UNIVERSITY OF CALIFORNIA
RIVERSIDE

Shining the Light on Aging and Isomers Using Photodissociation and Radical Chemistry

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

Doctor of Philosophy

in

Biochemistry and Molecular Biology

by

Dylan L. Riggs

June 2019

Dissertation Committee:
Dr. Ryan Julian, Chairperson
Dr. Joseph Genereux
Dr. Jeff Perry

The Dissertation of Dylan L. Riggs is approved:

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

I am fortunate to have been motivated and inspired by wonderful mentors at every step along my scientific journey. To each of you, I owe my most sincere gratitude. Your guidance, your support, and your friendship mean more to me than you could possibly know, and more than I could possibly convey in this brief acknowledgement.

First, I must thank Bill Tanner, who introduced me to the challenges and rewards associated with chemistry in a way that grasped me and never let go. I thank Nancy Deans whose contagious passion guided me toward biochemistry. I thank Paul Schnier for concisely refining my interests toward biotechnology. I thank Blake Gillespie for inviting me into his lab, for sharing his curiosity with me, and for setting an example of how to embrace and appreciate the chemistry of the natural world in our daily lives. I thank Simone Aloisio and Phil Hampton for sculpting my view of instrumental analysis and all things analytical in a way that forever changed my trajectory.

During these formative years as an undergraduate, I was lucky enough to make lasting friendships with many of my amazing peers, and I thank each of you for the fun times and for the meaningful support over the years. Devon Dally, Susan Ly, Lorenzo DeSantiago, and Claudina Cammack, you all welcomed me into your lab many years ago, you shaped my perspective of the research lab setting and

taught me to have fun to stay productive. Joel Sellars and Aurora Ginzburg, I thank you for your insight, for your advice, and for your irreplaceable friendship.

I thank my PhD advisor, Ryan Julian, for taking a chance on a biochemist who surely seemed out of place when I showed up to his lab's group meeting five years ago. I would like to thank Ryan for the guidance, for the feedback, and for the opportunity to grow as a scientist and as a person. I thank my committee members who advised me with sincerity over the years. Joseph Genereux, I thank you for sharing your genuine scientific curiosity and passion; it has been uplifting and infectious. Jeff Perry I thank you for your direct advice, for encouraging me to set well-defined objectives in my research and to outline a path toward my goals. Seán O'Leary, I thank for sharing your unique scientific connections with me, and for your ability to convey the fundamental chemistry that underlies complex biochemical experiments.

I also thank the many excellent researchers in the Julian lab, and I preemptively apologize for grouping most of you into one list. I thank Nicole Pham, Nathan Hendricks, Chris Nellesson, Lance Talbert, James Bonner, Tyler Lambeth, Ting Ting Wu, Jacob Silzel and Jin Tang. I have learned from each of you in the lab and been comforted by your friendships. I also thank Yuanqi Tao for pioneering the research within the Julian lab that paved the way and guided many of my own projects. I thank my personal hero, Ben Moore, for his lasting and unmatched contributions to the Julian lab; my research would not have been possible without the foundation that you built. I thank Guy Quanrud, Macon Abernathy, Taylor

Dennis, and especially Pablo Martinez for their friendship, valuable advice, and for the motivation they have offered. I thank Sonia Gomez for being the world's most uplifting coworker. Your positive attitude was absolutely refreshing, I hope you learned from me some fraction of what I learned from you. I thank Amrik Kang for sharing his gifted mind with me. I owe a most sincere and special thank you to Yana Lyon whose hard work, thoughtful feedback, and leadership helped me more times than I could possibly count. Omar Hamdy, I thank you for all your advice and for your encouraging and wholesome demeanor. Although we overlapped for the shortest duration in the lab, I thank you for continuing to look out for me, and for setting such a positive example for me to follow. I also thank you for introducing me to my most recent mentor, Matt Schenauer. Thank you, Matt, for so quickly sharing with me your admirable perspective on validation and the broader scientific process, and for setting an example of leadership that I hope to one day emulate. More importantly, I thank you for your honesty, for your reflection, for your genuine advice on life and for guiding me during the pivotal moments of my career.

Lastly, and above all else, I thank my family. I thank you for your guidance, for your encouragement, for molding me into the person I am today, and for providing me the opportunity to pursue my dreams. Sharon, I thank you for teaching me to always measure twice and cut once. This has proven useful. Chuck, I thank you for demonstrating that it is conceivably possible to measure just once under the premise that adequate provisions were in place so that Sharon may never find out. Cory, I thank you for a lifetime of loving competition that has

forced me to challenge myself, and for being an exemplary role model. Brandon, I thank you for sharing your creativity with me, and so that Cory does not hold the previous statement over your head. Linda, I thank you for your optimism, and Donald, I thank you balancing out said optimism. Finally, Liisa, I wholeheartedly thank you for your uncanny ability to motivate, to inspire, and, when all else fails, to shepherd me. More importantly, I thank you for your boundless support and for the leadership you provide every day. I thank you for the many amazing experiences and memories we have shared and for being an inexhaustible source of buoyancy and joy in my life.

I thank the following institutions for funding: UC Riverside, the National Science Foundation, the National Institutes of Health, the Wedding family and their generous contribution to the Biochemistry and Molecular Biology department, and the Graduate Division at UCR.

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

Chapter 2: Riggs, D. L., Gomez, S. V., and Julian, R. R. (2017) Sequence and Solution Effects on the Prevalence of D-Isomers Produced by Deamidation. *ACS Chem. Biol. 12*, 2875–2882.

Chapter 3: Riggs, D. L., Hofmann, J., Hahm, H. S., Seeberger, P. H., Pagel, K., and Julian, R. R. (2018) Glycan Isomer Identification Using Ultraviolet Photodissociation Initiated Radical Chemistry. *Anal. Chem. 90*, 11581–11588.

Dedication

To Margie, my endless source of inspiration.

