of age (McFall-Ngai et al., 1985). The nucleus is the oldest and most isolated part of the lens, while newly-synthesized proteins are continuously added to the cortex (Augusteyn, 2007). By separating the nucleus and cortex, proteins of different ages can be examined within a single lens from one person, allowing many variables to remain constant. The results in Fig. 3 illustrate a clear trend, with proteins from the WS cortex being least degraded, followed by the WI cortex, WS nucleus (when detectable), and finally the WI nucleus. Not surprisingly, the paucity of WS proteins in the nucleus limits detection of peptides from this fraction.

One of the most striking features of Figure 3.5c is the limited amount of L-isoAsp detected in any peptide from the cortex. Although some increase in abundance is noted between the WS and WI cortical fractions, a more dramatic shift is always observed by comparing the cortex and nucleus. On average, the relative amount of L-isoAsp increases from 6.5% (WI cortex) to 17.5% (WI nucleus). No other isomers exhibit such a reproducible shift in abundance between these fractions. The regional activity of PIMT can explain these results. We detected PIMT in the cortex of the 39 and 55 y/o lenses, with each digest yielding 3-4 unique peptides and confident assignment of the protein. These results agree with McFadden and Clarke, who previously identified PIMT in human lenses (McFadden and Clarke, 1986). Additionally, they found PIMT activity decreased with increasing age, including a reduction in activity in the nucleus compared to the cortex. L-isoAsp is the primary target for PIMT (Griffith et al., 2001), and our results suggest that PIMT efficiently suppresses the accumulation of L-isoAsp in the cortex but is unable to effectively repair either D-isomer, in agreement with previous

studies on substrate affinity (Johnson and Aswad, 1985; Murray and Clarke, 1984). The modest increase in L-isoAsp in the WI cortex suggests that PIMT cannot repair soluble and insoluble proteins with the same efficiency. Greater accumulation of L-isoAsp is also detected as a function of age in the WI nucleus (Figure 3.5a). The relative amount of L-isoAsp in the 39 y/o nucleus was 2.8% but increased nearly 5x in the 55 y/o lens, which was dramatically greater than the increase in the D-isoAsp and D-Asp content. The significant jump in L-isoAsp in the WI nucleus has two important implications, 1) PIMT has lost significant activity in the nucleus for adults over ~40-years-old, and 2) isomerization continues to occur in the nucleus after loss of PIMT activity. If proteins in the nucleus remained static after the loss of PIMT activity, then an increase in the abundance of L-isoAsp would not be expected.

Further support for the role of PIMT is provided in Figure 3.7g,h. Shortly after addition of PIMT and SAM to the peptide mixture shown in Figure 3.7f, preferential modification of L-isoAsp is detected (Figure 3.7g). Both the succinimide and methylated versions of the peptide have different masses and are easily distinguished by MS. After 6 hours of incubation, the relative amount of L-isoAsp decreases from 57% to 2.5% (Figure 3.7h). While the L-Asp product increases from 32% to 77%, also accompanied by increases in D-Asp and D-isoAsp. PIMT does not directly produce L-Asp, it merely methylates L-isoAsp, which facilitates reformation of the succinimide ring. All four isomers are produced upon ring-opening, but the L-isoAsp will be methylated again. Eventually, this cycle drives down the population of L-isoAsp and also increases the abundance of L-Asp and D-isomers. These *in vitro* results are consistent with the isomer

populations observed in the WS fractions from the cortex and suggest that PIMT is responsible for suppression of L-isoAsp in the younger portions of the lens. These results suggest that PIMT serves to prolong the timescale over which crystallin degradation occurs, delaying loss of solubility and subsequent aggregation that may eventually lead to cataract formation.

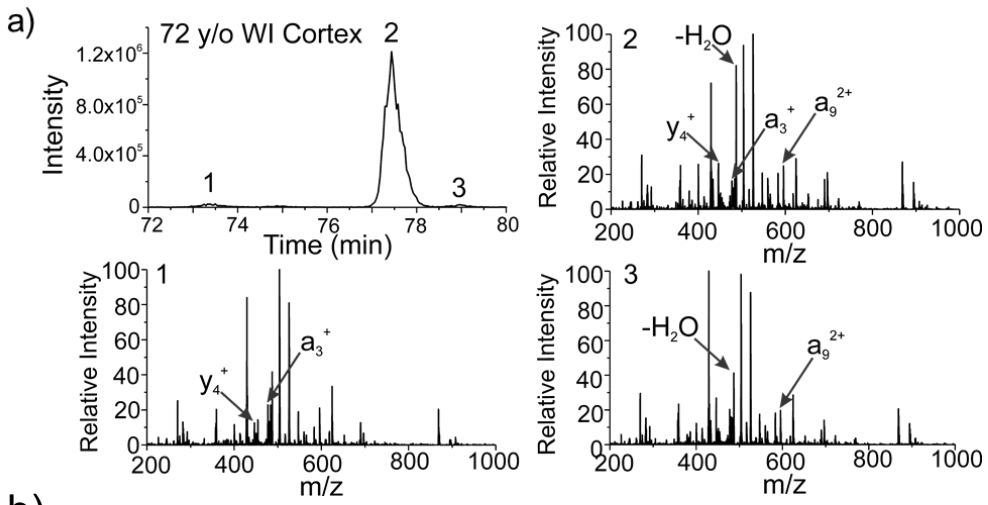
Previous experiments have demonstrated that sequential aspartic acid residues lead to isomerization hotspots (Lyon et al., 2017; Yi et al., 2013; Zhang et al., 2011). It is known that PIMT repair of L-isoAsp is hampered if the n+1, n+2 or n+3 residue is acidic, which would contribute to the hotspot effect (Lowenson and Clarke, 1991). Aspartic acid hotspots may also arise in by deamidation of Asn in the "ND" or "DN" motifs that are found in γS and βA1, respectively (Hooi et al., 2012). However, other contributing factors, such as increased backbone flexibility may also be important but have not been tested. The results in Figure 3.8 reveal additional insight into the nature of isomerization hotspots resulting from sequential Asp residues. Evaluation of peptides from αA and αB that adopt identical secondary structure within the ordered ACD and are comprised of similar sequences reveals a remarkable difference in isomerization. Importantly, ^{88}VQDDFVEIHGK^{98} (from αA) is highly isomerized in all fractions, yet ^{93}VLGDVIEVHGK^{103} (from αB) is significantly less modified. This confirms that the hotspot effect due to sequential Asp residues observed previously in ~1-year-old lenses from animals is also relevant in humans. Furthermore, hotspot isomerization does not result in a simple doubling of normal isomerization since the amount of modification in ^{88}VQDDFVEIHGK^{98} is not twice that in ^{93}VLGDVIEVHGK^{103}. Furthermore, the

presence of Glu99 should inhibit PIMT, suggesting that the hotspot effect is not solely a consequence of reduced enzymatic repair. Given the susceptibility of consecutive Asp residues to age-related isomerization, why are such sequences found in αA?

The answer may be surmised after further comparison of αA and αB as shown in Fig. 6. Although the ordered structures are nearly indistinguishable (Figures 3.9d and 3.9e), the sequences (Fig. 3.9c) vary significantly in Asp content and distribution. As a consequence, the isomerization profiles of αA and αB are strikingly different, compare Fig. 3.9a versus 3.9b and Fig. 3.9f versus 3.9g. Close inspection reveals that αB does not contain any sequential Asp residues and therefore does not contain any Asp-Asp hotspots. The disordered C-terminal domain of αB also lacks Asp, making it less susceptible to isomerization. Overall, isomerization in αB appears to drive protein insolubility to a greater extent than αA, as noted by the paucity of isomerization in the WS cortex (Fig. 3.9b, upper). Collectively, αB is less susceptible to isomerization, begging the question, why is αA needed? Knockout studies in mice revealed normal development without αB but not without αA, which led to premature cataract formation. Furthermore, the chaperone activity of α-crystallin is highest with a ratio of 3:1 αA to αB, suggesting that a combination of the two proteins may be optimal (Srinivas et al., 2010). Although αA is more prone to isomerize, it differs from αB in another important respect, refractive index. Asp is one of the key amino acids required to increase the refractive index of a protein (Glu is not) (Zhao et al., 2011). The increased balance of Asp rather than Glu within αA contributes to its greater refractive index increment (0.1938 αA, 0.1922 αB, d*n*/d*c* (ml/g)). The liability of Asp incorporation into αA in terms of

susceptibility to isomerization over time appears to be outweighed by the need for greater refractive index.

There are also notable differences in the location of serine residues between αA and αB, particularly in the alpha-crystallin domain. ^{11}RPFFPHSPSR^{21} from the N-terminal region of αB is significantly more epimerized than the corresponding peptide in αA. RDD identified three isomers of ^{11}RPFFPHSPSR^{21} as shown in Figure 3.10, indicating that both serine residues are epimerized. ^{48}YLRPPSFLR^{56}, also from the N-terminal region of αB, is 12 ± 1% epimerized in the WI cortex and 15.7% ±0.6% in the WI nucleus. This exceeds the degree of epimerization (2.5% ±0.2%) detected in ^{83}HFSPEELK^{90}, located in the alpha-crystallin domain of αB. In total, serine epimerization is noted for positions 19, 21, 53 and 85 for αB and 20 and 162 for αA. Extensive serine racemization has been detected at both Ser59 and Ser62 in ^{55}TVLDSGISEVR^{65} from αA. By age 40, it was found that ^{55}TVLDSGISEVR^{65} contained 44% D-Ser between these two sites in cataract lenses, compared to only 17% in normal lenses (Hooi et al., 2013). Interestingly, serine epimerization is strongly correlated with loss of solubility, indicating that, while less abundant than aspartic acid isomerization, it may contribute significantly to age-related pathologies associated with the lens.



Spectrum 1	Spectrum 2	R_{isomer}	Fragments Compared
Peak 1	Peak 2	3.0	y_4^+, a_3^+
Peak 2	Peak 3	3.4	$[M+2H]^{2+}$ -H_2O, a_9^{2+}

Figure 3.10 Detection of two sites of serine isomerization in RPFFPHSPSR. (A) LC chromatogram of the 4-iodobenzoyl modified RPFFPHSPSR from the 72 y/o WI Cortex digest. Three isomeric species are present in the chromatogram and are labeled based on elution. The three resulting spectra after performing radical directed dissociation (RDD) on each of the LC peaks are also shown. Comparison of the y_4^+ and a_3^+ fragment intensities in peak 1 vs. peak 2 yield an R_{isomer} value of 3.0 (above the threshold of 2.4 for RDD) indicating that these two peaks contain unique isomers. Similar analysis for peak 2 vs. 3 also yields an R_{isomer} score of 3.4, indicating that these two peaks are unique isomers.

3.5 Conclusions

While it is known that long-lived proteins fall victim to various spontaneous degradations as they age in the body, it has been difficult to monitor subtle, yet prolific, modifications such as isomerization and epimerization. By tracking differences between the newly synthesized proteins in the cortex of the lens and the oldest, most highly degraded proteins in the nucleus, we establish the time-course for protein isomerization

as a function of influencing factors such as local sequence, tertiary structure, and solubility. Peptides from the cortical fractions were found to lack L-isoAsp, which accumulated significantly in the water-insoluble fractions of the nucleus. PIMT is more prevalent in the cortex than the nucleus, and *in vitro* studies confirm that PIMT activity can account for the observed absence of L-isoAsp. Consequently, isomerization occurs in the cortex but it is actively repaired by PIMT, whereas PIMT activity declines in the nucleus, although isomerization and protein aging continue. Isomerization tendencies can also differ between highly similar proteins, such as αA and αB crystallin. Although αB appears to be less prone to isomerization, αA excels at sustaining damage while remaining soluble and makes a greater contribution towards refractive index. These differences in age-induced isomerization and function may explain why both αA and αB are needed in the lens. In many cases, isomerization appears to facilitate protein aggregation and loss of solubility. Given the long list of age-related diseases where protein aggregation is observed, including cataracts, Alzheimer's, Parkinson's, myofibrillar myopathy, and many others, it is likely that many of the observations detailed herein will extend to other systems beyond the lens and may offer clues for understanding their pathogenesis.

Augusteyn, R.C., 2007. Growth of the human eye lens. Mol. Vis. 13, 252–7.
Chavous, D.A., Jackson, F.R., O'Connor, C.M., 2001. Extension of the Drosophila lifespan by overexpression of a protein repair methyltransferase. Proc. Natl. Acad. Sci. 98, 14814–14818. https://doi.org/10.1073/pnas.251446498
Fujii, N., 2005. D-amino acid in elderly tissues. Biol. Pharm. Bull. 28, 1585–1589. https://doi.org/10.1248/bpb.28.1585
Fujii, N., Kaji, Y., Fujii, N., Nakamura, T., Motoie, R., Mori, Y., Kinouchi, T., 2010. Collapse of homochirality of amino acids in proteins from various tissues during aging. Chem. Biodivers. https://doi.org/10.1002/cbdv.200900337

Fujii, N., Sakaue, H., Sasaki, H., Fujii, N., 2012. A rapid, comprehensive liquid chromatography-mass spectrometry (LC-MS)-based survey of the Asp isomers in crystallins from human cataract lenses. J. Biol. Chem. 287, 39992–40002. https://doi.org/10.1074/jbc.M112.399972
Fujii, N., Takata, T., Fujii, N., Aki, K., 2016. Isomerization of aspartyl residues in crystallins and its influence upon cataract. Biochim. Biophys. Acta - Gen. Subj. 1860, 183–191. https://doi.org/10.1016/j.bbagen.2015.08.001
Griffith, S.C., Sawaya, M.R., Boutz, D.R., Thapar, N., Katz, J.E., Clarke, S., Yeates, T.O., 2001. Crystal structure of a protein repair methyltransferase from Pyrococcus furiosus with its L-isoaspartyl peptide substrate. J. Mol. Biol. 313, 1103–16. https://doi.org/10.1006/jmbi.2001.5095
Groenen, P.J.T.A., Merck, K.B., De Jong, W.W., Bloemendal, H., 1994. Structure and Modifications of the Junior Chaperone α-Crystallin: From Lens Transparency to Molecular Pathology. Eur. J. Biochem. https://doi.org/10.1111/j.1432-1033.1994.00001.x
Hood, C.A., Fuentes, G., Patel, H., Page, K., Menakuru, M., Park, J.H., 2008. Fast conventional Fmoc solid-phase peptide synthesis with HCTU. J. Pept. Sci. 14, 97–101. https://doi.org/10.1002/psc.921
Hooi, M.Y.S., Raftery, M.J., Truscott, R.J.W., 2013. Age-dependent racemization of serine residues in a human chaperone protein. Protein Sci. 22, 93–100. https://doi.org/10.1002/pro.2191
Hooi, M.Y.S., Raftery, M.J., Truscott, R.J.W., 2012. Racemization of two proteins over our lifespan: Deamidation of asparagine 76 in γS crystallin is greater in cataract than in normal lenses across the age range. Investig. Ophthalmol. Vis. Sci. 53, 3554–3561. https://doi.org/10.1167/iovs.11-9085
Horwitz, J., Bova, M.P., Ding, L.L., Haley, D. a, Stewart, P.L., 1999. Lens alpha-crystallin: function and structure. Eye (Lond). 13 (Pt 3b, 403–408. https://doi.org/10.1038/eye.1999.114
Horwitz, J., Huang, Q.L., Ding, L., Bova, M.P., 1998. Lens ?-Crystallin: Chaperone-like properties. Methods Enzymol. 290, 365–383. https://doi.org/10.1016/S0076-6879(98)90032-5
Jia, C., Lietz, C.B., Yu, Q., Li, L., 2014. Site-specific characterization of d-amino acid containing peptide epimers by ion mobility spectrometry. Anal. Chem. 86, 2972–2981. https://doi.org/10.1021/ac4033824
Johnson, B.A., Aswad, D.W., 1985. Enzymatic Protein Carboxyl Methylation at Physiological pH: Cyclic Imide Formation Explains Rapid Methyl Turnover. Biochemistry 24, 2581–2586. https://doi.org/10.1021/bi00331a028
Lowenson, J.D., Clarke, S., 1992. Recognition of D-aspartyl residues in polypeptides by the erythrocyte L- isoaspartyl/D-aspartyl protein methyltransferase. Implications for the repair hypothesis. J. Biol. Chem. 267, 5985–5995.
Lowenson, J.D., Clarke, S., 1991. Structural elements affecting the recognition of L-isoaspartyl residues by the L-isoaspartyl/D-aspartyl protein methyltransferase. Implications for the repair hypothesis*. J. Biol. Chem. 266, 19396–19406.
Lund, A.L., Smith, J.B., Smith, D.L., 1996. Modifications of the water-insoluble human lens alpha-crystallins. Exp. Eye Res. 63, 661–72. https://doi.org/10.1006/exer.1996.0160

Ly, T., Zhang, X., Sun, Q., Moore, B., Tao, Y., Julian, R.R., 2011. Rapid, quantitative, and site specific synthesis of biomolecular radicals from a simple photocaged precursor. Chem. Commun. 47, 2835. https://doi.org/10.1039/c0cc03363d
Lyon, Y.A., Sabbah, G.M., Julian, R.R., 2017. Identification of Sequence Similarities among Isomerization Hotspots in Crystallin Proteins. J. Proteome Res. acs.jproteome.7b00073. https://doi.org/10.1021/acs.jproteome.7b00073
Ma, Z., Hanson, S.R., Lampi, K.J., David, L.L., Smith, D.L., Smith, J.B., 1998. Age-related changes in human lens crystallins identified by HPLC and mass spectrometry. Exp. Eye Res. 67, 21–30. https://doi.org/10.1006/exer.1998.0482
Masters, P.M., Bada, J.L., Zigler, J.S., 1978. Aspartic acid racemization in heavy molecular weight crystallins and water insoluble protein from normal human lenses and cataracts. Proc. Natl. Acad. Sci. U. S. A. 75, 1204–8.
McFadden, P.N., Clarke, S., 1987. Conversion of isoaspartyl peptides to normal peptides: implications for the cellular repair of damaged proteins. Proc. Natl. Acad. Sci. U. S. A. 84, 2595–2599. https://doi.org/10.1073/pnas.84.9.2595
McFadden, P.N., Clarke, S., 1986. Protein carboxyl methyltransferase and methyl acceptor proteins in aging and cataractous tissue of the human eye lens. Mech. Ageing Dev. 34, 91–105. https://doi.org/10.1016/0047-6374(86)90107-7
McFall-Ngai, M.J., Ding, L.-L., Takemoto, L.J., Horwitz, J., 1985. Spatial and temporal mapping of the age-related changes in human lens crystallins. Exp. Eye Res. 41, 745–758. https://doi.org/10.1016/0014-4835(85)90183-6
Murray, E.D., Clarke, S., 1984. Synthetic peptide substrates for the erythrocyte protein carboxyl methyltransferase. Detection of a new site of methylation at isomerized L-aspartyl residues. J. Biol. Chem. 259, 10722–10732.
Noguchi, S., 2010. Structural changes induced by the deamidation and isomerization of asparagine revealed by the crystal Structure of Ustilago sphaerogena ribonuclease U2B. Biopolymers 93, 1003–1010. https://doi.org/10.1002/bip.21514
O'Connor, P.B., Cournoyer, J.J., Pitteri, S.J., Chrisman, P.A., McLuckey, S.A., 2006. Differentiation of aspartic and isoaspartic acids using electron transfer dissociation. J. Am. Soc. Mass Spectrom. 17, 15–19. https://doi.org/10.1016/j.jasms.2005.08.019
Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera—A Visualization System for Exploratory Research and Analysis. J Comput Chem 25, 1605–1612. https://doi.org/10.1002/jcc.20084
Qin, Z., Dimitrijevic, A., Aswad, D.W., 2015. Accelerated protein damage in brains of PIMT+/- mice; a possible model for the variability of cognitive decline in human aging. Neurobiol. Aging 36, 1029–1036. https://doi.org/10.1016/j.neurobiolaging.2014.10.036
Rao, P.V., Huang, Q. ling, Horwitz, J., Zigler, J.S., 1995. Evidence that α-crystallin prevents non-specific protein aggregation in the intact eye lens. BBA - Gen. Subj. 1245, 439–447. https://doi.org/10.1016/0304-4165(95)00125-5
Riggs, D.L., Gomez, S. V., Julian, R.R., 2017. Sequence and Solution Effects on the Prevalence of d -Isomers Produced by Deamidation. ACS Chem. Biol. acschembio.7b00686. https://doi.org/10.1021/acschembio.7b00686
Samuel Zigler, J., Goosey, J., 1981. Aging of protein molecules: lens crystallins as a model system. Trends Biochem. Sci. https://doi.org/10.1016/0968-0004(81)90050-5