Initial Research Evaluation Committee:

Dr. Ryan Julian
Dr. Joseph Genereux
Dr. Jeff Perry

Qualifying Exam Committee:

Dr. Ryan Julian
Dr. Joseph Genereux
Dr. Jeff Perry
Dr. Seán O'Leary
Dr. Min Xue

Dissertation Committee:

Dr. Ryan Julian, Chairperson
Dr. Joseph Genereux
Dr. Jeff Perry

ABSTRACT OF THE DISSERTATION

Shining the Light on Aging and Isomers Using Photodissociation and Radical Chemistry

by

Dylan L. Riggs

Doctor of Philosophy, Graduate Program in Biochemistry and Molecular Biology
University of California, Riverside, June 2019
Dr. Ryan R Julian, Chairperson

Biological aging is a complex and nuanced chemical process that proceeds along numerous routes at the molecular level. Spontaneous deamidation of asparagine and isomerization of aspartic acid are among the most prevalent age-related chemical modifications and are associated with a growing list of human diseases. Although both degradation pathways are common throughout the body, they are often unnoticed because the resulting chemical modification is relatively minor and is exceptionally difficult to detect. Structural characterization is further complicated because both deamidation and isomerization produce four isomers of aspartic acid

(L-Asp, D-Asp, L-isoAsp, and D-isoAsp). To better understand the aging process, we utilized radical directed dissociation (RDD) in conjunction with mass spectrometry to identify and quantify the products of deamidation and isomerization. We began by outlining intrinsic factors that govern the deamidation rate, and external factors that influence product outcomes to develop models of peptide and protein aging. Subsequent studies revealed the specific structural and functional perturbations associated with the unnatural isomers present in aged proteins. To further expand our isomer detection capabilities, we applied our radical based fragmentation method to the glutamine deamidation, which exhibits several parallels to asparagine deamidation, but has remained largely uncharacterized. Importantly, we demonstrate that radical chemistry generated diagnostic and informative fragment ions for both glutamic acid and isoglutamic acid. We apply this technique in a manner that is amenable to shotgun proteomics and reveal several key differences between the two aging processes. Finally, we tailor our radical based analytical methodology toward the analysis of isomeric glycans. After successfully discriminating a comprehensive family of isomeric glycans with RDD, we demonstrate similar capabilities with 213 nm ultraviolet photodissociation, and outline how such a versatile approach may unify the closely related fields of glycomics and proteomics.

TABLE OF CONTENTS

CHAPTER 1 1
Developing Methods to Explore Molecular Aging 1
Introduction 1
Proteins 2
Isomers 6
Mass Spectrometry 9
Fragmentation 10
Applications of Radical Chemistry 12
References 15
CHAPTER 2 18
Sequence and Solution Effects on the Prevalence of D-isomers Produced by Deamidation 18
Abstract 18
Introduction 19
Results and Discussion 24
Methods 38
Conclusions 40
Acknowledgements 41

Supporting Information....42

References....50

CHAPTER 3 55

Ultraviolet Photodissociation Initiated Radical Chemistry Enables Facile Glycan Isomer Identification....55

Abstract....55

Introduction....56

Experimental Section....62

Results and Discussion....65

Conclusions....76

Acknowledgements....76

Supporting Information....77

References....86

CHAPTER 4 90

Exploration of Glutamine Deamidation: Products, Pathways, and Kinetics 90

Abstract....90

Introduction....92

Experimental Section....96

Results and Discussion....98

Conclusion 112

Acknowledgements 114

Supporting Information 115

References 125

CHAPTER 5 127

Concluding Remarks 127

LIST OF FIGURES

Figure 1.1. Asparagine deamidation and aspartic acid isomerization proceed through the same cyclic succinimide intermediate. Both reactions yield four potential isomeric products (L-Asp, D-Asp, L-isoAsp, and D-isoAsp). 5

Figure 1.2. Skeletal and three-dimensional representations of the four Asp isomers reveal that the structural rearrangements associated with isomerization are not subtle.....7

Figure 1.3. Crystal structures of RNAse U2 as modelled from PDB entries 3AGN and 3AHS. **a)** The native structure contains a deamidation-prone Asn at position 32 in an α-helix. **b)** Following deamidation, the isoAsp generated unwinds the helix and reduces enzyme activity. **c)** Ambiguities in the electron density prevent atomistic characterization within the C-chain of the RNAse U2 crystal. 8

Figure 1.4. Peptide fragment nomenclature. While CID often produces b/y ions, RDD generates additional backbone and sidechain cleavages. ...11

Figure 2.1. (a) LCMS chromatogram of the 4IB-VKLDSG deamidation products which yields only two chromatographic peaks. **(b–e)** RDD tandem mass spectra on [4IB-VKLDSG+H]1+ from the leading and trailing edges of the peaks in **(a)** show differences in fragment intensities that reveal co-eluting isomers within both peaks.....26

Figure 2.2. (a,b) Standards of the L-Asp and D-Asp isomers of 4IB-VKLDSG are distinguished by diagnostic fragment peaks. **c)** The average fragmentation abundance for the LC peak that contains both isomers. **d)** A calibration curve generated from purified standards reveals the LC peak contains 88% L-Asp (red square). 28

Figure 2.3. Deamidation products at 37 °C in 50mM Tris at pH 7.8. Averages are indicated with dashed lines. 29

Figure 2.4. Total product outcomes from 6 different buffer conditions. Clockwise from top left: L-Asp, L-isoAsp, D-isoAsp, D-Asp). Averages are depicted with horizontal bars. Green box indicates ammonia catalysis. The asterisk indicates 80 °C. 35

Figure 2.5. Deamidation rate increases relative to Tris 7.8 for each accelerated buffer condition. The green box indicates ammonia catalysis. The asterisk indicates 80 °C. 37

Figure 2.S1. Typical LCMS chromatograms depicting the separation of the deamidation products for all four peptide sequences in Tris 7.8 at 37 °C.42

Figure 2.S2. Calibration curves for each of the coeluting isomers with the dissociation method indicated. The y-axis is labeled as the difference/sum of the two diagnostic fragment ions observed m/z. For the SG peptide L-isoAsp vs D-isoAsp comparison (f) the linear fit (blue) exhibits systematic error of a few percent at high L-isoAsp percentages – to reduce this error, a polynomial fit (black) is used instead. 45

Figure 2.S3. Peptide structures following 4IB attachment. Note only the histidine sidechain can participate in typical acid/base chemistry in the examined pH range. 47