Srinivas, P., Narahari, A., Petrash, J.M., Swamy, M.J., 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. https://doi.org/10.1002/iub.373
Stephenson, R.C., Clarke, S., 1989. Succinimide formation from aspartyl and asparaginyl peptides as a model for the spontaneous degradation of proteins. J. Biol. Chem. 264, 6164–6170.
Tao, W.A., Zhang, D., Nikolaev, E.N., Cooks, R.G., 2000. Copper(II)-assisted enantiomeric analysis of D,L-amino acids using the kinetic method: Chiral recognition and quantification in the gas phase. J. Am. Chem. Soc. 122, 10598–10609. https://doi.org/10.1021/ja000127o
Tao, Y., Julian, R.R., 2014. Identification of amino acid epimerization and isomerization in crystallin proteins by tandem LC-MS. Anal. Chem. 86, 9733–9741. https://doi.org/10.1021/ac502296c
Tao, Y., Quebbemann, N.R., Julian, R.R., 2012. Discriminating d-amino acid-containing peptide epimers by radical-directed dissociation mass spectrometry. Anal. Chem. 84, 6814–6820. https://doi.org/10.1021/ac3013434
Truscott, R.J.W., Friedrich, M.G., 2016. The etiology of human age-related cataract. Proteins don't last forever. Biochim. Biophys. Acta. https://doi.org/10.1016/j.bbagen.2015.08.016
Yamamoto, a, Takagi, H., Kitamura, D., Tatsuoka, H., Nakano, H., Kawano, H., Kuroyanagi, H., Yahagi, Y., Kobayashi, S., Koizumi, K., Sakai, T., Saito, K., Chiba, T., Kawamura, K., Suzuki, K., Watanabe, T., Mori, H., Shirasawa, T., 1998. Deficiency in protein L-isoaspartyl methyltransferase results in a fatal progressive epilepsy. J. Neurosci. 18, 2063–74.
Yi, L., Beckley, N., Gikanga, B., Zhang, J., Wang, Y.J., Chih, H.W., Sharma, V.K., 2013. Isomerization of Asp-Asp motif in model peptides and a monoclonal antibody fab fragment. J. Pharm. Sci. 102, 947–959. https://doi.org/10.1002/jps.23423
Zhang, J., Yip, H., Katta, V., 2011. Identification of isomerization and racemization of aspartate in the Asp-Asp motifs of a therapeutic protein. Anal. Biochem. 410, 234–243. https://doi.org/10.1016/j.ab.2010.11.040
Zhao, H., Brown, P.H., Schuck, P., 2011. On the distribution of protein refractive index increments. Biophys. J. 100, 2309–2317. https://doi.org/10.1016/j.bpj.2011.03.004
Zheng, X., Deng, L., Baker, E.S., Ibrahim, Y.M., Petyuk, V.A., Smith, R.D., Govind, N., Ibrahim, Y.M., Kabanda, M.M., Dubery, I.A., Heyman, H.M., Smith, R.D., Madala, N.E., Baker, E.S., McLean, J.A., 2017. Distinguishing d - and l -aspartic and isoaspartic acids in amyloid -β peptides with ultrahigh resolution ion mobility spectrometry. Chem. Commun. 8, 1381–1388. https://doi.org/10.1039/C7CC03321D

Chapter 4

Structural and Functional Consequences of Age-Related Isomerization in α Crystallins

4.1 Introduction

Long-lived proteins are important but often underappreciated, with recent findings illustrating their pervasiveness within critical organs and suggesting that their chemistry and biology should not be ignored.[1] Longevity renders proteins susceptible to degradation and accumulation of post-translational modifications (PTMs). Among these modifications are non-enzymatic, spontaneous changes including truncation, cross-linking, oxidation, deamidation, isomerization, and epimerization.[1,2] Aspartic acid residues are most prone to isomerization,[3,4] readily forming a succinimide ring following attack of the sidechain by the peptide backbone. The succinimide is susceptible to racemization and can reopen in two ways, ultimately yielding four isomers: L-Asp, D-Asp, L-isoAsp, and D-isoAsp.[5] After deamidation, asparagine can also yield four isomers of aspartic acid, but this process involves chemical modification in addition to isomerization. Aspartic acid isomerization is one of the most prevalent degradation pathways of crystallin proteins in aged human lenses and is known to increase with age and influence water-solubility and oligomerization in general.[6-9] Serine is also frequently found to undergo epimerization in long-lived proteins, readily forming D-Ser.[4,10] Herein, we use isomerization to describe the conversion of Asp residues and epimerization for the stereoinversion of Ser. Modification of both residues is difficult to detect because they do not lead to a change in mass and are consequently invisible to mass spectrometry (MS)-

based methods typically employed during proteomic analyses.[11,12] As a result, they are not widely studied, frequently overlooked, and their consequences are largely unknown.[13]

The crystallin proteins of the eye lens, in which there is no protein turnover, are among the longest-lived proteins in the body.[14] They are ideal targets for studying spontaneous degradation pathways that occur due to aging. The most abundant crystallins in humans, αA and αB, are important molecular chaperone[15] and regulatory proteins.[16] While αA is localized almost exclusively to the eye lens,[17] αB is found throughout the body.[18,19] Crystallin malfunction due to mutation or accumulation of PTMs is associated with a variety of diseases including cataract, cardiomyopathies, motor neuropathies, and neurodegeneration.[20-22]

αA and αB are members of the small heat-shock protein (sHSP) family,[15] with structures characterized by a highly conserved α-crystallin domain,[23] flanked in both proteins by less ordered N- and C-terminal regions (Fig. 1A). Outside of the α-crystallin domain, αA and αB share modest sequence homology,[17] and in both cases structural details have proven scarce due to their flexibility and inherent disorder. αA and αB use a number of interfaces to self- and co-assemble into large, polydisperse and dynamic oligomers. The α-crystallin domain mediates dimerization, and the dimers associate into oligomers via an interface between a palindromic sequence in the C-terminal region of one dimer and the α-crystallin domain of another, as well as interactions involving N-terminal regions (Fig 1B).[24-26] Perturbation of the dimer interface, as occurs in the well-known R120G variant of αB,[27] or mutations in the terminal regions, can lead to protein

aggregation and malfunction.[28] The function and localization of αB is also partially regulated by phosphorylation of three serines in the N-terminal region, with dysregulation leading to disease.[29-33] These data are consistent with modification of interfacial residues leading to aberrant assembly and compromised function in the cell.

We recently reported 81 sites of isomerization in human αA and αB isolated from the eye lenses of aged donors[8] using high-throughput radical-directed dissociation mass spectrometry (RDD-MS) experiments.[3] Here we examine the influence of isomerization at four sites in αA and αB ($^{\alpha B}$Ser59, $^{\alpha A}$Ser162, $^{\alpha B}$Asp62, $^{\alpha B}$Asp109). These sites reside in regions critical for oligomerization, and our experiments demonstrate how key structural interactions or functionality are disrupted by isomerization. Native MS[34,35] reveals that the interface between the C-terminal region and the α-crystallin domain is dramatically weakened upon conversion of L- to D-Ser. Epimerization abrogates serine phosphorylation, with isomerization of neighboring residues also interfering with kinase recognition. Molecular dynamics (MD) simulations reveal that isomerization of an aspartic acid at the dimer interface leads to cleavage of a salt-bridge, and precipitates the aberrant oligomerization observed in native MS data. Our results demonstrate that age-related, isobaric modifications of individual amino-acid residues can have significant impacts on the quaternary structure and function of α-crystallins, and suggest that similar "invisible" PTMs in other long-lived proteins may have an important influence on age-related diseases.

4.2 Results and discussion

4.2.1 αA and αB accumulate isobaric PTMs at their interfaces

As discussed above, the ordered structural domains of αA and αB crystallin are very similar. Therefore, to identify the relevant structural regions and sites of interest for both proteins, the α-crystallin domain of αB is used as a representative model in Fig. 4.1A. In the oligomeric assembly, monomers first come together to form dimers. As seen in Fig. 4.1A (left), the dimer interface is defined by the same sequence region from two different monomers interacting in an anti-parallel beta-sheet interface. Hydrogen bonding along this interface and complementary Arg120-Asp109 salt bridges stabilize the dimer. Another crucial interaction occurs in a groove where the normally disordered C-terminal tail can bind, Fig. 4.1A (right). This interface is one of several that facilitates the assembly of dimers into larger oligomers. The disordered N-terminal tail also forms quaternary interactions favoring oligomerization in the general region shown in Fig. 4.1A (left), though the interaction partners are omitted for clarity. Each of these regions are critical for proper crystallin assembly and function, suggesting that structural perturbations of these interfaces could have undesirable consequences.

Previous analysis[8] identified (without detailed study) four sites of isomerization within these regions. Given the potential structural impact, we now examine these sites in detail. Analysis of isomerization and epimerization first requires proteolysis. Tryptic digestion of the crystallins yields all four sites (underlined) in three peptides: αA-^{158}AIPV<u>S</u>R^{163}, which lies in the C-terminal region; αB-^{57}AP<u>S</u>WF<u>D</u>TGLSEMR^{69}, which encompasses residues involved in an N-terminal interface; and αB-^{108}Q<u>D</u>EHGFISR^{116} which lies in the

α-crystallin domain at the dimer interface. Each of these peptides corresponds to the relevant pink region in Fig. 4.1A. The degree of isomerization or epimerization at each site was determined with RDD-MS, as described in detail previously,[3,8] and the results are shown in Figs.4.1B-E. This data was obtained from the nucleus and cortex of a 72-year-old human lens, although the results are similar to those from younger donor lenses (data not shown). The complete crystallin sequences, including all known isomerized and epimerized residues, are listed in the Fig. S1 along with additional details about isomer identification, Fig. 4.2.

Performing RDD-MS across the chromatographic peak for αA-^{158}AIPVSR^{163} reveals product ions at 350.25 m/z and 364.00 m/z, whose ratio varies with elution time for the water-insoluble (WI) fraction, as illustrated by the snapshots in Fig. 4.1B where the L-Ser spectra are black and D-Ser is red. The RDD spectra for the water-soluble (WS) fraction do not vary as a function of retention time, and the mass spectra collected at the front and back end of the peak both match L-Ser. The changing ratios for the WI fraction are meaningful because in RDD experiments,[36] a radical is photolytically created at the same, atomically precise location in both isomers. Subsequent collisional activation stimulates migration of the radical and fragmentation of the peptide. Differences in three-dimensional structure due to isomerization lead to differences in the abundance of certain fragmentation channels. Therefore, the varying product ion ratios for the WI fraction in Fig. 4.1B reveal epimerization at $^{\alpha A}$Ser162, even though this modification is not readily apparent from the chromatography. Further analysis enables quantification of the abundance of $^{\alpha A}$D-Ser162 at 8% (see Fig. 4.3 and Table 4.1 for details).[37] This approach

can also be applied to more complex systems, as illustrated by the chromatograms for separation of αB-^{57}APSWFDTGLSEMR^{69}, from which nine different isomers were identified in the WI fraction (Fig. 4.1C bottom). Synthetic standards of selected candidates were used to determine that the original all-L peptide comprises only 14% of the total abundance in the WI fraction (peak 7, which dominates the WS fraction). The $^{\alpha B}$D-Ser59 epimer is present in 5.6% abundance (peak 4) in the WI fraction and 1.4% in the WS fraction (see Fig. 4.4, Table 4.2, and Table 4.3 for more information). Interestingly, when αB-^{57}APSWFDTGLSEMR^{69} is phosphorylated at $^{\alpha B}$Ser59, only a single isomer is detected in the WS fraction (Fig. 1D) and a small number of isomers are detected in the WI fraction. More discussion on the relevance of phosphorylation at this site follows below. The WS and WI fractions of αB-^{108}QDEHGFISR^{116} both contain multiple isomers of $^{\alpha B}$Asp109 (Fig. 4.1E), but the relative abundance of $^{\alpha B}$L-isoAsp109 increases from 5.1% to 55.4% in the WI portion, suggestive that this structural perturbation influences aggregation propensity (see Figure 4.5 and Table 4.4 for details). Upon detailed examination, it is clear that significant amounts of isomerization and epimerization are occurring in regions known to influence the structure and assembly of the α-crystallins, but further investigation is required to establish whether these modifications exert structural influence.

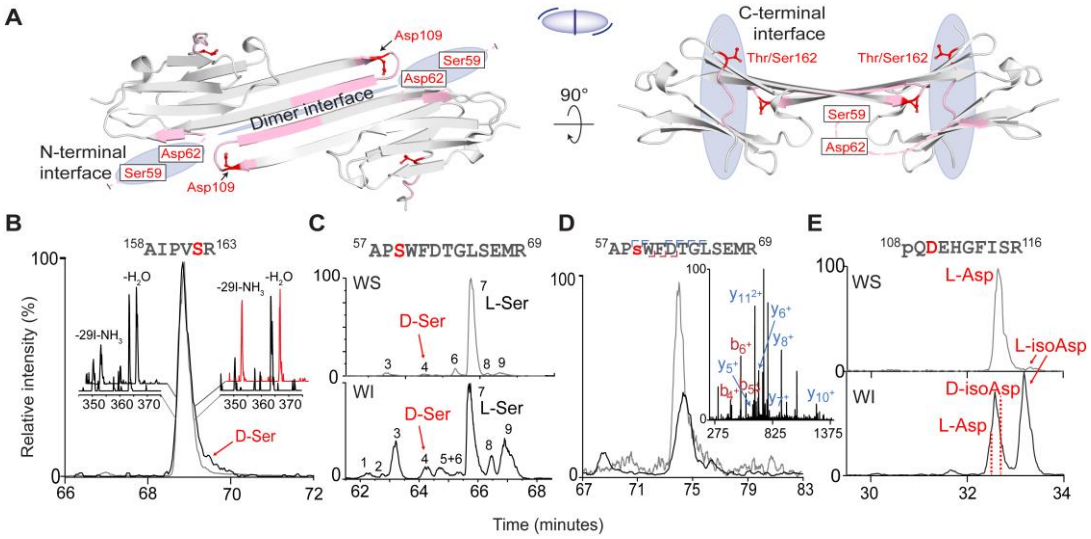


Figure 4.1 (A) Two views of the partial crystal structure of α-crystallin (αB, PDB: 4M5S). The structure of αB is used for illustration purposes since αB and αA intermix freely and share high structural similarity. Blue shaded regions indicate crucial oligomeric interfaces. Pink ribbons denote the isomer-containing peptides, with specific isomerization sites labeled in red. The small cartoon in the middle represents the assembly, with each half ellipse representing a monomer, the central line indicating the dimer interface, and the peripheral lines representing bound C-terminal peptides. Extracted ion chromatograms: (B) $^{\alpha A}$AIPVSR from the WI (black trace) and WS (grey trace) fractions of the cortex. Insets, RDD mass spectra from the front and back of each peak. (C) αB- ^{57}APSWFDTGLSEMR^{69} from the WI/WS nucleus, revealing abundant isomerization in the WI fraction including D-Ser. (D) Phosphorylated ^{57}APsWFDTGLSEMR^{69} detected in the WS cortex (grey trace) and WI cortex (black trace) revealing far less isomerization, where s = phosphoserine. Inset, MS/MS pinpoints the site of phosphorylation to Ser59. (E) αB- ^{108}pQDEHGFISR^{116} from the cortex (pQ=pyroglutamate). In the WS fraction, the major peak contains L-Asp and the minor peak is L-isoAsp. In the WI fraction, the peak at ~33 minutes contains both L-Asp and D-isoAsp; and the final peak is L-isoAsp.

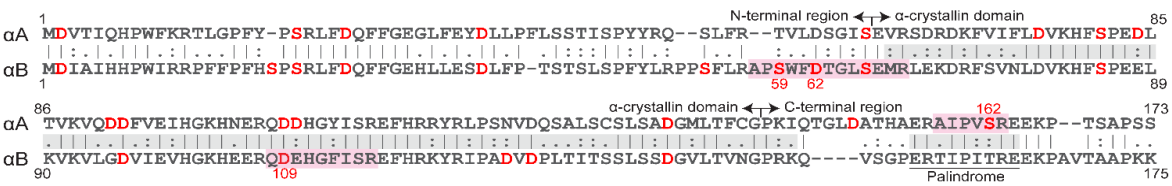


Figure 4.2 Sequence alignment of human αA and αB with designated boundaries between domains. Red residues indicate Asp/Ser sites within peptides from 72-year-old human lens identified in numerous isomeric forms by RDD-MS. Pink regions indicate isomer-containing peptides mapping to oligomeric interfacial regions

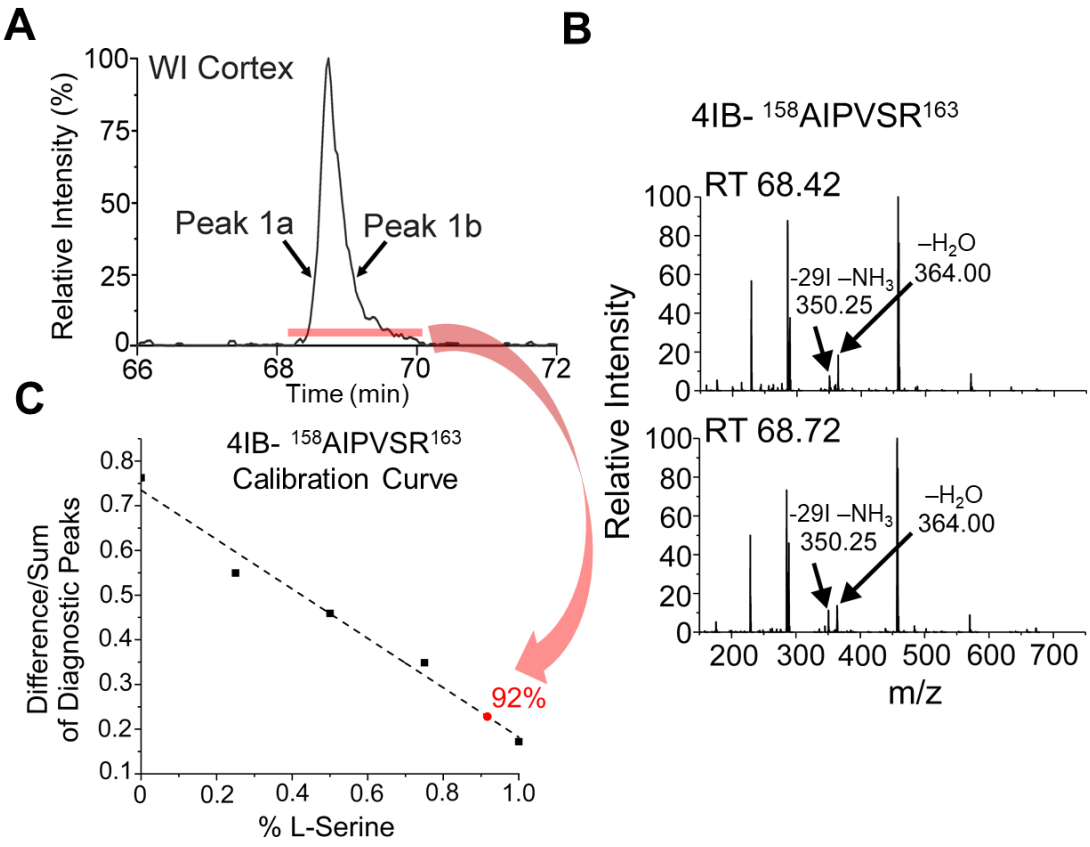


Figure 4.3 (A) LC chromatogram of 4IB-AIPVSR in the WI cortex digest of the 72 y/o lens. (B) RDD spectra from the front-end (peak 1a) and back-end (peak 1b) of the corresponding LC peak. (C) A calibration curve is then used to quantify the amount of D-Ser that co-elutes in the LC chromatogram. The curve is generated by making standard solutions that contain known amounts of both isomers and taking the difference over the sum of the two peaks that have the largest differences in the fragmentation spectra. For this peptide, the -29I-NH_3 losses from the precursor ion and the -H_2O loss from the precursor ion were chosen as the diagnostic peaks. The percent D-Ser/L-Ser in the digest is then determined by averaging the RDD spectra for the entire peak in part A (indicated by the red bar). This value maps to the red point in part (C), 92% L-Ser and 8% D-Ser.