Figure 2.S4. Figure 2.1 from the main text with additional fragment labels. * indicates the radical precursor produced via photodissociation. The same Figure 2.labels apply: **(a)** LCMS chromatogram of the 4IB-VKLDSG deamidation products which yields only two chromatographic peaks. **(b-e)** RDD tandem mass spectra on $[4IB\text{-}VKLDSG+H]^{1+}$ from the leading and trailing edges of the peaks in (a) show differences in fragment intensities that reveal co-eluting isomers within both peaks. 48

Figure 3.1. A succinct family of glycan isomers that differ in one or more of three structural aspects: connectivity, configuration, and composition. To simplify the structural comparisons, the R- group containing a ß-(1→4) glucose and $C_5H_{10}\text{-}NH_3$ linker that remains consistent between the oligomers is not shown. 58

Figure 3.2. Collisional activation fragmentation spectra (CID and HCD) of isomers **1**-**6** and peak identifications for each oligomer. † indicates a fragment containing a double bond. The $[M+H]^{+1}$ precursor observed at 590 *m/z* is indicated with an arrow..... 66

Figure 3.3. Radical fragmentation spectra afforded by UVPD and RDD. Peak identifications for each oligomer. The dot indicates radical products, the dagger indicates that the product forms a double bond. The $[4IB\ M+Na]^{+1}$ precursor observed at 842 *m/z* during UVPD, and the $-I^{\bullet}$ radical precursor observed at 715 *m/z* during RDD, is indicated with an arrow. 70

Figure 3.4. RDD fragmentation discriminates oligomers based on fragmentation abundances. The dagger indicates that the observed fragment contains a double bond. 73

Figure 3.5. Bond dissociation energies (kJ mol^{-1}) for C-H bonds in galactose. 75

Scheme 3.S1. Radical directed dissociation mechanisms that cleave the glycosidic linker to produce 552 and 554 Da fragments. 78

Scheme 3.S2. The radical directed cross ring cleavage mechanism that produces the 582 Da fragment is affected by the relative orientation of an axial versus equatorial hydrogen. 80

Scheme 4.1. Deamidation of glutamine and isomerization of glutamic acid can both proceed through the glutarimide intermediate, potentially yielding four isomers of glutamic acid. 93

Figure 4.2. RDD fragmentation enables unambiguous identification of Glu and isoGlu containing peptides. Fragments common to all isomers are labelled in black. Fragments that positively identify Glu are labelled in blue, while fragments unique to isoGlu are labelled in red. 102

Scheme 4.2. **(a)** Radical fragmentation pathways depict formation of an a-type and x•-ion pair at the canonical Glu residue following β-hydrogen abstraction. **(b,c)** Lacking a β-hydrogen, isoGlu residues yield mass-shifted a- and x-type ions following α- or γ- hydrogen abstraction. 104

Figure 4.3. **(a)** Incubation of the 4IB-VHLGGQGYK peptide yields 19.6% deamidated products after 176 days. **(b)** Addition of ammonia greatly accelerates the deamidation process, yielding 9.4% deamidation within 14 days. **(c)** RDD confirms that the leading deamidation peak contains the canonical Glu residue, while the latter peak contains the isoGlu product. 106

Figure 4.4. Relative deamidation product ratios following ammonia accelerated deamidation for 4IB-VHLGGQGYK (t=14d), 4IB-VHLGGQAYK (t=127d), 4IB-VHLGGQSYK (t=127d). 108

Figure 4.5. Aging results for 4IB-VHLGGEGYK peptide under acidic conditions. **(a)** Only one peak is observed by LCMS, suggesting that no isoGlu has formed within 61 days. **(b)** Trace amounts of a coeluting species at -18 Dalton, consistent with the glutarimide, is observed during the trailing edge of the main peak. **(c)** RDD fragmentation confirms the presence of the glutarimide intermediate. 112

Figure 4.S1. The -59E and -72E side chain losses associated with glutamic acid in radical directed dissociation arise from α-hydrogen and γ-hydrogen abstraction. These side chain losses are not possible in isoglutamic acid. 115

Figure 4.S2. Full RDD spectra for the 4IB-VHLGGEGYK isomer standards. Label colors match Figure 4.2 in the main text. 116

Figure 4.S3. RDD spectra for the leading and trailing edge of the deamidated product peak generated upon incubation of 4IB-VHLGGQGYK for 127 days. The diagnostic fragments confirm the presence of both Glu and isoGlu. 117

Figure 4.S4. The RDD spectra averaged across the coeluting deamidation product peak for the 4IB-VHLGGQSYK peptide contains both sets of diagnostic fragments, revealing the presence of both Glu and isoGlu. 118

Figure 4.S5. Ammonia accelerated deamidation results for **(a)** 4IB-VHLGGQGYK (t=14d), **(b)** 4IB-VHLGGQAYK (t=127d), **(c)** 4IB-VHLGGQSYK (t=127d). **(d)** A calibration curve based on RDD fragmentation can be used to characterize the coeluting Glu and isoGlu isomers produced by the GQS peptide (red circle). 120

Figure 4.S6. Calibration curves for each diagnostic fragment produce consistent results when evaluating the amount of isoAsp in the GQS deamidated product peak. Avg isoAsp = 17.48%, stdev = 1.45%. 122

Figure 4.S.5. RDD fragmentation patterns for synthetic forms of 4IB-VHLGGEGYK in the L-isoGlu form **(a)** and D-isoGlu form **(b)** with significant differences in the relative abundances of the a_6-14 and -106Y fragments. The RDD fragmentation for the LC peak corresponding to the isoGlu products following deamidation of the 4IB-VHLGGQGYK peptide **(c)**. A calibration curve quantitates the level of L-isoGlu within the LC peak by plotting the difference/sum of the relative abundance for the two diagnostic peaks **(d)**. 123

LIST OF TABLES

Table 3.S1. R-values for each oligomer as obtained by each dissociation method. 85

CHAPTER 1

Developing Methods to Explore Molecular Aging

Introduction

The completion of the Human Genome Project at the turn of the 21st century ushered in a new and exciting era of scientific discovery. As the world's most ambitious collaborative scientific endeavor, the sequencing effort offered an unprecedented level of potential to many previously disconnected fields of study. Along with the great promise, however, came the more nuanced understanding that an organism's genetic blueprint is only one component of the very complicated biochemical processes associated with life.