Table 4.1 Idenfication of Ser epimers from 4IB-^{158}AIPVSR^{163} from αA in WI Cortex of 72 y/o Lens using R_{isomer} scores

Peak Observed from Lens	Synthetic standard	R_{isomer}
RT 68.42	L-Ser	4.6
RT 68.42	**D-Ser**	**1.7**

Peak Observed from Lens	Synthetic standard	R_{isomer}
RT 68.42	**L-Ser**	**1.3**
RT 68.42	D-Ser	6.0

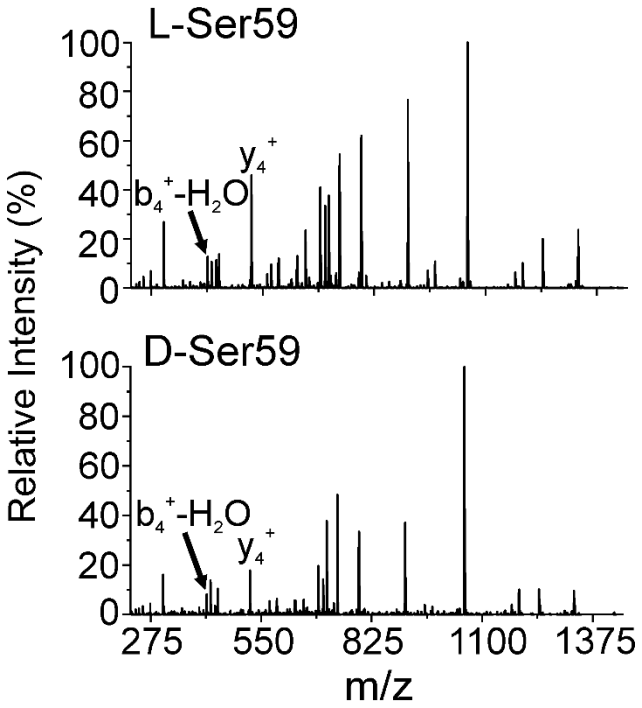


Figure 4.4 Resulting CID spectra of the APSWFDTGLSEMR L-Ser59 and D-Ser59 synthetic standards. Comparison of the b_4^+ -H2O and y_4^+ ions yields an R_{isomer} score of 4.1 indicating that CID can be used to detect D-Ser59 in this peptide. The fragmentation patterns of these two synthetics are then checked against each of the 9 peaks, allowing for identification of both L-Ser59 and D-Ser59.

Table 4.2 Identification of L-Ser59 of ^{57}APSWFDTGLSEMR^{69} from αB in WI Cortex

Synthetic standard	Peak Observed from Lens	R_{isomer}
L-Ser59	Peak 1	5.7
L-Ser59	Peak 2	3.4
L-Ser59	Peak 3	3.8
L-Ser59	Peak 4	4.2
L-Ser59	Peak 5	2.5
L-Ser59	Peak 6	8.3
L-Ser59	**Peak 7**	**1.3**
L-Ser59	Peak 8	2.9
L-Ser59	Peak 9	3.5

Table 4.3 Identification of D-Ser59 of ^{57}APSWFDTGLSEMR^{69} from αB in WI Cortex

Synthetic standard	Peak Observed from Lens	R_{isomer}
D-Ser59	Peak 1	5.4
D-Ser59	Peak 2	4.8
D-Ser59	Peak 3	4.9
D-Ser59	**Peak 4**	**1.4**
D-Ser59	Peak 5	3.3
D-Ser59	Peak 6	7.6
D-Ser59	Peak 7	4.9
D-Ser59	Peak 8	4.7
D-Ser59	Peak 9	4.5

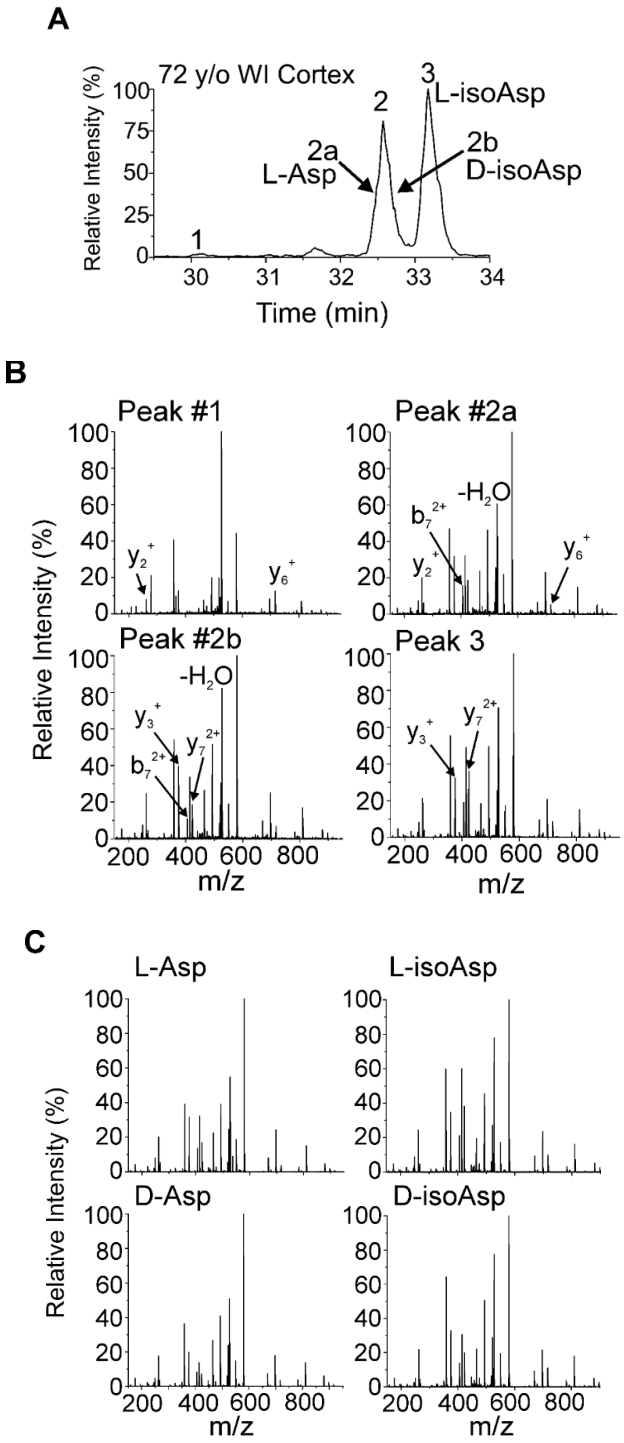

A
72 y/o WI Cortex
Relative Intensity (%)
Time (min)
1
2
2a
L-Asp
3
L-isoAsp
2b
D-isoAsp
B
Peak #1
Peak #2a
Peak #2b
Peak 3
-H2O
m/z
C
L-Asp
L-isoAsp
D-Asp
D-isoAsp

Figure 4.5 (A) LC chromatogram of the ^{108}pQDEHGFISR^{116} isomers from a 72 y/o WI cortex digest.(B) Four spectra resulting from collision-induced dissociation from three isomer peaks in the LC chromatogram. Peak 2 contains two co-eluting isomers and are labeled "2a and 2b". Comparison of the fragmentation intensities in peak #1 and peak #2 show differences in the b_6^+ ion, and by also comparing the fragmentation intensities of y_2^+ an R_{isomer} score can be calculated. In this example:

$$R_{isomer} = \frac{R_{Peak\ 1}}{R_{Peak\ 2a}} = \frac{\frac{I_A}{I_B}}{\frac{I_A}{I_B}} = \frac{\frac{12.48}{5.13}}{\frac{7.95}{20.04}} = 6.1;$$

Where I_A = Relative intensity of y_6^+, I_B = Relative Intensity of y_2^+.

This calculation is then used to compare the front end of peak 2 (2a) to the back end of peak 2 (2b), and then peak 2b to peak 3. The R_{isomer} scores for each of these subsequent calculations are all above the threshold of 1.9, indicating that they are different isomers. (C) Synthetic standards of the pQDEHGFISR are then used to identify the specific isomer in each of the peaks. This is done by comparing the fragmentation spectrum in each of the peaks to the fragmentation spectra from the authentic L-Asp, L-isoAsp, D-Asp and D-isoAsp synthetic versions. An R_{isomer} score below the 1.9 threshold allows for confident confirmation of the isomer.

Table 4.4 Idenfication of Asp isomers from ^{108}pQDEHGFISR^{116} from αB in WI Cortex of 72 y/o Lens using R_{isomer} scores

Peak Observed from Lens	Synthetic standard	R_{isomer}
Peak 1	L-Asp	5.4
Peak 1	L-isoAsp	3.9
Peak 1	D-Asp	5.9
Peak 1	D-isoAsp	4.1

Peak Observed from Lens	Synthetic standard	R_{isomer}
Peak 2a	**L-Asp**	**1.2**
Peak 2a	L-isoAsp	2.6
Peak 2a	D-Asp	2.3
Peak 2a	D-isoAsp	2.4

Peak Observed from Lens	Synthetic standard	R_{isomer}
Peak 2b	L-Asp	2.7
Peak 2b	L-isoAsp	3.8
Peak 2b	D-Asp	2.4
Peak 2b	**D-isoAsp**	**1.5**

Peak Observed from Lens	Synthetic standard	R_{isomer}
Peak 3	L-Asp	2.8
Peak 3	**L-isoAsp**	**1.2**
Peak 3	D-Asp	5.1
Peak 3	D-isoAsp	2.3

Table 4.4 lists the results of the isomer identification of the pQDEHGFISR peaks from the WI cortex of the 72 y/o lens, and the best matching isomer is indicated in red. When matching observed peptides to synthetic standards, low R_{isomer} values are desired to indicate similarity (the opposite of what is done to distinguish isomers from each other, where high R_{isomer} values are meaningful). Peak 1 (Figure S1a) centered at 30 min is an isomer of this peptide, but does not match any of the Asp synthetics suggesting that it contains a different site of isomeriation.

4.2.2 Epimerization of the C-terminal region compromises binding to α-crystallin domain

Ser162 from αA is an epimerized residue located adjacent to the highly conserved palindromic IXI/V motif within the C-terminal region that binds to a groove in the α-crystallin domain (Fig. 4.1A-B). This part of the sequence enables bi-directional domain-swap interactions by binding in the groove between β4 and β8 on an adjacent α-crystallin domains. [38] Examination of the available crystal structures (Fig. 4.6A) reveals that the C-terminal tail can occupy the groove in the β8 →β4 direction or the inverse direction (β4 →β8), stabilized by hydrogen bonding between the peptide backbones in either case. Although the binding to this groove is dynamic and can occur through two different states, these interactions are crucial for proper oligomerization, and prior work has shown that point mutations can influence the kinetics and thermodynamics of assembly.[39]

To quantify the effect of $^{\alpha B}$Ser162 epimerization on binding within this groove, we conducted a series of native MS experiments involving the α-crystallin domains of αA and αB, and isomeric variations of the peptide ERAIPVSRE, which includes the palindromic part of the C-terminal tail of αA. The core domains, when analyzed alone, yield peaks corresponding to monomer (αA), and both monomer and dimer (αB) (Fig. 4.6B-C, lower spectra). This is suggestive of the αA dimer interface being slightly weaker than that of αB. Competition binding experiments between the α-crystallin domain and palindromic peptides containing either D- or L-Ser162 are shown in Fig. 4.6B-C, and summarized in Fig. 4.6D. The canonical all L-amino acid form of the peptide, ERAIPVSRE, and a D-isomer variant made distinguishable by addition of glycine to both termini, GERAIPVsREG (s = D-Ser), were added at a 4:1 peptide/α-

crystallin domain molar ratio. The dominant adduct observed corresponds to the binding of the canonical peptide, with the D-isomer representing only a small fraction of the binding (see Fig. 4.6D).

To eliminate the possibility that the additional glycine residues might be responsible for the difference in binding, the inverse experiment was conducted with the glycine-shifted L-isomer (Fig. 4.6B-C, and insets, upper spectra). Binding of the L-isomer dominates again in both systems (Fig. 4.6D), and the data can be used for quantification[40] of the dissociation constants: αA core L-Ser = 48 μM, αB core L-Ser = 115 μM, αA core D-Ser = 650 μM, αB core D-Ser = 1380 μM. Based on these values, epimerization leads to destabilization of $\Delta\Delta G$ = ~6 kJ/mol. All of these results are consistent with D-$^{\alpha A}$Ser162 significantly inhibiting proper interaction between the C-terminal palindrome and the β4-β8 groove in the α-crystallin domain.

We have previously shown that removal of the side-chain at the equivalent position in αB ($^{\alpha B}$Thr162Ala) leads to a weaker interaction in the peptide-α-crystallin domain system and faster subunit-exchange of full-length αB.[39] The present data suggests that epimerization of $^{\alpha A}$Ser162 similarly affects the dynamics, which would be expected to also perturb the chaperone activity of αA. Furthermore, because the diversity of binding modes between the C-terminal tail and the α-crystallin domain promotes polydispersity and aids in preventing crystallization,[38] disruption of this binding may explain the increased abundance of the isomerized peptide in the WI lens fraction. Since both chaperone activity and polydispersity are presumed to help maintain lens transparency,

[28,41] epimerization of [αA]Ser162 is a likely contributor to age-related protein aggregation within the lens.

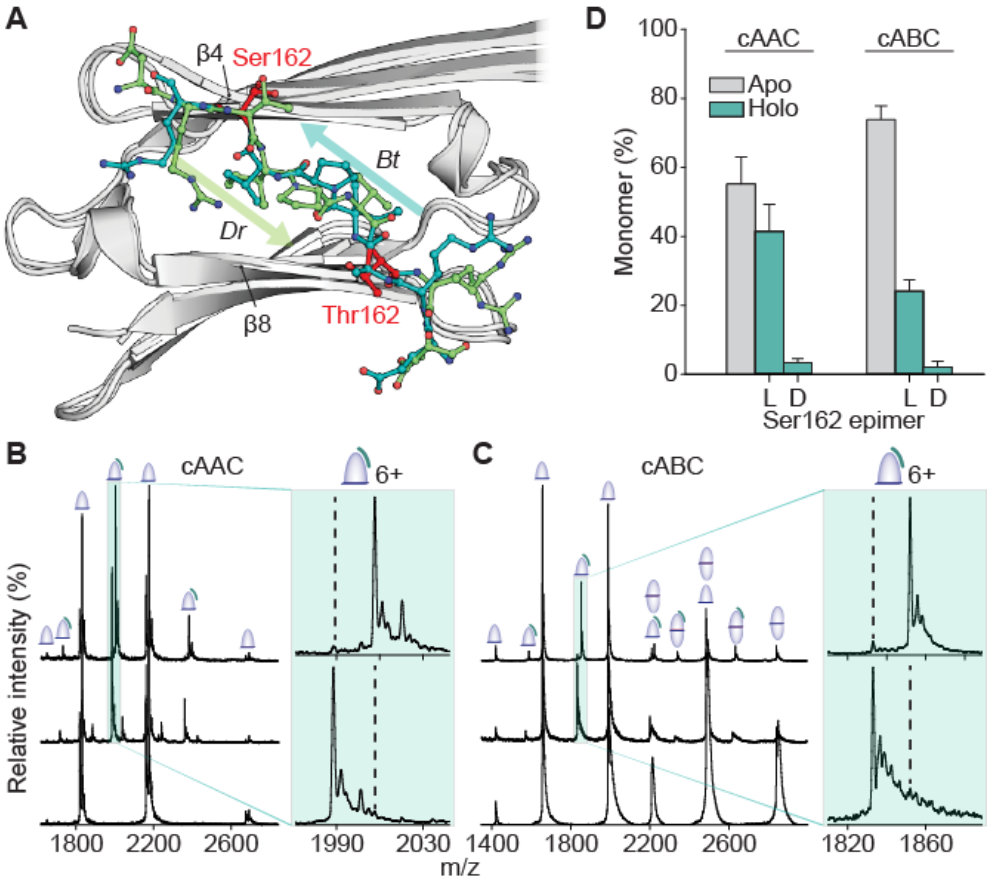


Figure 4.6 Competition experiments reveal a strong preference for L- over D-Ser162 binding to both αA and αB. (A) Aligned crystal structures of the α-crystallin domain (grey) with C-terminal peptide bound in two alternate orientations. Arrows indicate orientation (N → C) of bound peptides (green and teal). Equivalent isomerization sites (Ser162 in αA or Thr162 in αB) are shown in red. (B) Native mass spectra of αA core alone (bottom) and mixed with 4:1 ERAIPVSRE and GERAIPVsREG (middle, s=D-Ser). As seen in the magnification of 6+ peak, bottom spectrum, the binding of the lighter mass L-epimer is preferred. The upper trace corresponds to the reverse experiment, i.e. ERAIPVsRE and GERAIPVSREG. Dashed lines guide the eye to expected positions of D-epimer-bound peaks. (C) Identical experiment with the core of αB. (D) Relative average fractions of free versus bound cores from competition experiments, using all 5+, 6+ and 7+ charge states for quantitation. Error bars represent 95% confidence intervals. Crystal structures: bovine, *Bt*, in teal, PDB 3L1F, and zebrafish ,*Dr*, in green, PDB 3N3E.