Unlike DNA, which is characterized by an elegant and highly ordered series of repeating building blocks, lipids, carbohydrates, and proteins can be somewhat more chaotic. Much of our effort has been directed toward making sense of the structural complexity that underlies this chaos. For reasons discussed below, the primary focus of this dissertation is on proteins, although the work in chapter three outlines our recent efforts towards the analysis of carbohydrates, and the methodologies we are developing have shown similar promise in the analysis of lipids.[1]

Proteins

By many metrics, proteins are the most important biomolecules on Earth. Relative to nucleic acids, the original catalysts of life,[2] proteins differ in several key ways that enable drastic improvements in terms of structural variety, chemical specificity, and biological function. Like nucleic acids, proteins are initially constructed as linear polymers from a fixed pool of twenty canonical amino acid building blocks which constitute the primary structure of a protein. This larger pool of building blocks greatly increases the functional permutations relative to DNA and RNA which are limited to four bases each. Additionally, proteins derive complexity from the highly flexible and dynamic nature of secondary structure. While DNA is nearly universally associated with the iconic double helix, protein structure is far more variable. Barring intervention, or with the help of chaperone proteins, nascent polypeptide chains are prone to spontaneous rearrangement to form ‘secondary structures’, with α-helical structures forming as an energetic minimization and ß-sheets forming as a means to avoid entropic penalties in the aqueous environment of the cell. Collectively, these secondary structures associate to define a given proteoform, which can be rigid as is often the case in structural proteins, or highly flexible and dynamic as is often observed in signaling proteins. Perhaps more significantly, proteins are subjected to post-translational modifications (PTMs) as a higher order level of regulation.

Conventionally, post-translational modifications are covalent modifications that fine-tune protein products to a specific subset of cellular functions. The

importance of these modifications is exemplified in the various roles associated with PTMs, which range from cellular communication, signal transduction/signal amplification, cellular localization, to molecular recycling and even cell death.[3,4] Collectively, PTMs allow for a complicated and nuanced cellular communication system, where two protein molecules, derived from the same genetic code, may be subjected to vastly different fates. Because PTMs profoundly affect protein behavior, they are tightly controlled and dysregulation is strongly associated with numerous disease states.[5] As such, a considerable portion of the proteome is known to be associated with at least one PTM.[6] As the tools for analyzing the proteome (namely, mass spectrometry) have become more rapid, more sensitive, and more detailed, we have begun to identify PTMs at a rate that far exceeds our understanding of PTMs.[7] Nonetheless, it has become abundantly clear that protein structure is closely tied to function, that small perturbations to structure can have disproportionately large impacts on protein behavior, and that the proteome is simultaneously robust enough to withstand dozens of genetically derived mutations from generation-to-generation, but also so fragile that a single amino acid substitution can be fatal.[8,9]

Among these intended and highly regulated post-translational modifications are a collection of spontaneous chemical degradations such as isomerization and deamidation that compromise protein integrity over time. These decay pathways arise from the inherent instability of amino acids, primarily at asparagine, aspartic acid, and serine, and are compounded by the harsh chemical environment within

the cell.[10,11] Unlike traditional PTMs, age-related decay is often spontaneous and non-enzymatic; as a result, these degradation pathways are unregulated and typically irreversible. Because these deterioration pathways operate at rates slower than typical protein turnover,[10] they are often overlooked and remain poorly characterized. We have begun investigating amino acid isomers generated from several distinct pathways to better understand the structural and functional consequences of these age-related decomposition products.

Isomerization, which refers to the rearrangements of atoms or bonds within a molecule, is in some respects the smallest chemical change that a protein can undergo. With that subtlety comes an assumption of insignificance, however there is a growing body of evidence that highlights the importance of isomeric integrity. Deamidation, like isomerization, is a relatively minor change, and one that is often grouped with isomerization because the products of both reactions are isomeric amino acids. Both deamidation of asparagine and isomerization of aspartic acid proceed as the backbone nitrogen of the C-terminal neighboring amino acid attack the side chain of asparagine or aspartic acid to yield a cyclic succinimide intermediate (**Figure 1.1**).

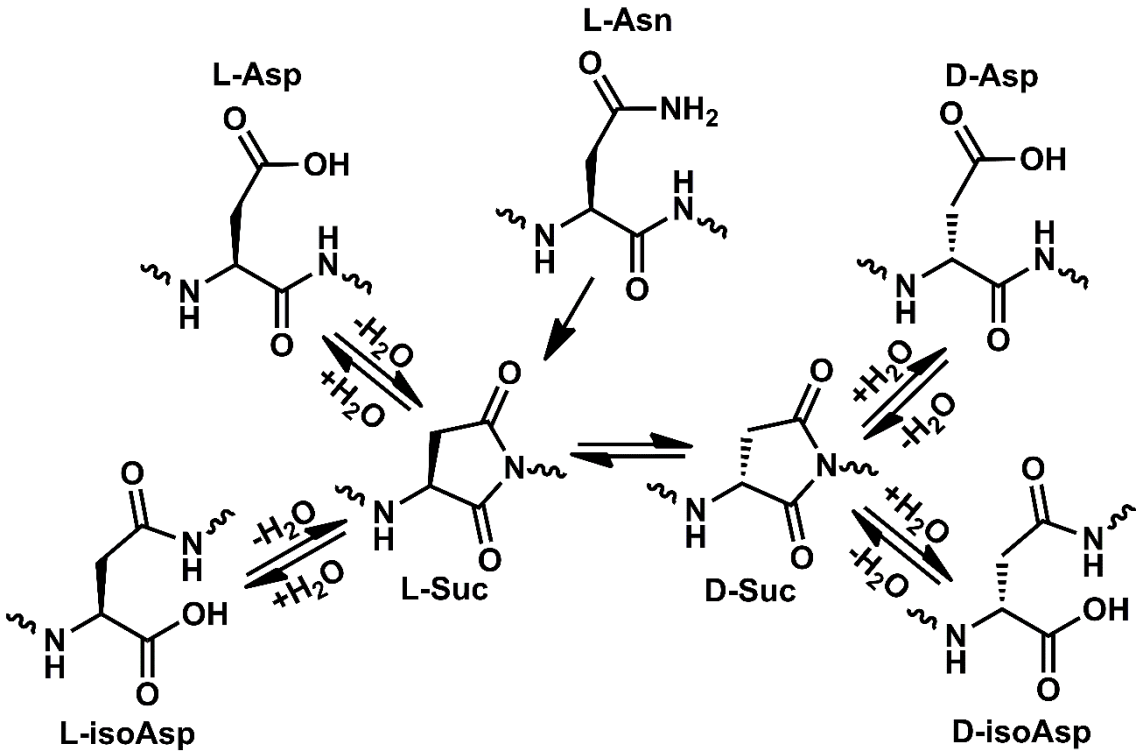


Figure 1.1. Asparagine deamidation and aspartic acid isomerization proceed through the same cyclic succinimide intermediate. Both reactions yield four potential isomeric products (L-Asp, D-Asp, L-isoAsp, and D-isoAsp).