4.2.3 Phosphorylation is precluded by epimerization of $^{\alpha B}$Ser59 and affected by isomerization of $^{\alpha B}$Asp62

Many observations suggest that Ser59 is a structurally important site, including solid-state NMR data that suggests Ser59 is involved in various inter-monomer contacts.[42]. As shown in Fig. 4.1C, the peptide containing $^{\alpha B}$Ser59 and $^{\alpha B}$Asp62 is highly isomerized in a normal lens. Furthermore, the degree of epimerization at Ser59 is increased in cataractous lenses.[43] Ser59 is also a primary site of phosphorylation in αB, which helps regulate the oligomeric size and activity.[44] However, phosphorylation is primarily observed in the younger cortex within the lens.[45]

To investigate the influence of $^{\alpha B}$Ser59 epimerization and $^{\alpha B}$Asp62 isomerization, we incubated the corresponding synthetic isomers of FLRAPSWFDTG-NH_2 with the native kinase, MAPKAPK-2.[46] The extracted ion chromatograms for both peptides reveal that the L-Ser59 isomer is the best substrate for the kinase (Fig. 4.7A,B). The ratio of L/D-Ser59 phosphorylation is ~240/1 after 2 hours and ~350/1 after 12 hours incubation. Interestingly, examination of the phosphorylated peptide in the lens reveals only minor isomerization (Fig. 4.1D), offering sharp contrast to the abundant isomerization of the unmodified peptide. This suggests that modifications elsewhere on the peptide may also inhibit phosphorylation, or that phosphorylation prevents isomerization, or both. Additional *in vitro* experiments confirmed that phosphorylation is inhibited by D-isoAsp62, but slightly enhanced by D-Asp or L-isoAsp at the same position (Fig. 4.7B). In sum, the native kinase activity is affected by all three non-native Asp isomers. These

results agree with previous experiments that failed to isolate phosphorylated D-Ser from erythrocytes.[47]

Ser59 and Asp62 have also been suggested to be involved in an interfacial salt-bridge cluster with Arg163 and Glu165 of an adjacent monomer.[48] In this model, all four residues are in close proximity, with isomerization of Asp62 or Ser59 both likely to perturb the dynamics of the salt bridge and disrupt the interface (Fig. 4.7C). Similarly, phosphorylation of Ser59, which resides 2.4 Å (O-O distance) from Asp62, is likely to disrupt this salt bridge network and may account for the reduced oligomer size observed for phosphorylation mimics of Ser59.[49]

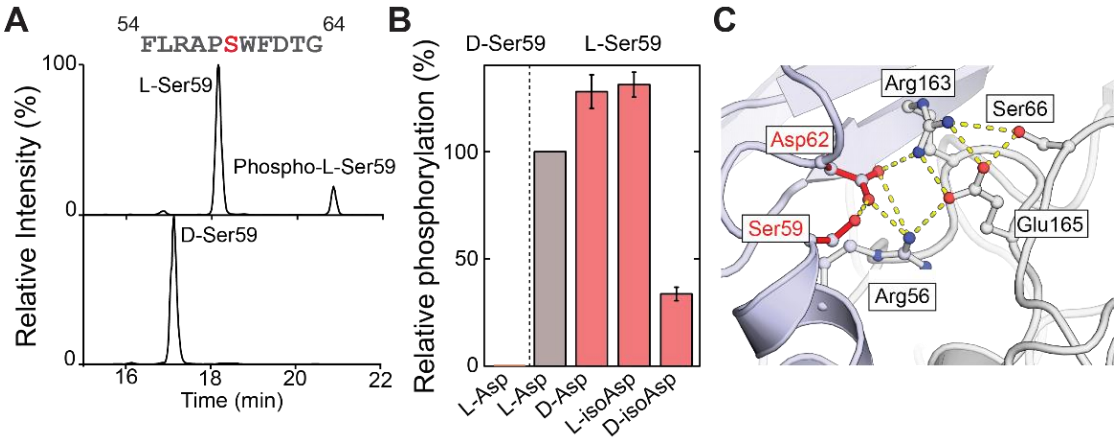


Figure 4.7 (A) Extracted ion chromatograms following incubation of FLRAPSWFDTG-NH_2 and FLRAPsWFDTG-NH_2 (s=D-Ser) with MAPKAPK-2 reveal that D-Ser is not a competitive phosphorylation substrate. (B) Relative degree of phosphorylation for Asp and Ser isomers of FLRAPSWFDTG-NH_2. (C) Salt-bridge model (PDB 2YGD) of an N-terminal oligomeric interface involving Ser59 and Asp62. Hydrogen bonds are shown using dashed yellow lines.

4.2.4 Isomerization of $^{\alpha B}$Asp109 breaks dimer stabilizing salt-bridge and leads to insolubility

Figure 4.1E illustrates abundant isomerization of $^{\alpha B}$Asp109, which is known to form an inter-monomer salt-bridge with $^{\alpha B}$Arg120 in the AP_{II} register,[38,50] the most populated

state in solution.[51] Mutation at either site in the salt bridge frequently leads to malfunction and disease.[50-53] For example, the R120G mutant is genetically linked to desmin-related myopathy,[54] while mutation of $^{\alpha B}$Asp109 is associated with myofibrillar myopathy[55] and cardiomyopathy.[56] To probe the structural consequences of $^{\alpha B}$Asp109 isomerization on the dimer interface of αB, *in silico* mutation and MD simulations were utilized, after parameterizing the force field for the isomeric amino acids.

The results obtained from all-atom simulations extending >150 ns are illustrated in Fig. 4.8. The Asp109-Arg120 salt bridge is stable for the L-Asp isomer, with a mean acceptor-to-donor hydrogen bond distance under 2 Å (Fig. 4.8A, left). For all other isomers, D-Asp109, L-isoAsp109, or D-isoAsp109, the salt-bridge with Arg120 is disrupted, and the average distances increase significantly (Fig. 4.8A, left). To determine the influence of Asp isomerization on the stability of the dimer interface, we monitored the distances between the final backbone hydrogen bond partners (His111 and Arg123). Reasonable hydrogen bond distances are only maintained for the L-Asp isomer, with all other isomers producing elongated distances (Fig. 4.8A, right). These results are illustrated by structural snapshots in Fig. 4.8B. For the L-Asp isomer (upper left), a backbone hydrogen bond between His111 and Arg123 links together the ends of β6+7 strands. For the other three isomers, these partners have been shifted to non-interacting distances due to disruption of the β-sheet dimer interface. It should be emphasized that these simulations model only one of two identical salt-bridges that stabilize the dimer interface. In the situation where two modified residues occupied both ends of the dimer interface, the resulting destabilization would be expected to be significantly worse. The

results from these simulations offer an explanation for the observed partitioning of each isomer extracted from the lens in Fig. 4.1E. To clarify, the abundance of isomerized Asp109 residues is much higher in the WI fraction, while L-Asp109 is the virtually the only isomer present in the WS fraction, suggesting isomerization leads to insolubility. The isomerization induced loss of the dimer interface mimics the effects of the R120G mutation, which is known to cause protein aggregation.[53]

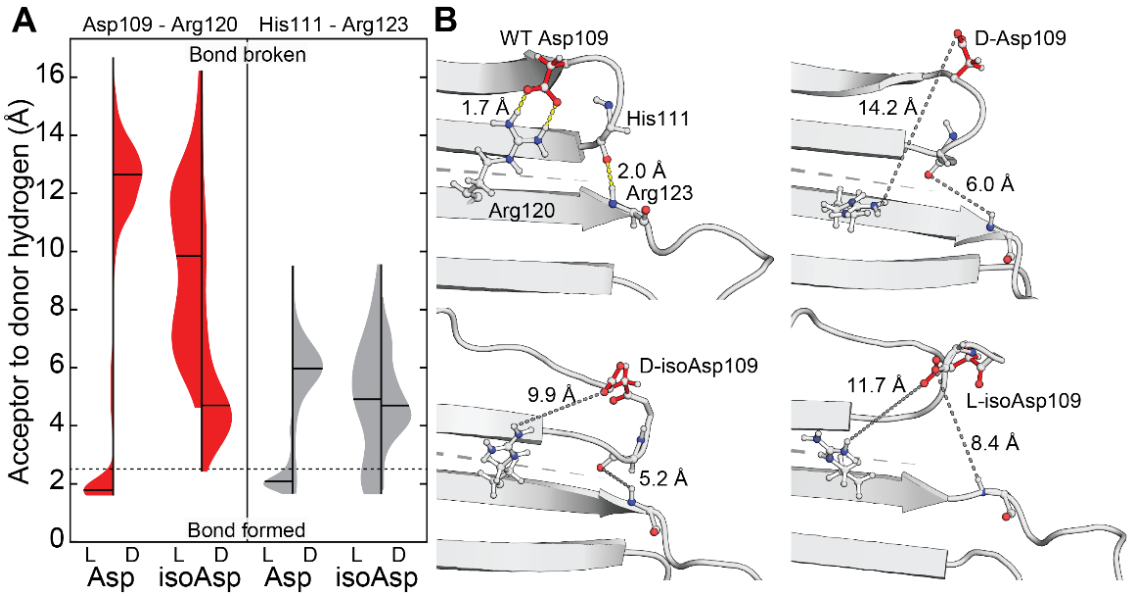


Figure 4.8 (A) Distance distributions between Asp109 and Arg120 (left, red), and His111 and Arg123 (right, grey) from MD simulations. Violin plots are shown for each isomer of Asp109; means are marked with black lines. Lengths <2.5 Å can be considered to correspond to bond formation (neglecting consideration of the bond angles), while those longer represent absence of the bond (boundary demarcated by dashed line). In all isomers other than L-Asp, the hydrogen-bond donor and acceptors are located too far apart for bond formation, the vast majority of the time. (B) Representative frames from MD simulations showing breakage of hydrogen bonds profiled in (A) and resultant interface destabilization. Yellow dashes indicate H-bonds; short gray dashes show concomitant distances following isomerization of Asp109; long gray dashes mark the antiparallel dimer interface. His111 and Arg123 side-chains have not been shown, for clarity.

4.2.5 Mimicking breakage of interfacial bonds by isomerization leads to aberrant oligomerization

Our examination of the amino-acid environment around both Ser59 and Asp109 revealed that in both cases modification at these sites would be likely to impact oligomerization. Specifically, we noted that $^{\alpha B}$Ser59 appears to be part of a network of salt bridges, into which a bulky negative charge seems unlikely to be accommodated without significant rearrangement. To test this prediction, we generated the phosphorylation mimic $^{\alpha B}$Ser59Asp and compared it to the wild-type with native MS. Both proteins gave mass spectra featuring a broad region of signal at high m/z, indicative of a polydisperse ensemble of oligomers, and consistent with previous spectra of αB (Fig. 4.9A).[57] Notably, the signal is at slightly lower m/z values for the phosphomimic, consistent with a shift to smaller stoichiometries. In order to quantify this change, we performed collisional activation to remove highly charged monomers from the parent oligomers, resulting in lower 'charge-stripped' oligomers that are well resolved and can be deconvolved into an oligomeric distribution (Fig. 4.9B,C).[57] Ser59Asp yields a distribution centered on an 18mer, smaller than the wild-type, and with a stronger preference for oligomers with an even number of subunits (Fig. 4.9C). This suggests that phosphorylation of Ser59, which regulates activity and localization of αB,[44] leads to destabilization of the larger oligomers, likely at inter-dimer interfaces involving the N-terminal region.

In addition to posing analytical challenges, isomerization and epimerization are also problematic from the perspective of molecular biology. It is not possible to use site-directed mutagenesis to insert D-residues or isoAsp into proteins, nor, given that both

modifications are spontaneous, is it feasible to induce isomerization or epimerization in a site-specific and controllable fashion. Direct investigation of the influence of isomerization at Asp109 on oligomeric assembly is therefore not feasible in the absence of whole protein synthesis. Nevertheless, the MD results show that the major consequence of L-Asp isomerization is disruption of the salt bridge with Arg120. This outcome can be mimicked with an Asp109Ala mutant, where the acidic partner in the salt bridge has been removed. Native MS experiments with this mimic revealed oligomeric assemblies larger than those observed with the wild-type protein, and a lower preference for even stoichiometries (Fig. 4.9C). Similar results were also obtained by light-scattering experiments (see Fig. 4.10). These results are consistent with trends observed for the Arg120Gly variant, in which the same salt-bridge is disrupted.[27] In both cases, the dimer interface is weakened and oligomer size increases. This may explain why the L-isoAsp variant is only observed in appreciable amounts in the insoluble fraction obtained from lenses (see Fig. 4.1E). Overall, the results in Fig. 5 confirm that subtle changes that mimic isomerization can impact oligomerization towards larger or smaller sizes even in the context of the full-length proteins.

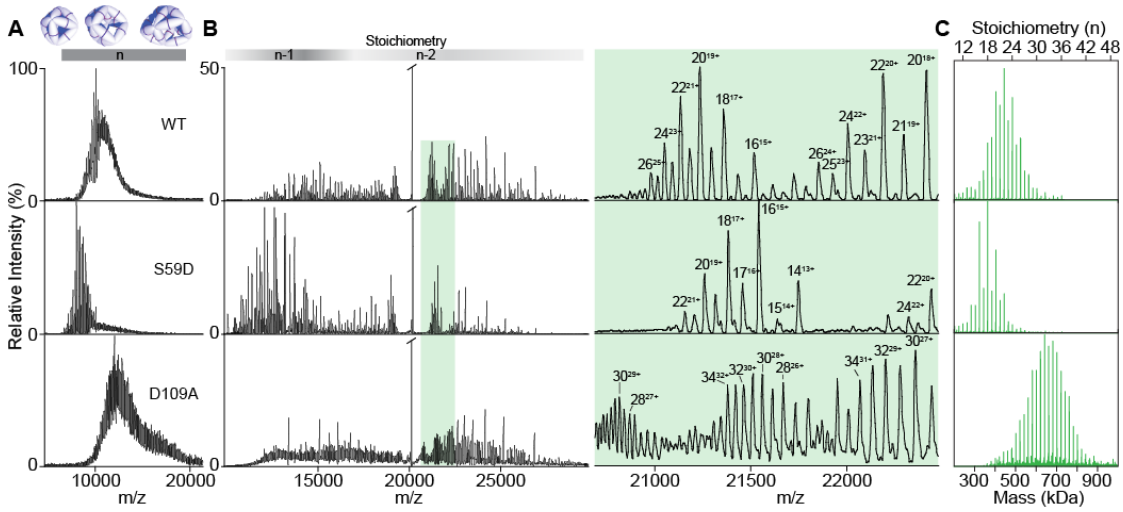


Figure 4.9 (A) Native MS of intact oligomeric assemblies of WT, Ser59Asp (phosphomimic), and Asp109Ala (isoAsp-mimic) αB, not directly assignable due to overlap of a multitude of charge states and stoichiometries. (B) Native MS with collision-induced dissociation (CID) of the a narrow slice from (A). Above shading shows regions where oligomers (n-mers) have lost one (n-1) or two (n-2) subunits. Detailed view of the region highlighted in green shows that with CID, charge states of unique oligomers are resolved. (C) Reconstructed oligomeric distributions. Data were charge deconvolved and then corrected to account for stripped subunits.

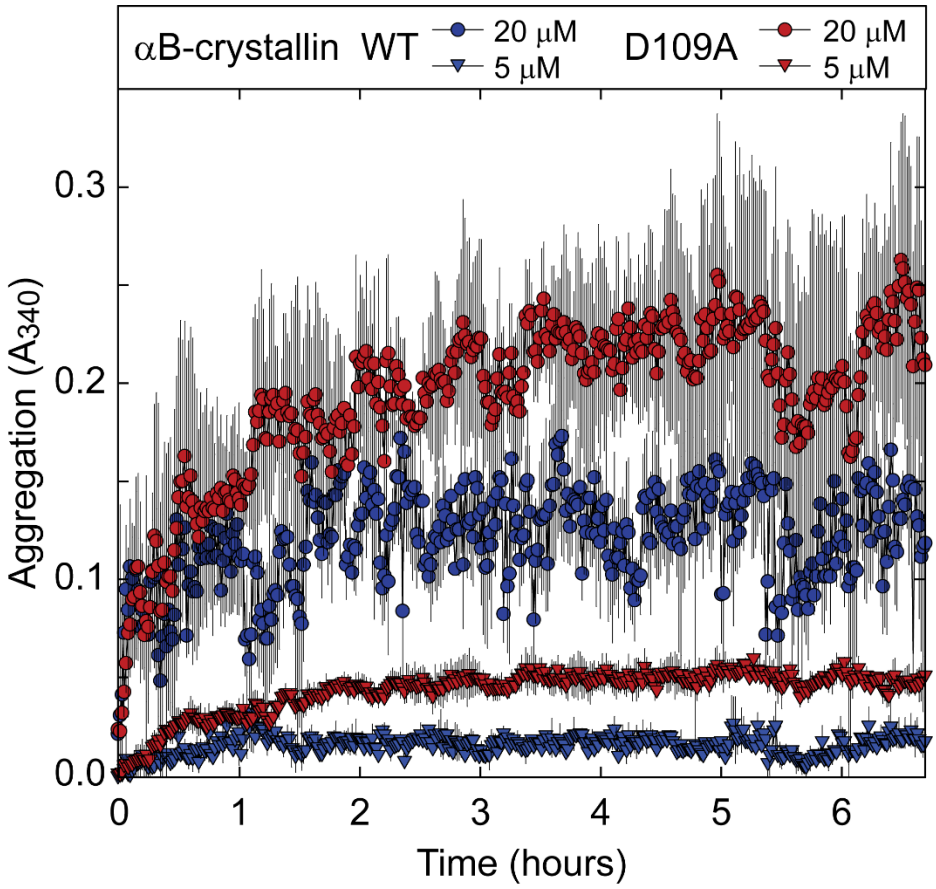


Figure 4.10. Aggregation of αB-crystallin (WT or D109A) over time at 42°C monitored by light scattering at 340 nm. Molar concentrations correspond to monomers. Error bars represent one s.d. (n = 3). The D109A substitution predisposes αB-crystallin to form large aggregates, in agreement with the native-MS results.

4.3 Conclusions

Isomerization and epimerization are prevalent PTMs in long-lived proteins such as the crystallins found in the eye lens. Although these modifications are difficult to detect and cannot be probed by site-directed mutagenesis, we have demonstrated that they can cause significant structural perturbation and loss of function. Epimerization of a single serine residue is sufficient to inhibit noncovalent recognition needed to maintain proper interface strengths and dynamics. Furthermore, epimerization of serine or isomerization

of nearby residues alters phosphorylation and any functionality derived from it. By redirecting and altering the peptide backbone, aspartic acid isomerization can also inhibit kinase recognition and disrupt native salt bridge interactions, leading to improper oligomer assembly and size.

Previous studies have also noted structural changes brought about by similar modifications to proteins. For example, hen egg-white lysozyme with an L-isoAsp substitution at Asp101 causes a backbone deflection of nearly 90° relative to the native structure,[58] and L-isoAsp32 insertion into ribonuclease forms a protruding U-shaped loop bent by nearly 90° instead of an α-helix.[59] Isomerization can also affect physical properties such as solubility and bioactivity, with isomerization of Asp92 in immunoglobulin γ2 (IgG2) leading to deactivation of the antigen-binding region.[60]

These observations show that what might appear to be innocuous PTMs can disrupt protein structure. Given that long-lived proteins are also associated with many other diseases, including Alzheimer's and Parkinson's,[1,61,62] it is likely that many of the structural issues highlighted herein also contribute to loss of function in these pathologies.

4.4 Experimental Section

4.4.1 Lens samples

Human lenses were acquired from the National Disease Research Interchange (NDRI) (Philadelphia, Pennsylvania). The nuclei and cortices of the thawed lenses were separated using a 4mm trephine. The endcaps of each nucleus were removed by gently scraping off the outer tissue until only the dense nuclear portion remained. The nucleus and cortex were then separately

homogenized in 600 µL of 50 mM Tris-HCl pH = 7.8. The lens fractions were then centrifuged at 15,100g for 20 min at 4ºC to separate the supernatant from the precipitate.

4.4.2 WS fraction.

The WS supernatant was purified by dialysis against water and lyophilized. 50 µg of lyophilized powder was added to 20 µL of 50 mM NH_4HCO_3 buffer, pH = 7.8. Disulfide bonds were then reduced with 1.5 µL of 100 mM of dithiothreitol (DTT) at 95°C for 5 minutes. After returning to room temperature, the solution was treated with 3 µL of 100 mM iodoacetamide (IAA) and incubated at room temperature in the dark for 20 minutes.

4.4.3 WI fraction.

The WI precipitate was solubilized in 6 M urea in 50 mM Tris-HCl pH=7.8 and purified by dialysis. Approximately 100 µg of the dialyzed protein (as determined by concentration based on absorbance at 280nm) was diluted to a total volume of 100 µL in 6 M urea in 50 mM Tris-HCl, pH = 8.0. Disulfide bonds were reduced with 5 µL of 200 mM DTT at 37°C for 20 minutes and then capped with 20 µL of 200 mM of IAA for one hour. The urea concentration was diluted to < 0.6 M using 50 mM Tris-HCl, 1 mM CaCl2, pH = 7.6.

Proteins were digested with trypsin for 12-16 hours at 37°C using a 50:1 protein to enzyme ratio. For samples requiring the iodobenzoate modification (i.e. RDD experiments), the digested peptides were desalted and cleaned using a peptide trap (Michrom Bioresource Inc). Approximately 5 nmoles of the digestion mixture, 10-100x excess of 4-iodobenzoic acid NHS-activated ester in dioxane and borate buffer (pH = 8.6) were combined and incubated for 1 hour at 37°C. Important: Dimethyl sulfoxide should not be substituted for dioxane in this step because it can cause aspartic acid isomerization. The modification side products at arginine and tyrosine side chains were removed by incubating the reaction mixture in 1 M hydroxylamine, pH = 8.5.

The same procedure was used to modify synthetic peptide standards. These procedures have been determined previously not to yield any detectable isomerization in control experiments.[36]

4.4.4 Peptide and Radical Precursor Synthesis

All synthetic peptides were synthesized manually using standard FMOC SPPS procedures with Wang resins, or Rink Amide resins as the solid support (29). N-hydroxysuccinimide (NHS) activated iodobenzoyl esters were synthesized by following a previous procedure. Synthetic model peptides were purified using a Jupiter Proteo column (250 mm x 4.6 mm, 4 μm, 90 Å, C12) (Phenomenex, Torrence, CA). These peptides were used to authenticate the identities of all peptide isomers and to generate calibration curves for accurate quantitation when full separation by LC was not achieved.

4.4.5 Phosphorylation Assay

Mitogen-activated protein kinase-activated protein kinase 2 (MAPKAPK-2, also MK2) was purchased from Genetex (Irvine, CA). The kinase substrate peptides were synthesized with Rink Amide resins to prevent a charged C-terminus from interfering with kinase recognition. 100μM of each purified synthetic peptide was assayed in 25 mM HEPES buffer at pH 7.5 with 25mM $MgCl_2$, and 2mM ATP in a 37 ºC sand bath. The assays were initiated by addition of 1μg kinase (for L-Ser and D-Ser assays) or 0.5μg kinase for (L-Asp, D-Asp, L-isoAsp, and D-isoAsp assays). Aliquots were taken at defined timepoints and the reaction was quenched by the addition of 1% TFA (v/v) before rapid freezing in liquid nitrogen. Samples were stored frozen until LCMS analysis. As a control, the all L- form of the FLRAPSWFDTG-NH_2 peptide was assayed without the addition of kinase to eliminate the possibility of autophosphorylation.

4.4.6 LC-MS Data Acquisition and Analysis-

An Agilent 1100 series HPLC system (Agilent, Santa Clara, CA) with a BetaBasic column (150 mm x 2.1 mm, 3 μm, 150 Å, C18) was coupled to an LTQ mass spectrometer. Peptides were

separated using a 0.1% formic acid in water (mobile phase A) and a 0.1% formic acid in acetonitrile (mobile phase B) binary system at a flow rate of 0.2 mL/min. The digestion mixtures were loaded onto the column and separated using the following gradient: 5% B to 20% B over 60 minutes, 20% B to 30% B over the next 45 minutes, 30% B to 50% B over the next 15 minutes and 50% B to 95% B over the final 10 minutes. The LTQ was operated in data-dependent mode using the Xcalibur program (Thermo Fisher Scientific). Specifically, in the CID-only LC-MS run, the first scan event was a full MS from m/z 400-2000 Da, followed by an ultrazoom (MS^2) and then CID (MS^3). In the RDD LC-MS experiments, the laser pulses were triggered during the MS^2 step and CID was performed as a pseudo-MS^3 step. Due to the high photodissociation yield of the 4-iodobenzoic acid chromophore, the major peak during this step is the loss of iodine, and it is the subsequent precursor for MS^3. The exclusion time was set to 60 seconds for the identification of peptides. For isomer identification, an inclusion mass list was added and the exclusion time was reduced to 16 seconds to enable repeated analysis of isomers.

4.4.7 R_{isomer} Calculations

To quantify isomer identification, the kinetic method using the R value, originally described by Tao and co-workers was utilized (31). In this paper, R_{isomer} represents the ratios of the relative intensities of a pair of fragments that varies the most between two isomers (R_A/R_B). Following acquisition of a tandem mass spectrum, R_{isomer} values are calculated for all pairs of peaks to reveal fragments that yield the best differentiation. If $R_{isomer} = 1$ then the two MS^n spectra are indistinguishable, and the species are likely not isomers. If $R_{isomer} > 1$, a larger number indicates a higher probability that the two species are isomers. To confidently identify each of these isomers by MS^n, we use a threshold that was determined by performing a *t*-test on the R_{isomer} values obtained by performing CID and RDD on a mixture of synthetic peptides separated by LCMS. Using 99% confidence intervals, the R_{isomer} threshold for CID is >1.9 and for RDD it is >2.4.[36]

4.4.8 Protein expression and purification

Core αB (cABC, residues 68-153) was expressed in *E. coli* and purified as described previously.[63] A gene insert encoding core αA (cAAC, residues 59-153) was purchased from Integrated DNA Technologies and inserted into a pET28a vector linearized with BamHI and XhoI (New England Biolabs) using an In-Fusion HD Cloning Kit (New England Biolabs) to generate a TEV-cleavable His-tagged construct. This was expressed and purified in the same manner as cABC with addition of 5 mM BME in all buffers prior to SEC, resulting in some population of BME-adducted protein visible in the spectra in Figure 2. Core domains were stored in 100 mM NaCl, 20 mM Tris, pH 8 at -80°C until use. Full length αB was expressed and purified as described previously.[63] Mutations S59D and D109A were introduced using a Quik-Change Site-Directed Mutagenesis kit (Agilent) and mutants were expressed and purified in the same manner as WT. Full length proteins were stored in MS buffer (200 mM ammonium acetate pH 6.9) at -20°C until use. Concentrations were determined by UV absorbance at 280 nm.

4.4.9 Native MS of core domains and peptides

Spectra were collected using a previously described protocol[64] on a Synapt G1 IM-QToF mass spectrometer (Waters) with parameters as follows: capillary 1.5 kV, sampling cone 40 V, extraction cone 3 V, backing pressure 3.1 mbar, trap gas (argon) 3 mL min^{-1}, trap cell voltage 10 V, transfer cell voltage 8 V. Ion mobility was enabled with parameters in the mobility cell: IMS gas flow 22 mL min^{-1}, IMS wave velocity 320 m s^{-1}, IMS wave height 5.5 V. Proteins were buffer exchanged into 200 mM ammonium acetate pH 6.9 using a Biospin-6 column (BioRad). All spectra were recorded at 10 μM based on

concentration measurement post- buffer exchange. Samples were introduced using gold-coated capillaries prepared in-house. Lyophilized peptides -- Ct (ERAIPVSRE) and G-Ct-G (GERAIPVSRE) with L or D-Ser – were resuspended in milliQ H_2O to a stock concentration of 1 mM and then diluted in MS buffer and mixed with protein immediately prior to analysis to a final concentration of 40 μM each for competition experiments. For quantitation, monomeric species were extracted in DriftScope (Waters) and intensities recorded from MassLynx using all resolved adduct peaks in addition to apo for 5+, 6+ and 7+ charge states. Data are reported as the mean +/- SD for three replicates.

4.4.10 Native MS of full length αB and mutants

Spectra were collected on a modified QExactive hybrid quadrupole-Orbitrap mass spectrometer (ThermoFisher Scientific) optimized for transmission of high-mass complexes.[65] Protein concentration was 15 μM by monomer. Capillary voltage was 1.4 kV in positive ion mode with source temperature 200°C and S-lens RF 200%. UHV pressure (argon) was between 1.4×10^{-9} and 1.7×10^{-9} mbar. In-source trapping fragmentation voltage ranged from -150 to -180 V. Ion transfer optics were as follows: injection flatapole 10 V, inter-flatapole lens 8 V, bent flatapole 6 V, transfer multipole 4 V, C-trap entrance lens 3 V. Nitrogen was used in the HCD cell and HCD energy was 0 V for intact spectra and tuned for optimal dissociation of each protein for CID spectra, ranging from 200 to 230 V. Resolution was kept at 17,500 at m/z = 200 for a transient time of 64 ms and the noise threshold was set to 3. For CID spectra, groupings of 30 microscans were combined to improve signal quality. Data were visualized using

Xcalibur (ThermoFisher Scientific) and calibrated manually according to expected peak positions for WT αB-crystallin. Calibrated CID data were processed using UniDec software[66] which allowed for stoichiometric assignment and post-hoc correction for dissociated subunits.

4.4.11 Molecular modeling

PDB 4M5T, featuring cABC (residues 67-151) with each subunit complexed with a C-terminal peptide (156-164), was exploited to assess the interactions between C-terminal peptide and crystallin domains. Two molecular dynamics simulations were prepared, one with C-terminal peptides bound as shown in the crystal structure, i.e. running from the β8 to the β4 strands of cABC, and one with the peptide bound in the opposite direction, i.e. β4 to β8. In order to build a model of the β4-β8 peptide, we selected only the alpha carbons of β8-β4 peptide, renumbered them in the opposite direction, and used them as template to align a peptide extracted from the crystal structure.

Structures were then simulated using the Amber ff14SB force field[67] on the NAMD molecular dynamics engine[68] Structures were first solvated in a box of TIP3P water, their box charge neutralized by addition of Na+ ions, and the resulting systems energy minimized with 2000 conjugate gradient steps. We then performed 0.5 ns steps in the NPT ensemble, with all protein's alpha carbons constrained by a harmonic potential. Langevin dynamics were used to impose a temperature of 300 K, using a damping of 1 /ps. A constant pressure of 1 Atm was imposed via a Langevin piston having a period of 200 fs, and a decay of 50 fs. The system was then further equilibrated in the NVT

ensemble for 1 ns, after which 200 ns production runs in the NPT ensemble were performed. In all simulation steps, Particle Mesh Ewald was used to treat long range electrostatic interactions, a cutoff distance of 12 Å was set on van der Waals interactions, and a 2 fs time step was exploited by restraining every covalent bond with SHAKE.

From each simulation, one frame per nanosecond was extracted from the production run for analysis. From each of these, hydrogen bonds between each residue part of a peptide and ones part of the crystallin domain were identified using VMD.[69] We counted the occurrences of each bond in order to determine the percentage of time two residues spend interacting. A bond was considered established if the distance between donor and acceptor atoms was less than 3 Å, and the acceptor-donor-hydrogen angle less than 20 degrees. Residue pairs forming simultaneously more than one H-bond in the same frame would be counted as bonded only once.

In order to study the effect of Asp109 epimerization we exploited PDB 2WJ7, featuring a cABC dimer in Aβ II register (thus enabling the formation of the Asp109-Arg120 salt bridge). Three models featuring D-Asp, D-isoAsp and L-isoAsp at position 109 in one of the two monomers, respectively, were produced by modifying PDB 2WJ7 in Schrödinger Maestro. Simulation parameters for non-standard amino acids were produced with Antechamber[70,71] All atom types could be assigned according to available ff14SB parameters. All models were simulated in explicit solvent for 150 ns following the same simulation protocol described above. His111-Arg123 and Asp109-Arg120 distances were measured every 100 ps. For the latter, we report the shortest distance between each of the hydrogens of Arg120 guanidinium, and oxygens of Asp109

carboxylate. The percentage of time a hydrogen bond is established was determined as described above, and averaged over the three simulations.

1 R. J. W. Truscott, K. L. Schey and M. G. Friedrich, *Trends Biochem. Sci.*, 2016, 41, 654.
2 R. J. W. Truscott and M. G. Friedrich, *Biochim. Biophys. Acta*, 2016, 1860, 192.
3 Y. A. Lyon, G. M. Sabbah and R. R. Julian, *J. Proteome Res.*, 2017, **16**, 1797.
4 M. Y. S. Hooi and R. J. W. Truscott, *Age (Omaha).*, 2011, **33**, 131.
5 T. Geiger and S. Clarke, *J. Biol. Chem.*, 1987, **262**, 785.
6 N. Fujii, T. Takata, N. Fujii and K. Aki, *Biochim. Biophys. Acta - Gen. Subj.*, 2016, **1860**, 183.
7 T. Takata and N. Fujii, *FEBS J.*, 2016, **283**, 850.
8 Y. A. Lyon, G. M. Sabbah and R. R. Julian, *Exp. Eye Res.*, 2018, **171**, 131.
9 P. M. Masters, J. L. Bada and J. S. J. Zigler, *Proc. Natl. Acad. Sci. U. S. A.*, 1978, **75**, 1204.
10 M. Y. S. Hooi, M. J. Raftery and R. J. W. Truscott, *Protein Sci.*, 2013, **22**, 93.
11 W. Ni, S. Dai, B. L. Karger and Z. S. Zhou, *Anal. Chem.*, 2010, **82**, 7485.
12 E. T. Jansson, *J. Sep. Sci.*, 2018, 41, 385.
13 J. B. Catterall, D. Barr, M. Bolognesi, R. D. Zura and V. B. Kraus, *Arthritis Res. Ther.*, 2009, **11**, R55.
14 J. Samuel Zigler and J. Goosey, *Trends Biochem. Sci.*, 1981, 6, 133.
15 J. Horwitz, *Proc. Natl. Acad. Sci. U.S.A*, 1992, **89**, 10449.
16 W. C. Boelens, *Prog. Biophys. Mol. Biol.*, 2014, **115**, 3.
17 J. Horwitz, M. P. Bova, L. L. Ding, D. A. Haley and P. L. Stewart, *Eye*, 1999, **13**, 403.
18 T. Iwaki, A. Kume-Iwaki and J. E. Goldman, *J. Histochem. Cytochem.*, 1990, **38**, 31.
19 F. Bennardini, A. Wrzosek and M. Chiesi, *Circ. Res.*, 1992, **71**, 288.
20 R. J. W. Truscott and M. G. Friedrich, *Biochim. Biophys. Acta - Gen. Subj.*, 2016, 1860, 192.
21 T. Iwaki, T. Wisniewski, A. Iwaki, E. Corbin, N. Tomokane, J. Tateishi and J. E. Goldman, *Am. J. Pathol.*, 1992, **140**, 345–56.
22 K. Renkawek, C. E. Voorter, G. J. Bosman, F. P. van Workum and W. W. de Jong, *Acta Neuropathol(Berl)*, 1994, **87**, 155.
23 G. J. Caspers, J. A. M. Leunissen and W. W. de Jong, *J. Mol. Evol.*, 1995, **40**, 238.
24 G. K. A. Hochberg and J. L. P. Benesch, *Prog. Biophys. Mol. Biol.*, 2014, **115**, 11.
25 M. Haslbeck, J. Peschek, J. Buchner and S. Weinkauf, *Biochim. Biophys. Acta - Gen. Subj.*, 2016, **1860**, 149–166.
26 E. Basha, H. O'Neill and E. Vierling, *Trends Biochem. Sci.*, 2012, 37, 106.
27 M. P. Bova, O. Yaron, Q. Huang, L. Ding, D. A. Haley, P. L. Stewart and J. Horwitz, *Biochemistry*, 1999, **96**, 6137.
28 A. R. Clark, N. H. Lubsen and C. Slingsby, *Int. J. Biochem. Cell Biol.*, 2012, **44**, 1687.

29 K. Kato, Y. Inaguma, H. Ito, K. Iida, I. Iwamoto, K. Kamei, N. Ochi, H. Ohta and M. Kishikawa, *J. Neurochem.*, 2001, **76**, 730.
30 R. Maddala and P. V. Rao, *Exp. Cell Res.*, 2005, **306**, 203.
31 J. A. Aquilina, J. L. P. Benesch, L. L. Ding, O. Yaron, J. Horwitz and C. V. Robinson, *J. Biol. Chem.*, 2004, **279**, 28675.
32 H. Ecroyd, S. Meehan, J. Horwitz, J. A. Aquilina, J. L. P. Benesch, C. V. Robinson, C. E. Macphee and J. A. Carver, *Biochem. J.*, 2007, **401**, 129.
33 R. Li and G. Reiser, *J. Neurochem.*, 2011, **118**, 354.
34 S. A. Chandler and J. L. Benesch, *Curr. Opin. Chem. Biol.*, 2018, 42, 130.
35 H. Li, Y. Sheng, W. McGee, M. Cammarata, D. Holden and J. A. Loo, *Anal. Chem.*, 2017, **89**, 2731.
36 Y. Tao and R. R. Julian, *Anal. Chem.*, 2014, **86**, 9733.
37 D. L. Riggs, S. V. Gomez and R. R. Julian, *ACS Chem. Biol.*, 2017, **12**, 2875.
38 A. Laganowsky, J. L. P. Benesch, M. Landau, L. Ding, M. R. Sawaya, D. Cascio, Q. Huang, C. V. Robinson, J. Horwitz and D. Eisenberg, *Protein Sci.*, 2010, **19**, 1031.
39 G. R. Hilton, G. K. A. Hochberg, A. Laganowsky, S. I. Mcginnigle, A. J. Baldwin and J. L. P. Benesch, *Philos. Trans. R. Soc. B Biol. Sci.*, 2013, **368**, 1–13.
40 A. El-Hawiet, E. N. Kitova, L. Liu and J. S. Klassen, *J. Am. Soc. Mass Spectrom.*, 2010, **21**, 1893.
41 A. Tardieu, *Int. J. Biol. Macromol.*, 1998, **22**, 211.
42 S. Jehle, P. Rajagopal, B. Bardiaux, S. Markovic, R. Kühne, J. R. Stout, V. A. Higman, R. E. Klevit, B. J. Van Rossum and H. Oschkinat, *Nat. Struct. Mol. Biol.*, 2010, **17**, 1037.
43 M. Y. S. Hooi, M. J. Raftery and R. J. W. Truscott, *Protein Sci.*, 2013, **22**, 93.
44 R. Bakthisaran, K. K. iran Akula, R. Tangirala and C. M. ohan Rao, *Biochim. Biophys. Acta*, 2016, 1860, 167.
45 A. C. Grey and K. L. Schey, *Mol. Vis.*, 2008, **14**, 171.
46 K. Kato, H. Ito, K. Kamei, Y. Inaguma, I. Iwamototo and S. Sagal, *J. Biol. Chem.*, 1998, **273**, 28346.
47 P. N. McFadden, L. L. Lou, K. Drickamer and S. Clarke, *Arch. Biochem. Biophys.*, 1987, **259**.
48 N. Braun, M. Zacharias, J. Peschek, A. Kastenmuller, J. Zou, M. Hanzlik, M. Haslbeck, J. Rappsilber, J. Buchner and S. Weinkauf, *Proc. Natl. Acad. Sci.*, 2011, **108**, 20491.
49 M. F. Ahmad, B. Raman, T. Ramakrishna and C. M. Rao, *J. Mol. Biol.*, 2008, **375**, 1040.
50 A. R. Clark, C. E. Naylor, C. Bagnéris, N. H. Keep and C. Slingsby, *J. Mol. Biol.*, 2011, **408**, 118.
51 G. K. A. Hochberg, H. Ecroyd, C. Liu, D. Cox, D. Cascio, M. R. Sawaya, M. P. Collier, J. Stroud, J. A. Carver, A. J. Baldwin, C. V. Robinson, D. S. Eisenberg, J. L. P. Benesch and A. Laganowsky, *Proc. Natl. Acad. Sci.*, 2014, **111**, E1562.
52 S. Simon, M. Michiel, F. Skouri-Panet, J. P. Lechaire, P. Vicart and A. Tardieu, *Biochemistry*, 2007, **46**, 9605.