The succinimide-mediated degradation pathway can be accessed by aspartic acid directly or through asparagine which generates the succinimide following deamidation. Interestingly, the succinimide ring is formed over an order of magnitude faster via deamidation than through direct aspartic acid cyclization; therefore, deamidation of asparagine is a major source of aspartic acid isomers in the body.[12] Because asparagine enters the succinimide ring more quickly, and is generally easier to detect, the kinetics of deamidation have been most thoroughly explored.[8] The role of pH,[13] temperature,[14] primary structure,[15] and secondary

structure[16] in determining the deamidation rate has been investigated, and it is clear that backbone flexibility, backbone amide acidity, and side chain sterics all govern the rate of the reaction.[17] Importantly, the trends observed for deamidation rates and behaviors have proven useful in predicting isomerization results, and the two reactions share many parallels.[18] Nonetheless, the isomeric products remain exceptionally difficult to characterize for both deamidation and isomerization.[19, 20]

Isomers

Regardless of the origin, the four isomeric products of aspartic acid are known to exhibit significant impacts on protein structure and function and are often associated with drastic biological effects. For instance, in addition to triggering autoimmune responses,[21] isoaspartic acid is widely associated with protein aggregation.[22,23] This association is most notable in the lens crystallin proteins where isomerization is thought to seed cataract formation[24] and in β-amyloid peptides which are linked to Alzheimer's Disease.[25,26] Recently, we have also revealed that isomerization can severely attenuate enzyme activity, both for kinases and for proteases.[27,28] While isomerization is broadly considered to be a subtle, or minor change, a close look at the specific atomic rearrangement localized to the isomerization site reveals how this 'minor' change can have major ramifications. **Figure 1.2** depicts the four isomers of aspartic acid, aligned along the peptide backbone. Notably, the side-chain inversion associated with L- to D-chiral inversion is evident, but perhaps more strikingly, the dramatic backbone kink

associated with isoAsp formation is also revealed. From this perspective, it is more apparent why, and how, isomerization exhibits such damaging effects within the cell.

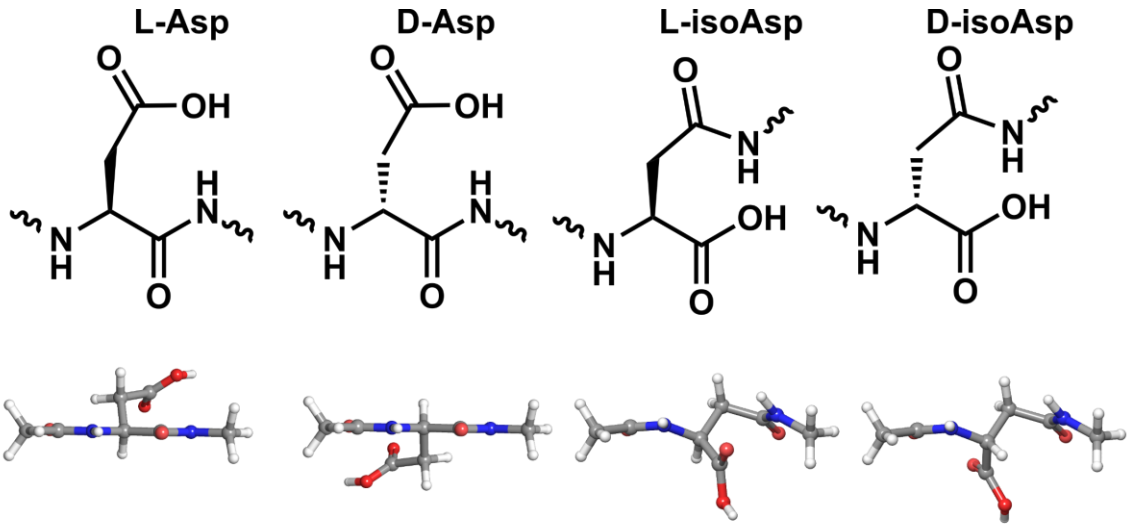


Figure 1.2. Skeletal and three-dimensional representations of the four Asp isomers reveal that the structural rearrangements associated with isomerization are not subtle.

RNase U2 represents a particularly informative example of how the relatively minor structural changes at an individual residue can have pronounced structural and functional impacts on the whole protein. RNAse U2 has long been known to lose activity as it deamidates.[29] More recently, x-ray crystal structures of the native and deamidated enzymes have detailed the specific structural modification responsible for the loss of enzyme activity. Following deamidation, the native α-helix is partially unfolded as L-isoAsp is generated (**Figure 1.3**).[30] While these crystal structures provide an excellent example of the impacts of this age-

related degradation, they also demonstrate the difficulty in studying isomers. In the PDB deposition, the RNAse U2 chain C is missing electron density from Asp29 to Asp34, due to 'disorder in the region' that 'led to ambiguous electron densities'.[30] One likely explanation for this observation is isomeric impurities within the crystallized material, (ie some inseparable combination of L-Asp, D-Asp, L-isoAsp, and D-isoAsp).