53 M. Michiel, F. Skouri-Panet, E. Duprat, S. Simon, C. Férard, A. Tardieu and S. Finet, *Biochemistry*, 2009, **48**, 442.
54 P. Vicart, A. Caron, P. Guicheney, Z. Li, M. C. Prévost, A. Faure, D. Chateau, F. Chapon, F. Tomé, J. M. Dupret, D. Paulin and M. Fardeau, *Nat. Genet.*, 1998, **20**, 92.
55 J. P. Fichna, A. Potulska-Chromik, P. Miszta, M. J. Redowicz, A. M. Kaminska, C. Zekanowski and S. Filipek, *BBA Clin.*, 2017, **7**, 1.
56 A. Brodehl, A. Gaertner-Rommel, B. Klauke, S. A. Grewe, I. Schirmer, A. Peterschröder, L. Faber, M. Vorgerd, J. Gummert, D. Anselmetti, U. Schulz, L. Paluszkiewicz and H. Milting, *Hum. Mutat.*, 2017, **38**, 947.
57 J. A. Aquilina, J. L. P. Benesch, O. A. Bateman, C. Slingsby, and C. V. Robinson, *Proc. Natl. Acad. Sci*., 2003, ***100(19)***, 10611.
58 S. Noguchi, K. Miyawaki and Y. Satow, *J. Mol. Biol.*, 1998, **278**, 231.
59 S. Noguchi, *Biopolymers*, 2010, **93**, 1003.
60 D. S. Rehder, D. Chelius, A. McAuley, T. M. Dillon, G. Xiao, J. Crouse-Zeineddini, L. Vardanyan, N. Perico, V. Mukku, D. N. Brems, M. Matsumura and P. V. Bondarenko, *Biochemistry*, 2008, **47**, 2518–2530.
61 J. N. Savas, B. H. Toyama, T. Xu, J. R. Yates and M. W. Hetzer, *Science*, 2012, **335**, 942.
62 A. Ben-Zvi, E. A. Miller and R. I. Morimoto, *Proc. Natl. Acad. Sci.*, 2009, **106**, 14914.
63 G. K. A. Hochberg, H. Ecroyd, C. Liu, D. Cox, D. Cascio, M. R. Sawaya, M. P. Collier, J. Stroud, J. A. Carver, A. J. Baldwin, C. V. Robinson, D. S. Eisenberg, J. L. P. Benesch and A. Laganowsky, *Proc. Natl. Acad. Sci.*, 2014, **111**, E1562
64 F. D. L. Kondrat, W. B. Struwe and J. L. P. Benesch, in *Structural Proteomics: High-Throughput Methods: Second Edition*, 2014, pp. 349.
65 R. J. Rose, E. Damoc, E. Denisov, A. Makarov and A. J. R. Heck, *Nat. Methods*, 2012, **9**, 1084.
66 M. T. Marty, A. J. Baldwin, E. G. Marklund, G. K. A. Hochberg, J. L. P. Benesch and C. V. Robinson, *Anal. Chem.*, 2015, **87**, 4370.
67 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696.
68 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781.
69 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33.
70 J. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Mol. Graph. Model.*, 2006, **25**, 247.
71 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157.

Chapter 5

Leveraging Electron Transfer Dissociation for Site Selective Radical Generation: Applications for Peptide Epimer Analysis

5.1 Introduction

Radical chemistry has played an important role in the development of mass spectrometry over the years, and the combination continues to evolve and produce useful applications. Unfortunately, many modern ionization sources are geared towards gentle transfer of large biomolecular ions into the gas phase, and do not typically generate odd-electron ions in significant abundance.[12] Methods for introducing radicals post-ionization are therefore required. Several options are available, including photodissociation[3] and collisional-activation based approaches.[4-6] Radical chemistry can also be combined with ion-ion reactions utilizing either activation method.[7] Once created, odd-electron chemistry has many uses in the gas phase including: cross-linking,[8] examining three dimensional structure,[9] accessing novel dissociation pathways,[10,11] and isomer identification.[12] Radical-directed dissociation (RDD) is often sensitive to fine structural details because dissociation is preceded by radical migration, which is guided by the relative orientation of hydrogen atom(s) that must be abstracted to allow relocation of the nascent radical to the ultimate site of fragmentation.[13]

A key for many of these applications is site-specific generation of a radical via homolytic photodissociation of labile bonds. Dissociation occurs due to electronic excitation to dissociative excited states, leading to fragmentation on the femtosecond timeframe.[14] Many bonds have been demonstrated to be suitable, including carbon-

iodine, carbon-sulfur, and sulfur-sulfur.[15] In optimal cases, radical yields can be quantitative with a single laser pulse,[16] but implementation requires coupling an ultraviolet laser to a mass spectrometer—an experimental configuration that is not currently commercially available. The need for customized instrumentation has likely limited accessibility to the technique.

In contrast, electron-transfer dissociation (ETD) is another radical based fragmentation method that is widely available on many commercial platforms. ETD was developed following discovery of electron capture dissociation (ECD), and both facilitate fragmentation by delivering an electron to a multiply charged cation, often yielding very similar results. In ETD, the electron originates from a molecular anion and transfer results in charge reduction, energy increase, and dissociation. For peptides, fragmentation into c/z ions typically dominates, with the side benefit that post-translational modifications (PTMs) are often preserved.[17] ETD can also be used to identify isomers, such as iso-aspartic acid, for which a unique fragment ion is generated.[18] ECD has been used to interrogate three-dimensional protein structure by mapping out the relative strength of noncovalent interactions between sequence remote regions.[19] Thus, there are many similarities between applications of RDD and ETD/ECD. Interestingly, electron transfer also converts an ion into a hydrogen abundant radical, which can spontaneously convert into a hydrogen deficient radical, the same type of radical operative in RDD. Indeed, some of the fragmentation observed in ETD/ECD is attributable to RDD,[20] but precise mechanisms accounting for all observed fragmentation in ETD/ECD are not universally agreed upon and have been the source of significant study and controversy.[21]

In the present work, we present a method for generating site-specific cleavage of carbon-iodine bonds via ETD rather than photodissociation. Both homolytic cleavage due to excitation of dissociative electronic states yielding loss of iodine radical (-I•), and loss of iodine accompanied by hydrogen (-HI) are observed. Interestingly, the ratio of these two losses varies as a function of additional collisional activation, with the amount of -HI loss increasing dramatically with greater activation for some peptides. A novel electron storage mechanism involving the side chain of arginine, which allows decoupling in time of electron transfer and dissociation, is proposed to account for these observations. Density-functional calculations support the proposed mechanism. Finally, the utility of ETD-generated radicals for isomer identification is illustrated by examination of several peptide isomers, dissimilar in structure due to isomerization or epimerization at a single residue. For these applications, the ability to simultaneously generate both even- and odd-electron species is shown to be beneficial in some cases.

5.2 Experimental

5.2.1 Materials

Organic solvents were purchased from Sigma Aldrich (St. Louis, MO) or Fisher Scientific and used without further purification. Water was purified to 18.2 MΩ by a Millipore 147 (Billerica, MA) Direct Q system. Amino acids and resins were purchased from Anaspec (Fremont, CA). Trypsin and urea were purchased from Sigma Aldrich (St. Louis, MO).

5.2.2 Peptide and Radical Precursor Synthesis

All synthetic peptides were synthesized manually using standard FMOC procedures with Wang and Rink-Amide Resins as the solid support.[22] N-hydroxysuccinimide (NHS) activated iodobenzoyl esters were synthesized by following a previous procedure.[16] Chloramine-T, sodium metabisulfite, and sodium iodide were purchased from Fisher Chemical (Fairlawn, NJ). Water was purified to 18.2 MΩ resistivity using a Millipore Direct-Q (Millipore, Billerica, MA. Peptides were iodinated by modification of a previously published procedure.[3]

5.2.3 Mass Spectrometry and dissociation of the peptides

Mass spectrometric analysis was performed using an Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific, San Jose, CA). Solutions containing ~1 μM of protein in 49.5:49.5:1 MeCN/H2O/AcOH were directly infused using the Nanospray Flex Ion source (Thermo Scientific, San Jose, CA) at 2 uL per min through a 500 uL syringe. 10 μm ID emitters were pulled using a laser-based micropipette puller (Sutter Instrument, Novato, CA). High energy collision induced dissociation (HCD) was used for the supplemental activation (from 0 to 35 normalized collisional activation) in ETD which is termed EThcD. For ETD and EThcD experiments, 100 ms was used for fluoranthene reagent and peptide ion reaction time. The isolation window used for each step was 2-3 m/z to maximize ion count. For ETD-CID experiments, ETD or EThcD was first performed to generate the -I/-HI losses, then these peaks were isolated and subjected to collision induced dissociation (CID) of appropriate energy.

5.2.4 R_{isomer} calculations

To quantify isomer identification, an R_{chiral} value approach, originally reported by Tao *et. al.*, was utilized.[23] In this paper, R_{isomer} represents the ratios of the relative intensities of a pair of fragments that varies the most between two isomers (R_A/R_B). If $R_{isomer} = 1$ then there the two mass spectra are the same and the species are not distinguishable. If $R_{isomer} >1$, a larger number indicates a better degree of recognition. CID and RDD values were calculated using the +1 precursor charge state, while ETD-CID was performed on the +2 as charge reduction to the +1 occurs before dissociation.

5.2.5 Calculations

Transition state calculations were carried out using density functional theory as implemented in Gaussian09. The B3PW91 functional was used with the following basis sets: LANL2DZ for iodine loss, 6-31+g(d) for disulfide cleavage, 6-31g(d) for c/z dissociation. Transition states were found using the QST3 approach and verified by calculating vibrational frequencies, for which a single imaginary value was obtained. The transition state for electron transfer to the disulfide functionality was not well behaved, and was therefore approximated via a step-wise scan of S-S bond length with all other coordinates freely optimizable. Time-dependent density functional theory at the B3LYP/LANL2DZ level was used to optimize the lowest energy excited state of iodobenzene.

5.3 Results and Discussion

ETD on the 2+ charge state of 4-iodobenzyl-modified VQEDFVEIHGK yields the spectrum shown in Figure 1a. As expected for an ETD experiment, c and z ions and an

abundant charge reduced peak, corresponding to ETnoD,[24] are generated. However, the most abundant fragment results from loss of iodine, which is not a dissociation pathway previously observed to be abundant in ETD/ECD experiments. Closer examination, as shown in the inset of Figure 5.1a, reveals that the loss of both iodine and HI are observed. The high intensity of these losses implies that they are produced via energetically favorable pathways. An obvious possibility would be transfer of an electron (i.e. during the ET part of ETD) to the iodo-functional group, facilitating loss of iodine. Indeed, addition of an electron to iodobenzene in *ab initio* calculations leads to spontaneous loss of iodide anion upon minimization as shown in Figure 5.1b. This prediction by theory is also in agreement with previous experiments.[25] The product following electron attachment is iodide anion, which is unlikely to be lost from a doubly protonated peptide due to strong Coulombic attraction. However, intramolecular transfer of a proton to the iodide anion would yield HI, which should be less strongly bound and is an observed dissociation channel. This pathway for the generation of HI loss is illustrated in the upper path of Figure 5.2.

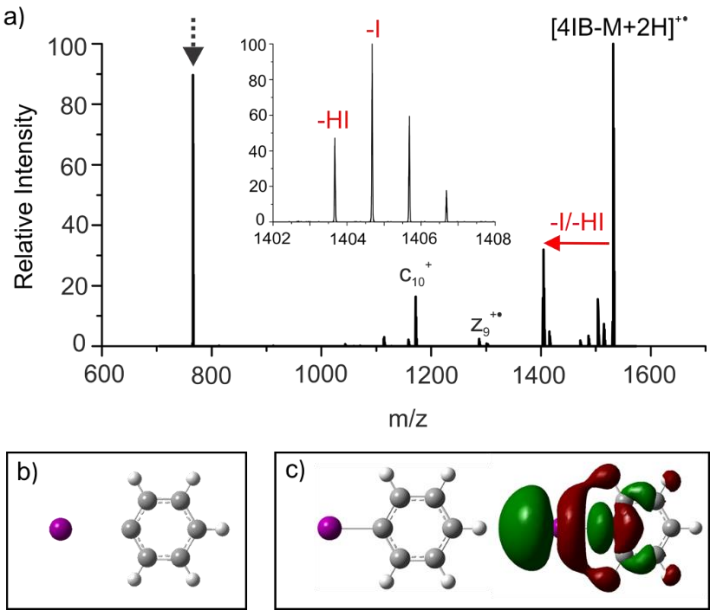


Figure 5.1 a) Figure 1. a) ETD on [4IB-VQEDFVEIHGK+2H]2+ leads to loss of –HI/-I. Inset shows the ratio of these losses. b) Addition of an electron to iodobenzene leads to spontaneous loss of iodide anion. c) Lowest unoccupied molecular orbital for iodobenzene is dissociative along the C-I bond.

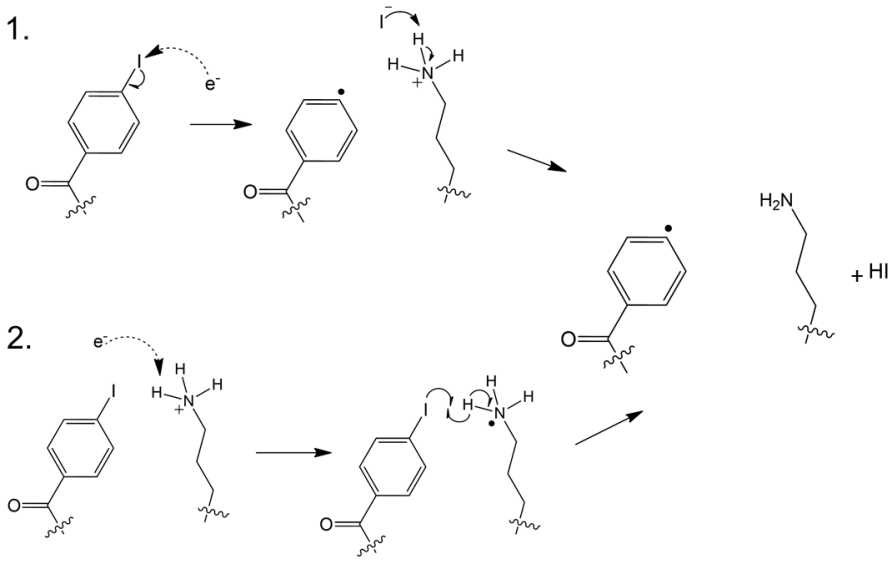


Figure 5.2 Two potential pathways that could account for observed loss of –HI.

The loss of neutral iodine is also observed in Figure 5.1a and is more difficult to explain because, as mentioned above, direct interaction of the electron with the iodobenzyl moiety should lead to dissociation of negatively charged iodide, not neutral iodine. Furthermore, iodide anion would not be expected to relinquish an electron back to the peptide. TDFT calculations on iodobenzene reveal that the lowest unoccupied molecular orbital is dissociative along the C-I bond, see Figure 5.1c. Population of this orbital via electronic excitation leads to prompt loss of neutral iodine, a pathway that is frequently accessed in photodissociation experiments, where the dissociative state is populated by absorption of a UV photon.[16] In order to account for the neutral iodine loss in ETD, we propose that electron transfer leads to temporary occupation of an excited state by the excess electron, which subsequently decays to a lower energy state while simultaneously exciting an electron in the iodobenzyl group, prompting homolytic cleavage of the C-I bond. This type of energy coupling between separate excited state electrons has been observed previously in the context of energy transfer.[26]

Another potential route for HI loss is illustrated in the lower portion of Figure 5.2. If the transferred electron is captured at the protonated side chain of lysine or the N-terminus, a hypervalent amine is generated, which could subsequently attack the iodobenzyl moiety, leading to loss of HI. This mechanism would likely require close proximity between the protonated and iodobenzyl groups. In order to explore the feasibility of the two pathways shown in Figure 5.2, experiments utilizing supplemental activation were carried out on 4IB-YLRPPSFLR-NH_2 and the results are shown in Figure 5.3. Less fragmentation is observed for this peptide compared to 4IB-VQEDFVEIHGK,

but loss of iodine is still one of the most abundant pathways. ETD without additional activation is shown in Figure 5.3a. On the right side of Figure 5.3 the ratio of I/HI loss is illustrated in the zoomed in spectra. With no additional activation, the loss of HI is minimal for 4IB-YLRPPSFLR-NH_2. However, with increasing collisional activation energy, the relative amount of HI loss increases dramatically. At the highest activation energy, loss of HI/I becomes the most abundant fragmentation channel (Figure 5.3e). There are two likely explanations that could account for these observations. First, it is possible that HI is created immediately after electron transfer, but not observed due to retention by noncovalent bonds.[27] Second, it is possible that HI is not generated until additional activation is applied.

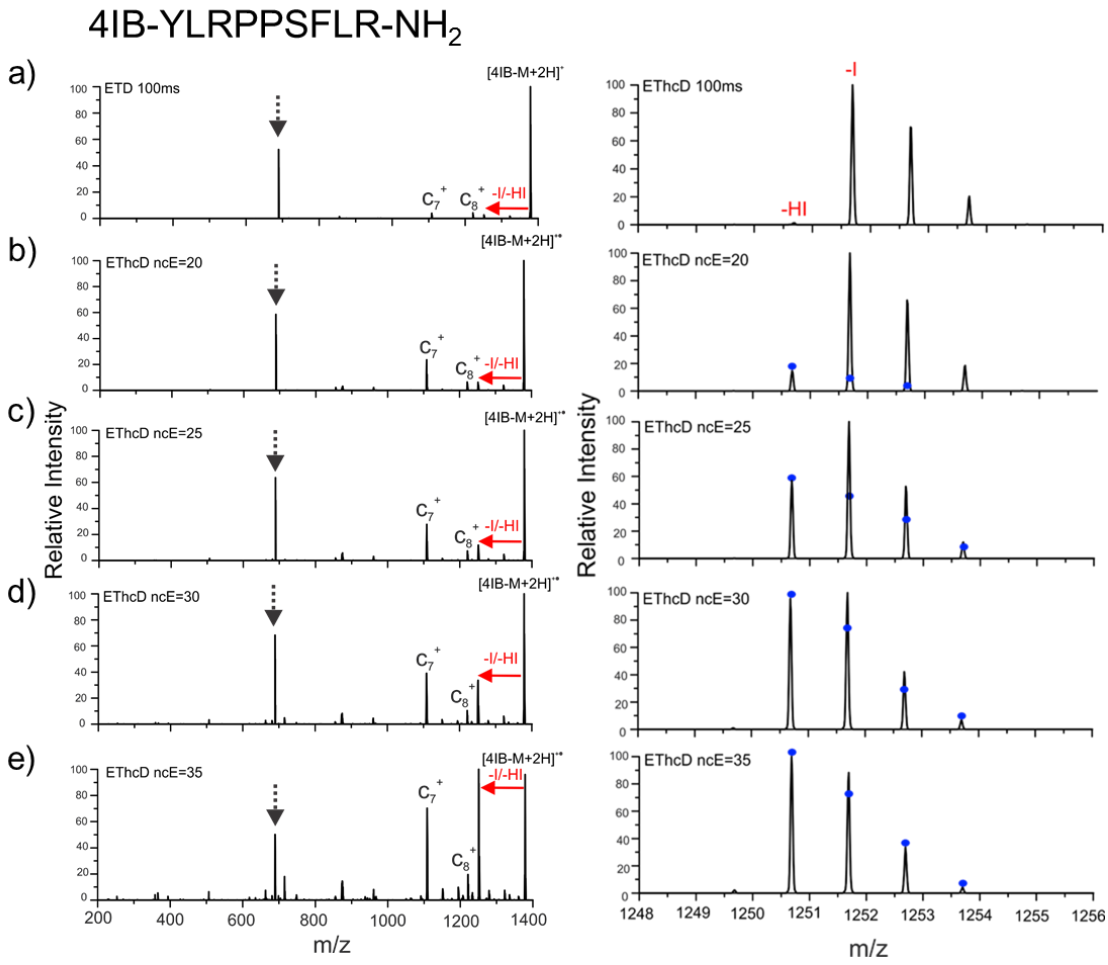


Figure 5.3. ETD on $[4IB\text{-}YLRPPSFLR\text{-}NH_2+2H]^{2+}$ with increasing supplemental activation. -I/-HI losses from the charge-reduced species are indicated by red arrows. The –I/-HI losses are shown more clearly in zoomed in spectra on right. Blue dots represent the theoretical isotopic distribution for the –HI peak that grows in, illustrating that the –I peak does not increase with increasing SA.

Examination of the ratios of HI loss to c_7^+ or c_8^+ in the 4IB-YLRPPSFLR-NH_2 peptide reveals that the relative abundance of HI increases dramatically with increasing activation energy while the relative abundance of c-type ions remains relatively flat (see Figure 5.4), which supports delayed generation of HI. Collision-induced dissociation of the -HI/I products reveals a dissociation spectrum that is similar to what is seen when performing

radical-directed dissociation (RDD), where the radical was generated by photodissociation (see Figure 5.5). In other words, ETD can be used for cleavage of the C-I bond, eliminating the need for a laser to carry out RDD experiments. Applications for this new fragmentation technique are detailed below.

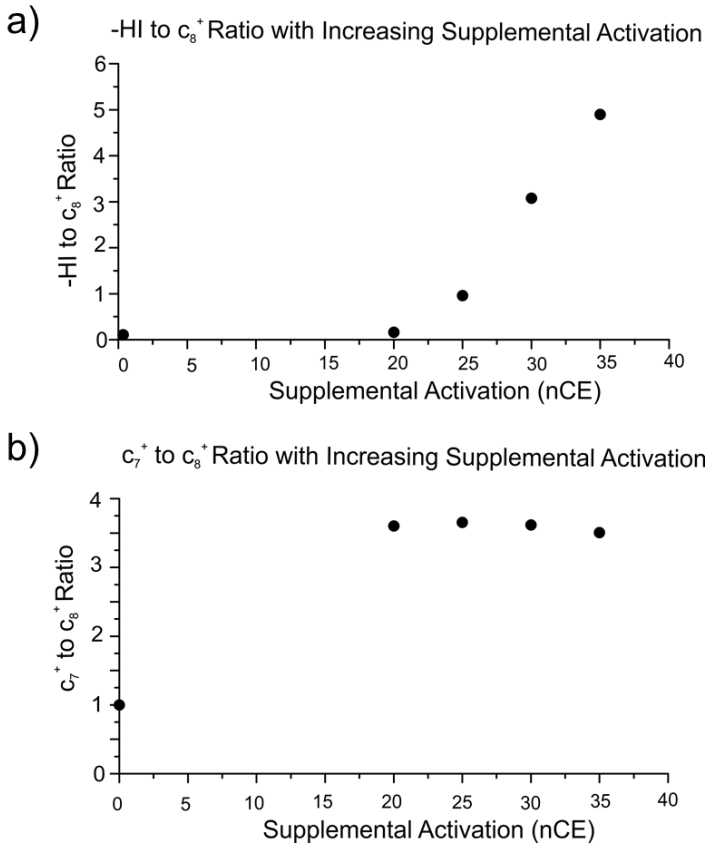


Figure 5.4 a) Ratio of –HI loss to c_8^+ with increasing supplemental activation in 4IB-YLRPPSFLR-NH_2. b) Ratio of c_8^+ to c_7^+ with increasing supplemental activation. While the ratio of –HI to c_8^+ increases with increasing SA, the ratio of c_7^+ to c_8^+ does not. These results support the idea that the –HI loss is a delayed process.

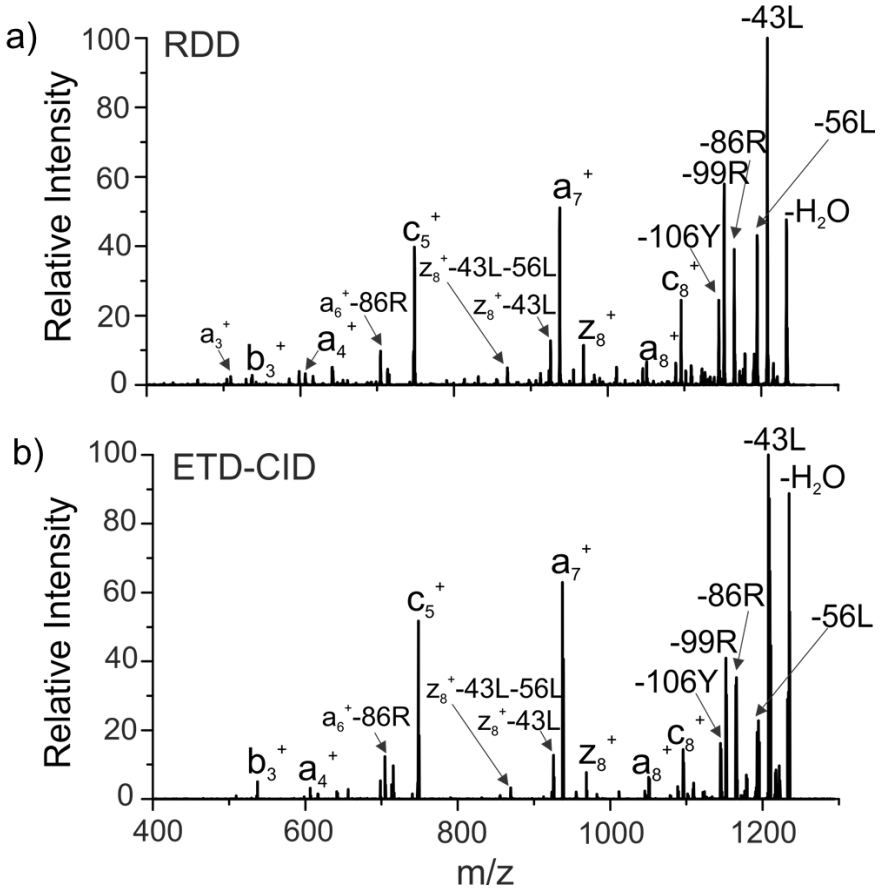


Figure 5.5 Comparison of the fragmentation spectra resulting from a) RDD on [4IB-YLRPPSFLR-NH_2+H]$^+$ and b) ETD-CID on [4IB-YLRPPSFLR-NH_2-2H]$^{2+}$. The spectra are very similar.

However, we first wish to explore a potential mechanism for delayed HI production. Delayed loss of HI requires a mechanism capable of separating, in time, electron transfer from HI formation. We have demonstrated previously that electron capture at protonated arginine and lysine side chains is not equivalent.[20] The hypervalent amine generated at lysine is not stable towards dissociative H-atom loss and therefore is not likely to be involved with delayed HI loss. In contrast, electron capture at protonated arginine creates a stable radical site by loss of a double bond. Theory was used to evaluate whether neutralized radical arginine side chain could attack the iodobenzyl functionality and

initiate loss of HI. DFT calculations identified a reaction pathway proceeding by electron transfer from the neutral arginine side chain to the iodobenzene, followed by loss of iodide, and eventual formation of an ionic complex between I^- and protonated arginine side chain, see Figure 5.6. The transition state for this pathway is only ~18 kJ/mol above reactants, and exhibits an elongated C-I bond and partially replanarized guanidinium. Upon completion, the reaction is exothermic and yields products ~-77 kJ/mol downhill from reactants. The system was calculated with freely translatable/rotatable functional groups and likely represents the lowest energy pathway (with respect to sterics). In an actual peptide, steric constraints would likely introduce additional barriers associated with rearrangement of the peptide structure to accommodate the transition state geometry. This pathway provides a viable mechanism for delayed HI production because transfer of the electron to the arginine side chain can be decoupled from subsequent attack of the iodobenzyl group by arginine. Importantly, the lysine analog, 4IB-YLKPPSFLK- NH_2, does not exhibit increased loss of HI following collisional activation (Figure 5.7).

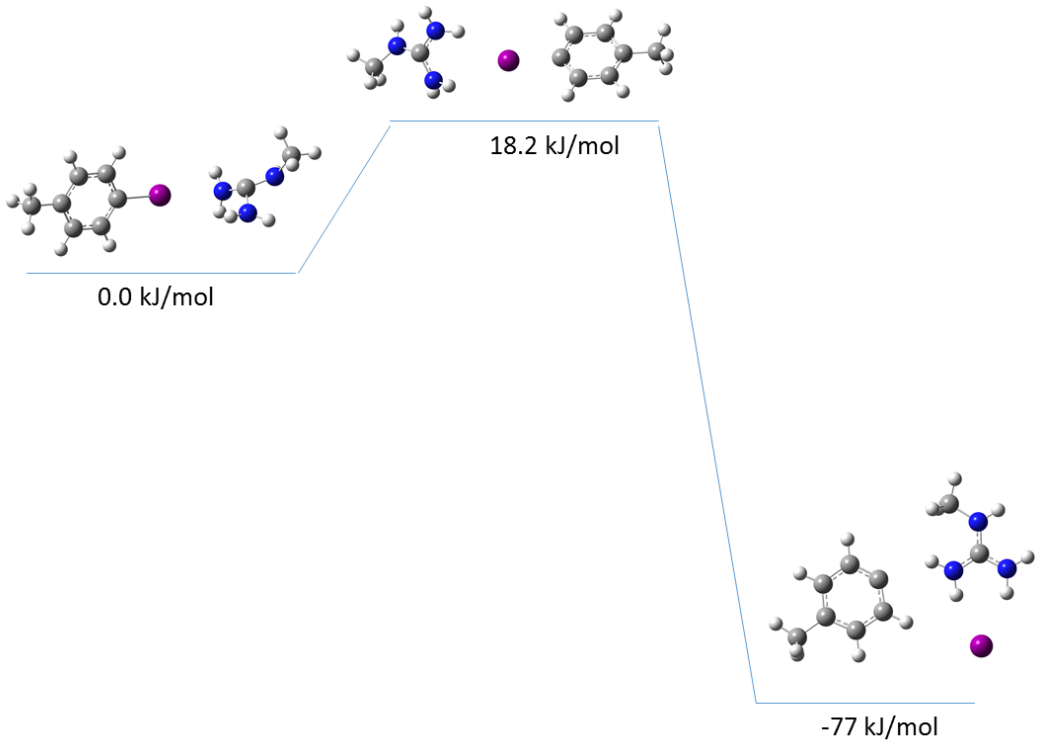


Figure 5.6 Structures and transition states revealed by DFT calculations as described in the text.

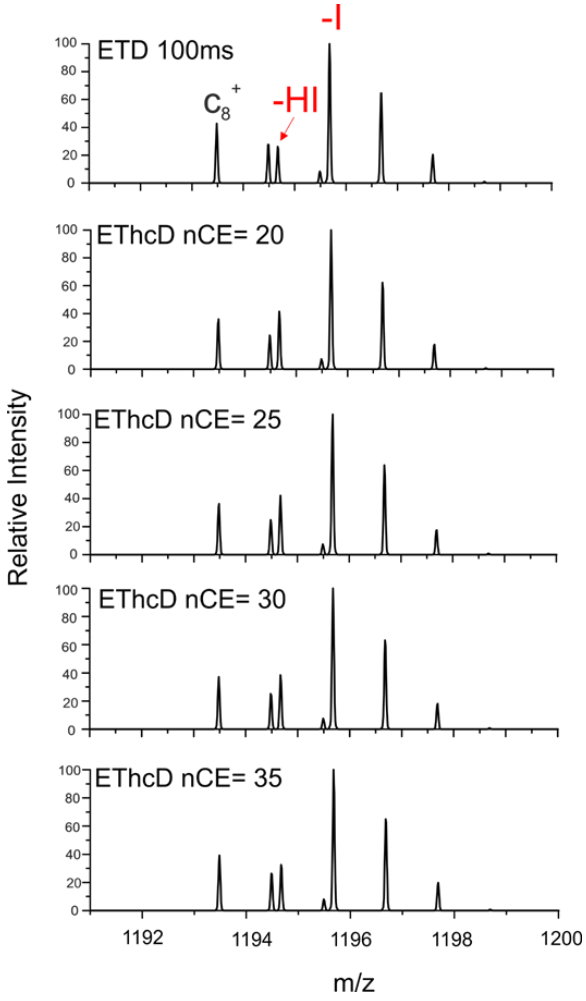


Figure 5.7 –HI/-I ratio with increasing supplemental activation in 4IB-YLKPPSFLK-NH2. The losses do not increase with activation. This mass range also includes the c8+ fragment, but can be discerned using high resolution.

Intrigued by the idea that radical arginine side chain could initiate delayed HI loss in ETD/ECD, we examined other well-known dissociation pathways. Calculations reveal that radical arginine side chain can also initiate c/z backbone fragmentation and disulfide bond cleavage via pathways analogous to the one shown in Figure 5.6. Complete details including energetics and structures are shown in Figures 5.8, 5.9 of the supporting information. Backbone cleavage to yield c/z ions is slightly endothermic, though still possible with supplemental collisional activation. Disulfide bond fragmentation is

exothermic and protected by a minimal ~30 kJ/mol transition state. This pathway may contribute to the favored cleavage of disulfide bonds observed previously in ETD/ECD.[28]

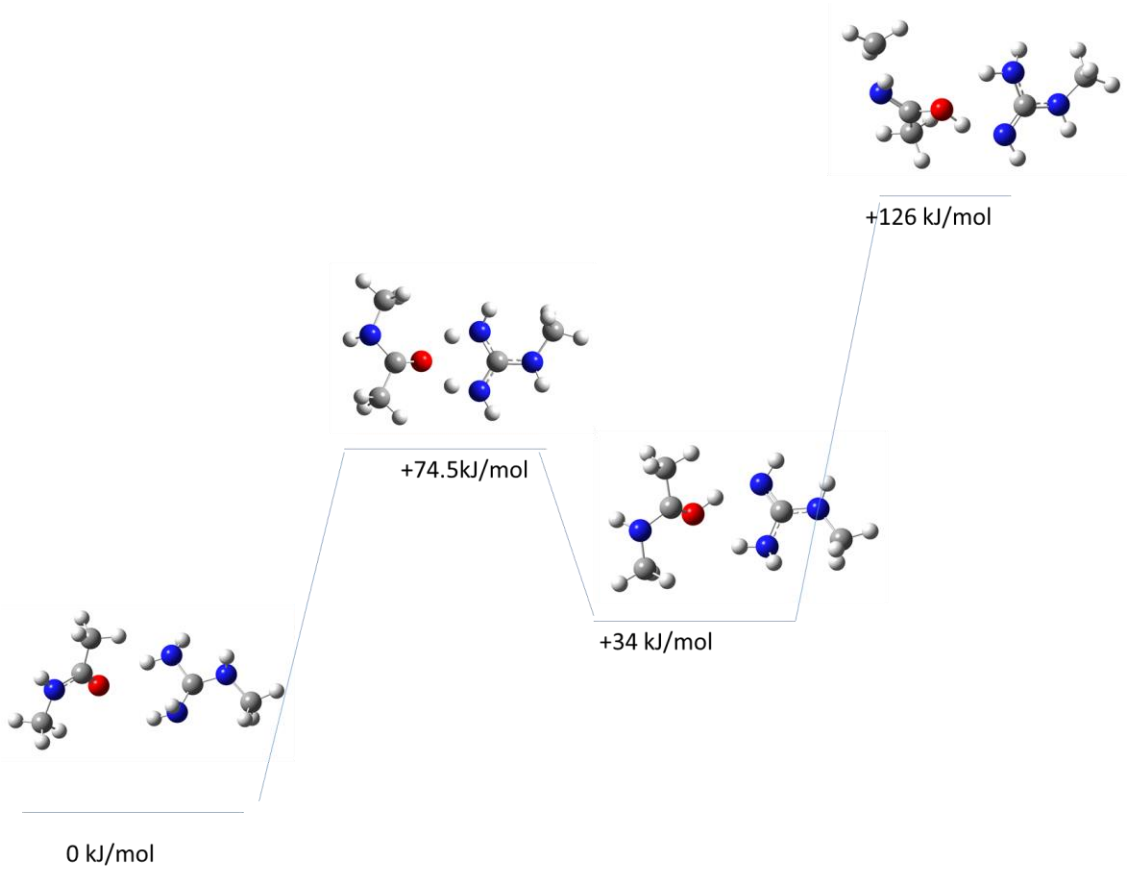


Figure 5.8. Energies associated with c/z ion formation. The first transition state allows electron transfer from the methylguanidine radical to the peptide backbone. The second transition state corresponds to cleavage of the backbone.

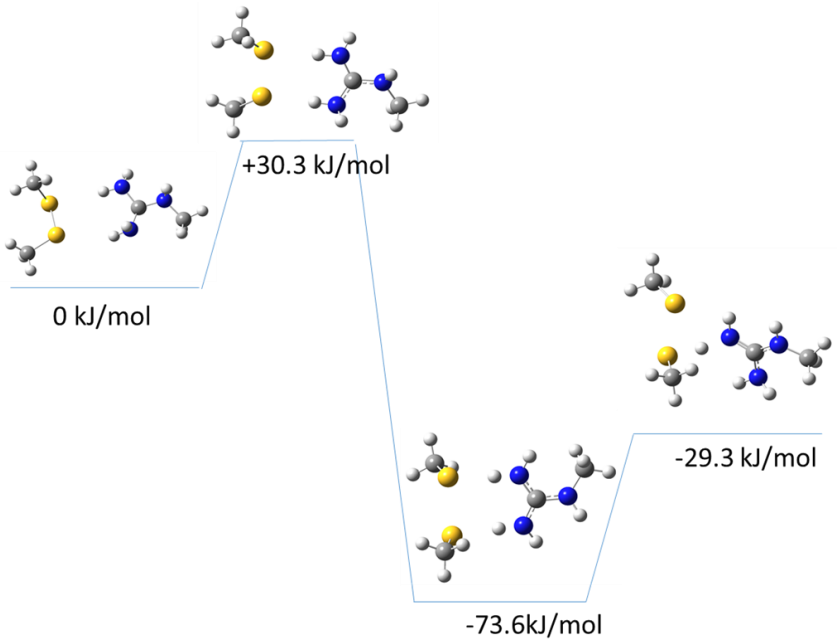


Figure 5.9. Energetics associated with disulfide bond cleavage. The first transition state corresponds to electron transfer from the methylguanidine radical to the disulfide bond. The second transition state leads to transfer of a hydrogen atom to form methyl-thiol.

5.3.1 Application

The activation of carbon-iodine bonds by ETD as described above can be used to generate radicals site-specifically, enabling structurally sensitive RDD experiments that have advantages for isomer identification as described in the introduction. In the past these experiments were only possible by photoactivation of C-I bonds, but now it is clear that ETD may be used as well. Although in photodissociation-based experiments the fragment of interest is the loss of iodine, the equivalent radical generated by ETD is obtained by loss of HI. The difference occurs because electron transfer to a multiply protonated peptide yields a hydrogen abundant, odd-electron species. Loss of radical iodine from this odd electron ion would therefore yield an even-electron product. However, loss of HI, which is even electron, maintains the radical on the peptide. To test

the structural sensitivity of fragmentation initiated by ETD-generated radicals, we conducted a series of experiments on peptide isomers, including epimers.

In vivo peptide/protein synthesis occurs exclusively from L-amino acids, but aging[29] or enzymatic manipulations[30,31] can lead to inversion of chirality for some residues. Peptides with a single D-amino acid are epimers of the all L-amino acid isomer, have identical exact masses, and are very difficult to identify. RDD initiated by photoactivation has been demonstrated previously to be the most sensitive method for epimer identification via comparison of the fragmentation spectra acquired independently from both epimers. Given that ETD activation of carbon-iodine bonds leads to generation of both -HI and -I losses, it is possible to interrogate both radical and even-electron products simultaneously. Figure 5.10a shows results from collisional activation of the -I/-HI products from the L-Asp isomer of 4IB-VQEDFVEIHGK. Both side-chain losses that are unique to RDD and proton-mediated backbone cleavages leading to b/y ions from CID are observed in Figure 5.10a. In Figure 5.10b, the typical CID spectrum for the same peptide is shown, revealing preferential formation of b/y ions, as expected. The pure RDD spectrum, obtained in a separate experiment by photodissociation of the C-I bond followed by collisional activation, is shown in Figure 5.10c. The RDD spectrum is dominated by typical side-chain losses, accompanied by a few backbone cleavages. It is clear that a combination of the results from Figures 5.10b and 5.10c would yield the spectrum in Figure 5.10a, confirming that both even and odd-electron peptides can be fragmented simultaneously in the ETD-activation experiment.