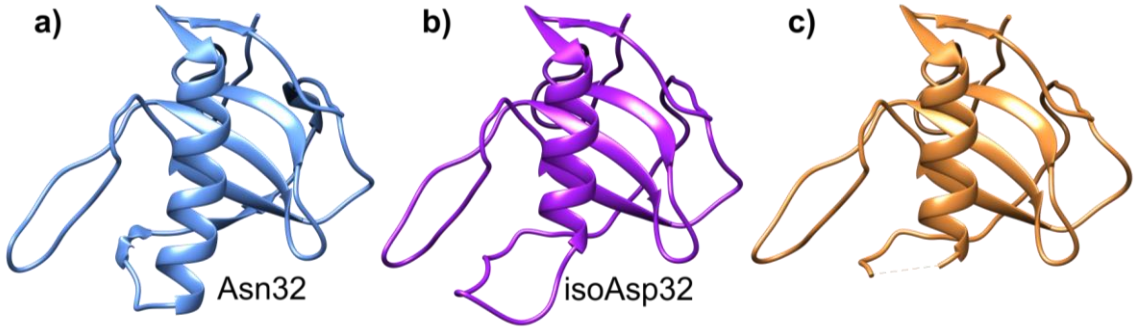


Figure 1.3. Crystal structures of RNAse U2 as modelled from PDB entries 3AGN and 3AHS. **a)** The native structure contains a deamidation-prone Asn at position 32 in an α-helix. **b)** Following deamidation, the isoAsp generated unwinds the helix and reduces enzyme activity. **c)** Ambiguities in the electron density prevent atomistic characterization within the C-chain of the RNAse U2 crystal.

Interestingly, the powerful effects of unnatural amino acids have long been utilized by nature. For instance, tree frogs (*Phyllomedusa sauvagei*) secrete the D-alanine form of the opioid dermorphin (YAFGYPS) through their skin in what is likely a defense mechanism. This D- isomer containing peptide is a powerful opioid, ~1,000 times more potent than morphine, while the L-alanine version is

completely benign.[31] Similar D-amino acid containing peptides have been observed across a wide range of life on earth including antibiotic peptides, platypi,[32] spiders, and numerous marine organisms.[33] While individual D- amino acids are known neurotransmitters,[34] D- amino acids within a human protein are unnatural, and are exclusively associated with protein damage and aging. We are primarily interested in understanding deamidation and isomerization as it arises through natural aging processes and have been developing mass spectrometry-based approaches to aid in this endeavor.

Mass Spectrometry

Electrospray ionization is a wonderful and convenient approach to introducing non-volatile biological molecules into the gas phase. One particularly impactful application of electrospray ionization is at the interface of liquid chromatography and mass spectrometry (LCMS). This, now seemingly ubiquitous union of two powerful techniques has become the preferred workhorse of proteomics and revolutionized the experimental approach to asking “what’s in this cell?”.

By separating analytes in time and delivering the output as a liquid stream ready to be ionized, liquid chromatography is an ideal preface to mass spectrometry, and one that interfaces seamlessly. Analytes (namely digested proteins) can be successively delivered to the mass spectrometer, rapidly

analyzed whole, before being dissociated and then sequenced based on fragmentation patterns. With great regularity, new advances in LCMS are accelerating the speed, depth, accuracy, and quantitation of protein and post-translational modification identification. For instance, in 2014 the Coon lab reported comprehensive coverage of the yeast proteome (~4,000 proteins) in just one hour, a state-of-the-art accomplishment,[35] but still had the foresight to predict an impending identification rate doubling within a few years. This identification doubling was recently realized by the Mann and Cox labs (albeit through different than predicted means) who reported a single-shot proteomics experiment detecting 10,000 proteins in just 100 minutes.[36] Given this flux of proteomic information achievable with a modern LCMS setup, it is no surprise that our ability to catalog proteomic information far surpasses our ability to understand it.[7]

Fragmentation

LCMS based proteomic studies are founded on peptide fragmentation. Conventionally, collision induced dissociation (CID) is employed to fragment polypeptides along the peptide backbone, ideally yielding consecutive losses of whole amino acids at every position (**Figure 1.4**).

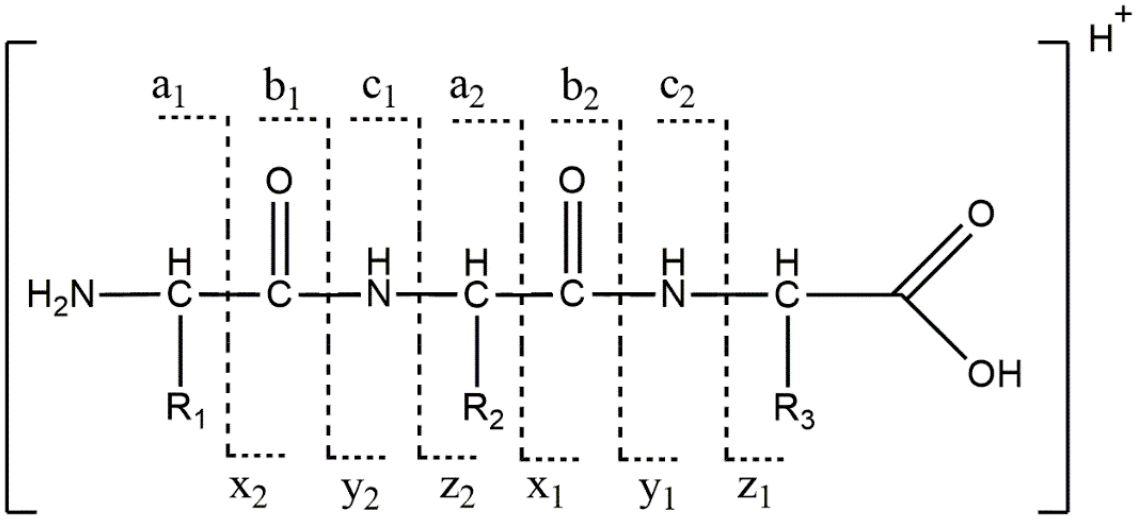


Figure 1.4. Peptide fragment nomenclature. While CID often produces b/y ions, RDD generates additional backbone and sidechain cleavages.

Following fragmentation, the peptide sequence can be deduced from the differences in fragment masses, and the finite number of permutations of amino acids that can yield the observed intact mass. While this approach has numerous advantages over conventional biochemical approaches to protein identification, it has several notable pitfalls as well. Namely, isomers are, by definition, identical in mass. Most proteomic studies disregard isomers entirely, with the exception of leucine and isoleucine which can be determined from the associated protein's gene if needed. However, long-lived proteins present a notable exception. As the biological lifetime of a protein increases, spontaneous age-related modifications such as isomerization and deamidation become almost inevitable. For instance, racemization, a form of isomerization where only chirality is inverted, has been

found to be the most abundant PTM in long-lived proteins,[37] but is completely invisible to conventional proteomic studies.