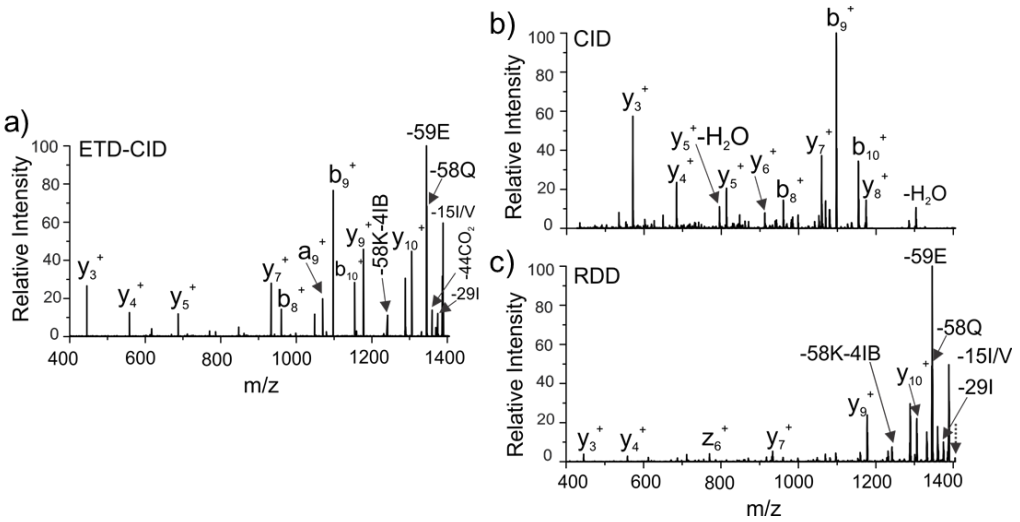


Figure 5.10. a). Collisional activation of –I/HI products from ETD for protonated 4IB-VQEDFVEIHGK. b) CID spectrum. c) RDD spectrum.

ETD was used for iodo-activation of other isomers of 4IB-VQEDFVEIHGK including: L-isoAsp, D-Asp, and D-isoAsp. Figure 5.11 shows the ratios of the -HI/-I losses that are observed for each isomer. Although the largest peak corresponds to -I loss in each case, there are noticeable differences in the intensity of the -HI losses. For example, L-Asp generates significantly less -HI loss than the D-isoAsp isomer. These results suggest that differences in the gas phase structures of the two isomers influence the interaction between the initial electron transfer site and the iodobenzyl group. Differences in the ratio of these two products could in theory be used directly for isomer identification, although clearly this strategy would not work for the D-Asp and D-isoAsp isomers in Figure 5.11. Fortunately, differences in the -HI/-I ratio between isomers should translate downstream to the subsequent activation step because the relative abundances of the even/odd-electron peptides will be different, altering the abundance of any products uniquely derived from either precursor. Therefore, these differences will

still be relevant in the subsequent activation step, and the additional possibility for differential fragmentation will be simultaneously explored.

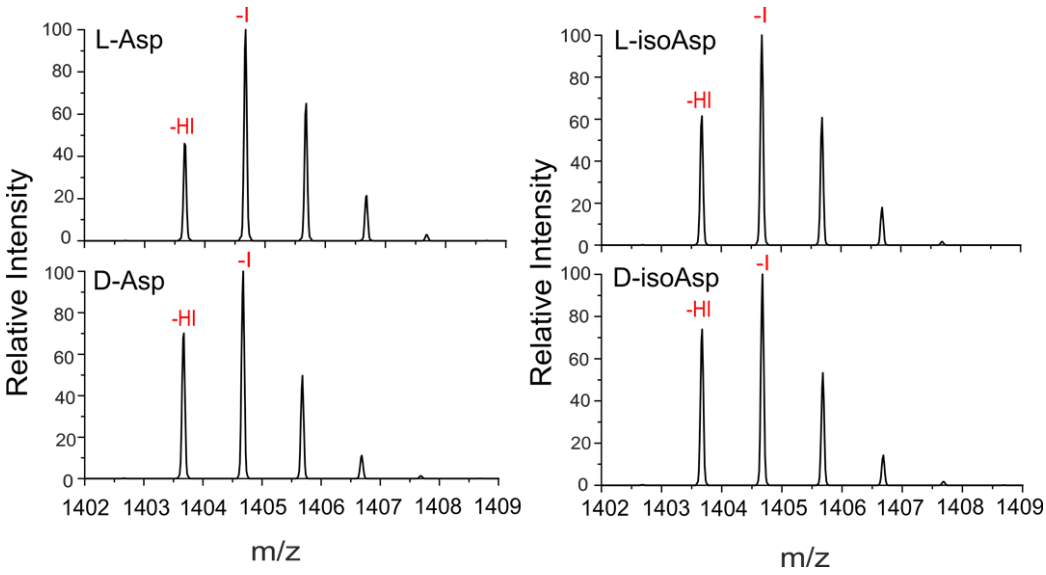


Figure 5.11 –I/HI losses for listed isomers of protonated 4IB-VQEDFVEIHGK. The ratio of products is sensitive to isomer structure.

ETD-CID spectra for L-Asp and D-isoAsp 4IB-VQEDFVEIHGK are shown in Figure 5.12, and the corresponding R_{isomer} values are listed in Table 1. Interestingly, the largest R-value is obtained by comparing the b_9 ion to the -58K-4IB lysine side chain loss. The b_9 ion clearly originates from the even electron peptide (see Figure 5.10b). The side chain loss occurs due to RDD. The L-Asp isomer generates a smaller amount of radical precursor, as illustrated in Figure 5.11a. This likely contributes to the lower abundance of lysine side-chain loss for the L-Asp isomer in Figure 5.12a. The b_9 ion is most sensitive to structural differences between the isomers, but its abundance in Figure 5.12a may be augmented by the higher intensity of nonradical precursor.

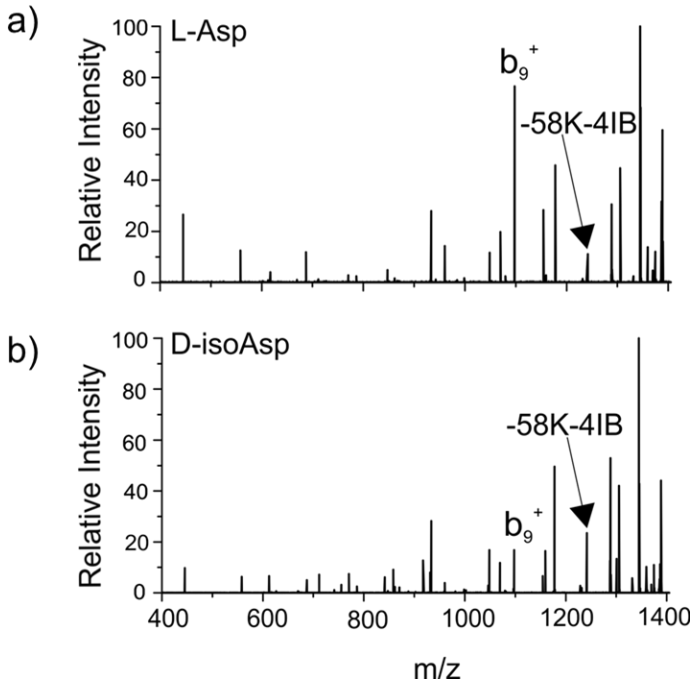


Figure 5.12 ETD-CID spectra for L-Asp and D-isoAsp isomers of protonated 4IB-VQEDFVEIHGK.

The R_{isomer} values obtained from a variety of different activation methods for a several peptides are listed in Table 1. The ETD-CID method is not only competitive, but actually yields the highest value for 2/4 of the examined peptides. The high structural sensitivity derives from the ability of ETD-CID to leverage structural discrimination from both even and odd electron processes simultaneously.

Table 5.1 Resulting R_{isomer} scores from the Comparison of Different Peptide Epimers/Isomers

	R_{isomer}		
Peptide	CID	ETD-CID	PD-CID
[a]VQEDFVEIHGK	8.1	7.0	3.5
[b]VQE*D*FVEIHGK	2.7	6.7	8.7
[c]VQE**D**FVEIHGK	4.6	9.6	3.6
QDEHGFISR	2.0	3.0	1.8
HFSPEDLTVK	2.8	8.7	6.4
IQTGLDATHAER	1.3	2.0	3.0

a) Aspartic acid in red represents L-Asp vs D-Asp
b) Italicized aspartic acid represents L-Asp vs L-isoAsp
c) Bold aspartic acid represents L-Asp vs D-isoAsp

5.4 Conclusions

ETD can be used to activate carbon-iodine bonds in a laser-free experiment, enabling access to RDD-like spectra. Observation of neutral iodine loss strongly indicates excited electronic states not populated directly by the transferred electron can be generated in ETD. Delayed loss of HI suggests that arginine is capable of "storing" the electron temporarily, allowing a variety of ETD-like fragmentation pathways to be accessed following subsequent additional collisional activation. ETD-generated radicals are capable of identifying peptide epimers, at times with greater sensitivity than other methods. The generation of both even- and odd- electron products, which can both be examined simultaneously, is a unique advantage of the ETD-activation method. It is clear that ion-ion reactions and radicals have many yet undiscovered uses that will continue to augment the capabilities of mass spectrometry in years to come.

---

1 McAllister, R. G.; Metwally, H.; Sun, Y.; Konermann, L. Release of Native-like Gaseous Proteins from Electrospray Droplets via the Charged Residue Mechanism:

Insights from Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **2015**, *137* (39), 12667–12676.
2 Zenobi, R.; Knochenmuss, R. Ion Formation in MALDI Mass Spectrometry. *Mass Spectrom. Rev.* **1998**, *17* (5), 337–366.
3 Ly, T.; Julian, R. R. Residue-Specific Radical-Directed Dissociation of Whole Proteins in the Gas Phase. *J. Am. Chem. Soc.* **2008**, *130* (1), 351–358.
4 Song, T.; Hao, Q.; Law, C.-H.; Siu, C.-K.; Chu, I. K. Novel C-Beta-C-Gamma Bond Cleavages of Tryptophan-Containing Peptide Radical Cations. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (2), 264–273.
5 Lee, M.; Kang, M.; Moon, B.; Oh, H. Bin. Gas-Phase Peptide Sequencing by TEMPO-Mediated Radical Generation. *Analyst* **2009**, *134* (8), 1706–1712.
6 Hodyss, R.; Cox, H. A.; Beauchamp, J. L. Bioconjugates for Tunable Peptide Fragmentation: Free Radical Initiated Peptide Sequencing (FRIPS). *J. Am. Chem. Soc.* **2005**, *127* (36), 12436–12437.
7 Gilbert, J. D.; Fisher, C. M.; Bu, J.; Prentice, B. M.; Redwine, J. G.; McLuckey, S. A. Strategies for Generating Peptide Radical Cations via Ion/ion Reactions. *J. Mass Spectrom.* **2015**, *50* (2), 418–426.
8 Zhang, X.; Julian, R. R. Photoinitiated Intramolecular Diradical Cross-Linking of Polyproline Peptides in the Gas Phase. *Phys. Chem. Chem. Phys.* **2012**, *14* (47), 16243.
9 Ly, T.; Julian, R. R. Elucidating the Tertiary Structure of Protein Ions in Vacuo with Site Specific Photoinitiated Radical Reactions. *J. Am. Chem. Soc.* **2010**, *132* (25), 8602–8609.
10 Sun, Q. Y.; Nelson, H.; Ly, T.; Stoltz, B. M.; Julian, R. R. Side Chain Chemistry Mediates Backbone Fragmentation in Hydrogen Deficient Peptide Radicals. *J. Proteome Res.* **2009**, *8* (2), 958–966.
11 Pham, H. T.; Julian, R. R. Radical Delivery and Fragmentation for Structural Analysis of Glycerophospholipids. *Int. J. Mass Spectrom.* **2014**, *370*, 58–65.
12 Tao, Y.; Julian, R. R. Identification of Amino Acid Epimerization and Isomerization in Crystallin Proteins by Tandem LC-MS. *Anal. Chem.* **2014**, *86*, 9733–9741.
13 Ly, T.; Julian, R. R. Tracking Radical Migration in Large Hydrogen Deficient Peptides with Covalent Labels: Facile Movement Does Not Equal Indiscriminate Fragmentation. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (6), 1148–1158.
14 Cheng, P. Y.; Zhong, D.; Zewail, A. H. Kinetic-Energy, Femtosecond Resolved Reaction Dynamics - Modes of Dissociation (in Iodobenzene) from Time-Velocity Correlations. *Chem. Phys. Lett.* **1995**, *237* (5–6), 399–405.
15 Agarwal, A.; Diedrich, J. K.; Julian, R. R. Direct Elucidation of Disulfide Bond Partners Using Ultraviolet Photodissociation Mass Spectrometry. *Anal. Chem.* **2011**, *83* (17), 6455–6458.
16 Ly, T.; Zhang, X.; Sun, Q. Y.; Moore, B.; Tao, Y. Q.; Julian, R. R. Rapid, Quantitative, and Site Specific Synthesis of Biomolecular Radicals from a Simple Photocaged Precursor. *Chem. Commun.* **2011**, *47* (10), 2835–2837.
17 Swaney, D. L.; Wenger, C. D.; Thomson, J. A.; Coon, J. J. Human Embryonic Stem Cell Phosphoproteome Revealed by Electron Transfer Dissociation Tandem Mass Spectrometry. *Proc. Natl. Acad. Sci.* **2009**, *106* (4), 995–1000.

18 Chan, W. Y. K.; Chan, T. W. D.; O'Connor, P. B. Electron Transfer Dissociation with Supplemental Activation to Differentiate Aspartic and Isoaspartic Residues in Doubly Charged Peptide Cations. *J. Am. Soc. Mass Spectrom.* **2010**, *21* (6), 1012–1015.
19 Breuker, K.; Brüschweiler, S.; Tollinger, M. Electrostatic Stabilization of a Native Protein Structure in the Gas Phase. *Angew. Chem. Int. Ed.* **2011**, *50* (4), 873–877.
20 Moore, B. N.; Ly, T.; Julian, R. R. Radical Conversion and Migration in Electron Capture Dissociation. *J. Am. Chem. Soc.* **2011**, *133* (18), 6997–7006.
21 Tureček, F.; Julian, R. R. Peptide Radicals and Cation Radicals in the Gas Phase. *Chem. Rev.* **2013**, *113* (8), 6691–6733.
22 Hood, C. A.; Fuentes, G.; Patel, H.; Page, K.; Menakuru, M.; Park, J. H. *J. Pept. Sci.* **2008**, *14*, 97-101.
23 Tao, W. A., Zhang, D., Nikolaev, E. N. & Cooks, R. G. Copper(II)-assisted enantiomeric analysis of D,L-amino acids using the kinetic method: Chiral recognition and quantification in the gas phase. *J. Am. Chem. Soc.* **2000,** *122*, 10598–10609.
24 Liu, J.; McLuckey, S. A. Electron Transfer Dissociation: Effects of Cation Charge State on Product Partitioning in Ion/ion Electron Transfer to Multiply Protonated Polypeptides. *Int. J. Mass Spectrom.* **2012**, *330–332*, 174–181.
25 Shimamori, H.; Sunagawa, T.; Ogawa, Y.; Tatsumi, Y. "Low-energy electron attachment to C6H5X (X = Cl, Br and I)" *Chem. Phys. Lett.* **1995,** *232*, 115–120.
26 Hendricks, N. G.; Lareau, N. M.; Stow, S. M.; McLean, J. A.; Julian, R. R. Bond-Specific Dissociation Following Excitation Energy Transfer for Distance Constraint Determination in the Gas Phase. *J. Am. Chem. Soc.* **2014**, *136* (38), 13363–13370.
27 Stephenson, J. L.; McLuckey, S. A. Gaseous Protein Cations Are Amphoteric. *J. Am. Chem. Soc.* **1997**, *119* (7), 1688–1696.
28 Bishop, A.; Brodbelt, J. S. Selective Cleavage upon ETD of Peptides Containing Disulfide or Nitrogen–nitrogen Bonds. *Int. J. Mass Spectrom.* **2015**, *378*, 127–133.
29 Truscott, R. J. W.; Schey, K. L.; Friedrich, M. G. Old Proteins in Man: A Field in Its Infancy. *Trends Biochem. Sci.* **2016**, *41* (8), 654–664.
30 Jia, C.; Lietz, C. B.; Yu, Q.; Li, L. Site-Specific Characterization of D-Amino Acid Containing Peptide Epimers by Ion Mobility Spectrometry. *Anal. Chem.* **2014**, *86* (6), 2972–2981.
31 Bai, L.; Romanova, E. V; Sweedler, J. V. Distinguishing Endogenous D-Amino Acid-Containing Neuropeptides in Individual Neurons Using Tandem Mass Spectrometry. *Anal. Chem.* **2011**, *83* (7), 2794–2800.

Chapter 6

Concluding Remarks

Our appreciation for the complexity of long-lived proteins will only continue to grow as we uncover more of them and learn about the mechanisms by which they misfold and induce aggregation. While this field of research is still in its infancy, major advancements have been made the lay the groundwork for future studies. Using mass spectrometry to detect an isobaric modification may seem counterintuitive, but we have shown that radical-directed dissociation can tease out the structural differences induced by isomerization and epimerization. Specifically, the work detailed in Chapter 2 that shows the emergence of Asp repeats being isomerization hotspots. By uncovering new regions that will spontaneously degrade faster than others in the crystallins, we can use this to identify and study similar repeats in other long-lived proteins. The work in Chapter 2 also showed multiple tryptic peptides contained only one Asp residue but more than four isomers. The second site of isomerization was typically pinpointed to a serine if present, but this was not always the case. So, while Asp and Ser might be the most highly-isomerized residues in long-lived proteins, it is apparent they are not the only two. More work will be necessitated to uncover which other amino acids are prone to spontaneous modification as well.

The work in Chapter 3 involved separating the different regions of the lens and studying the degree of isomerization in the nucleus compared to cortex. In doing so, we revealed that the degree of isomerization was significantly higher in the nuclear fractions, which contain the oldest proteins in the lens. Knowing that isomerization typically

increases as a function of age this result may have seem expected, however closer examination of the isomers generated in these two regions began to expose even more details. The increased L-isoAsp content in the nucleus, and subsequent *in vitro* aging studies, revealed that protein isoaspartyl transferase (PIMT) might be actively repairing damaged cortical fiber cells but loses activity in the nucleus. By continuing to study the isomerization patterns of other long-lived proteins where PIMT might be present we can uncover the role that this enzyme has in repairing damage due to isomerization. We also know that different variations of PIMT exist in a variety of organisms and more research should be done to determine if other repair enzymes are present in long-lived proteins.

Dissociation patterns cannot always tell us about the ramifications that these modifications have on the structure and function of proteins. The work in Chapter 4 shows that native mass spectrometry, enzyme assays, and molecular dynamics calculations can all be used to help reveal these details. We know that Asp sidechains are commonly found in salt-bridge interactions, while Ser participates in hydrogen bonding and can be enzymatically phosphorylated. We used this knowledge to probe the crystallins to see if they contained Asp/Ser residues that were participating in any of these interactions. In doing so, we revealed four different sites on the α-crystallins that are isomerized and involved or are near salt-bridges or sites of phosphorylation. These results should facilitate future studies where a similar approach of probing a protein for crucial interactions that define the structure or promote proper oligomerization can be used to flag specific amino acids during isomerization studies.

The work done in Chapter 5 revealed that RDD can be performed without the need of a laser. This is important, as many researchers have access to instruments with ETD capabilities, but not ones equipped with lasers. The work in this section revealed that the differences in -HI/-I from the charge reduced species might be correlated to the distance between the C-I containing chromophore and the initial radical generation site. More studies would need to be done to determine if this is in fact the case before any more conclusions can be drawn.

Together the work in this dissertation shows that long-lived proteins accrue a significant amount of damage over time, and this can have significant consequences on protein structure and function. Extending these techniques to other long-lived proteins will allow for us to uncover the mechanisms by which they unfold, aggregate, and promote disease.