To identify and distinguish the products of isomerization and deamidation, the Julian lab has developed a method for isomer proteomics which is founded on radical directed dissociation (RDD). Unlike traditional dissociation techniques, RDD utilizes a covalent modifier (*para*-iodobenzoic acid, or 4IB) which is photolabile. Modified analytes are sprayed into a modified ion trap mass spectrometer and isolated. Upon laser irradiation of the ion cloud, the carbon-iodine bond in 4IB is homolytically cleaved producing a radical at a fixed location on the analyte which leads to a series of radical-mediated bond cleavages through hydrogen abstraction.[38,39] Because radical migration is structurally sensitive, even subtle differences between isomers leads to marked differences in fragmentation. RDD has proven to be exceptionally capable of discriminating isomeric peptides,[40] and has been instrumental in identifying novel sites of isomerization, including sequence and structurally specific hotspots in the lens proteins. [18,41,42]

Applications of Radical Chemistry

Radical directed dissociation is utilized in chapter two to systematically characterize the products of deamidation in a series of model peptides. In these foundational studies, we explored the influence of external and internal factors on deamidation rates and isomer production. Importantly, we discovered that the

peptide sequence and buffer condition bear little influence on the deamidation product profile. However, we also identified several factors that promote racemization. In addition to providing important insight into the deamidation process, these experiments established the utility of RDD based calibration curves, which can be used to quantify isomers in coeluting LC peaks based on relatively minor differences in fragmentation. This was an important development which established that quantitation was possible even in the absence of unique fragments which were previously demonstrated by Yuanqi Tao.[40]

This capability was further utilized in chapter four to characterize the isomeric products arising from the very similar glutamine deamidation pathway. Comprehensive studies in this area have remained exceptionally elusive and difficult to carryout due to the very long timescale associated with glutamine deamidation. For instance, while the most rapid asparagine deamidation sequences (GNG) have half-lives of ~1 day, the corresponding glutamine deamidation (GQG) has a half-life exceeding 1 year.[43] Nonetheless, we were inspired by observations of Glu deamidation in the lens, and recent reports of Glu deamidation within the brain[44] to apply our approach toward characterizing this aging process. Our initial efforts were focused on utilizing RDD to discriminate isomers glutamic acid. Based on previous mechanistic studies on RDD,[38] we suspected that the structural rearrangement associated with isoGlu would suppress the side chain losses associated with canonical Glu. In fact, the side chain losses were found to be unique to canonical glue residues and therefore

diagnostic, but perhaps more importantly we identified additional fragmentation channels that positively identify isoGlu. The utility of this approach allowed us to investigate the sequence effects on glutamine deamidation, identify several parallels between asparagine deamidation and glutamine deamidation, and further study glutamic acid isomerization.

Ultimately, mass spectrometry and radical directed dissociation has proven an invaluable tool at every step toward the detection and characterization of age-related protein damage. Our work began with model peptide systems and followed a natural progression to full protein analysis and protein-substrate interactions before recently culminating with live cell studies. And while our attention has primarily focused on proteins, it must be mentioned that a complete biological and cellular understanding would be incomplete without giving merit to the most ubiquitous biomolecules on Earth, carbohydrates. In collaboration with the Pagel lab at the Fritz Haber Institute of the Max Planck Society, we demonstrate in chapter three that radical chemistry is particularly useful for characterizing isomeric carbohydrates which are known to be exceptionally challenging to analyze. Importantly, this work was performed in parallel using our conventional radical direction dissociation platform, as well as an emerging 213 nm photodissociation method which employs a commercially available laser for radical generation and should therefore be accessible to a broader range of chemists and biologists interested in studying these challenging, but important, chemical modifications.

References

[1] Pham, H. T., and Julian, R. R. (2016) Characterization of glycosphingolipid epimers by radical-directed dissociation mass spectrometry. *Analyst 141*, 1273–1278.

[2] Alberts B, Johnson A, Lewis J, et al. Molecular Biology of the Cell. 4th edition. New York: Garland Science; 2002. The RNA World and the Origins of Life.

[3] Roth, J.; Zuber, C.; Park, S.; Jang, I.; Lee, Y.; Kysela, K. G.; Le Fourn, V.; Santimaria, R.; Guhl, B.; Cho, J. W. Protein N-Glycosylation, Protein Folding, and Protein Quality Control. *Mol. Cells* **2010**, *30* (6), 497–506.

[4] Komander, D., and Rape, M. (2012) The Ubiquitin Code. *Annu. Rev. Biochem. 81*, 203–229.

[5] Huang, Q.; Chang, J.; Cheung, M. K.; Nong, W.; Li, L.; Lee, M.; Kwan, H. S. Human Proteins with Target Sites of Multiple Post-Translational Modification Types Are More Prone to Be Involved in Disease. *J. Proteome Res.* **2014**, *13* (6), 2735–2748.

[6] Khoury, G. A., Baliban, R. C., and Floudas, C. A. (2011) Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci. Rep. 1*, 90.

[7] Naegle, K. M., Gymrek, M., Joughin, B. a, Wagner, J. P., Welsch, R. E., Yaffe, M. B., Lauffenburger, D. a, and White, F. M. (2010) PTMScout, a Web resource for analysis of high throughput post-translational proteomics studies. *Mol. Cell. Proteomics 9*, 2558–2570.

[8] Veltman, J. A., and Brunner, H. G. (2012) De novo mutations in human genetic disease. *Nat. Rev. Genet. 13*, 565–575.

[9] Taniike, M., Fukushima, H., Yanagihara, I., Tsukamoto, H., Tanaka, J., Fujimura, H., Nagai, T., Sano, T., Yamaoka, K., Inui, K., and Okada, S. (1992) Mitochondrial tRNAlle mutation in fatal cardiomyopathy. *Biochem. Biophys. Res. Commun. 186*, 47–53.

[10] 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.

[11] Fujii, N.; Sakaue, H.; Sasaki, H.; Fujii, N. A Rapid, Comprehensive Liquid Chromatography-Mass Spectrometry (LC-MS)-Based Survey of the Asp Isomers in Crystallins from Human Cataract Lenses. *J. Biol. Chem.* **2012**, *287* (47), 39992–40002.

[12] Stephenson, R. C.; Clarke, S. Succinimide Formation from Aspartyl and Asparaginyl Peptides as a Model for the Spontaneous Degradation of Proteins. *J. Biol. Chem.* **1989**, *264* (11), 6164–6170.

[13] Patel, K.; Borchardt, R. T. Chemical Pathways of Peptide Degradation. III. Effect of Primary Sequence on the Pathways of Deamidation of Asparaginyl Residues in Hexapeptides. *Pharmaceutical Research: An Official Journal of the American Association of Pharmaceutical Scientists*. 1990, pp 787–793.

[14] Yüksel, K. Ü.; Gracy, R. W. In Vitro Deamidation of Human Triosephosphate Isomerase. *Arch. Biochem. Biophys.* **1986**, *248* (2), 452–459.
[15] Robinson, N. E.; Robinson, A. B. Molecular Clocks. *Proc. Natl. Acad. Sci.* **2001**, *98* (3), 944–949.
[16] Kosky, A.; Razzaq, U. O.; Treuheit, M. J.; Brems, D. N. The Effects of Alpha-Helix on the Stability of Asn Residues: Deamidation Rates in Peptides of Varying Helicity. *Protein Sci.* **1999**, *8*, 2519–2523.
[17] Radkiewicz, J. L.; Zipse, H.; Clarke, S.; Houk, K. N. Neighboring Side Chain Effects on Asparaginyl and Aspartyl Degradation: An Ab Initio Study of the Relationship between Peptide Conformation and Backbone NH Acidity. *J. Am. Chem. Soc.* **2001**, *123* (15), 3499–3506.
[18] Lyon, Y. A., Sabbah, G. M., and Julian, R. R. (2018) Differences in α-Crystallin isomerization reveal the activity of protein isoaspartyl methyltransferase (PIMT) in the nucleus and cortex of human lenses. *Exp. Eye Res.*
[19] Yang, H.; Fung, E. Y. M.; Zubarev, A. R.; Zubarev, R. A. *J. Proteome Res.* **2009**, *8* (10), 4615–4621.
[20] Lampi, K. J.; Wilmarth, P. A.; Murray, M. R.; David, L. L. *Prog. Biophys. Mol. Biol.* **2014**, *115* (1), 21–31.
[21] Doyle, H. A.; Mamula, M. J. Post-Translational Protein Modifications in Antigen Recognition and Autoimmunity. *Trends Immunol.* **2001**, *22* (8), 443–449.
[22] Hooi, M. Y. S.; Raftery, M. J.; Truscott, R. J. W. Interconversion of the Peptide Isoforms of Aspartate: Stability of Isoaspartates. *Mech. Ageing Dev.* **2013**, *134* (3–4), 103–109.
[23] Shimizu, T.; Matsuoka, Y.; Shirasawa, T. D -Amino Acid Biosystem Biological Significance of Isoaspartate and Its Repair System. *Biol. Pharm. Bull.* **2005**, *28* (9), 1590–1596.
[24] Truscott, R. J. W.; Friedrich, M. G. The Etiology of Human Age-Related Cataract. Proteins Don't Last Forever. *Biochim. Biophys. Acta* **2016**, *1860* (1 Pt B), 192–198.
[25] Shimizu, T.; Matsuoka, Y.; Shirasawa, T. D -Amino Acid Biosystem Biological Significance of Isoaspartate and Its Repair System. *Biol. Pharm. Bull.* **2005**, *28* (9), 1590–1596.
[26] Kozin, S. A.; Mitkevich, V. A.; Makarov, A. A. Amyloid-β Containing Isoaspartate 7 as Potential Biomarker and Drug Target in Alzheimer's Disease. *Mendeleev Commun.* **2016**, *26* (4), 269–275.
[27] Lyon, Y. A., Collier, M. P., Riggs, D. L., Degiacomi, M. T., Benesch, J. L. P., and Julian, R. R. (2019) Structural and functional consequences of age-related isomerization in α-crystallins. J. Biol. Chem.
[28] Lambeth, T. R., Riggs, D. L., Talbert, L. E., Tang, J., Coburn, E., Kang, A. S., Noll, J., Augello, C., Ford, B. D., and Julian, R. R. (2019) Spontaneous isomerization of long-lived proteins provides a molecular mechanism for the lysosomal failure observed in Alzheimer's disease. *bioRxiv* 605626.
[29] Uchida, T.; Shibata, Y. *J. Biochem.* **1981**, *90* (2), 463–471.
[30] Noguchi, S. *Biopolymers* **2010**, *93* (11), 1003–1010.

[31] Jilek, A., Kreil, G. (2008) D-Amino Acids in Animal Peptides, *Monatshefte für Chemie*. *139*, 1–5.
[32] Torres, A. M., Tsampazi, C., Geraghty, D. P., Bansal, P. S., Alewood, P. F., and Kuchel, P. W. (2006) D-Amino acid residue in a defensin-like peptide from platypus venom: effect on structure and chromatographic properties. *Biochem. J. 391*, 215–220.
[33] Bai, L., Sheeley, S., and Sweedler, J. V. (2009) Analysis of Endogenous D-Amino Acid-Containing Peptides in Metazoa. *Bioanal Rev. 1*, 7–24.
[34] Durrant, A. R., and Heresco-Levy, U. (2014) D-Serine in Neuropsychiatric Disorders: New Advances. *Adv. Psychiatry. 2014*, 1–16.
[35] Hebert, A. S., Richards, A. L., Bailey, D. J., Ulbrich, A., Coughlin, E. E., Westphall, M. S., and Coon, J. J. (2014) The One Hour Yeast Proteome. *Mol. Cell. Proteomics 13*, 339–347.
[36] Meier, F., Geyer, P. E., Virreira Winter, S., Cox, J., and Mann, M. (2018) BoxCar acquisition method enables single-shot proteomics at a depth of 10,000 proteins in 100 minutes. *Nat. Methods 15*, 440–448.
[37] Hooi, M. Y. S., Raftery, M. J., and Truscott, R. J. W. (2013) Age-dependent racemization of serine residues in a human chaperone protein. *Protein Sci. 22*, 93–100.
[38] Sun, Q., Nelson, H., Ly, T., Stoltz, B. M., and Julian, R. R. (2009) Side chain chemistry mediates backbone fragmentation in hydrogen deficient peptide radicals. *J. Proteome Res. 8*, 958–966.
[39] Moore, B. N., and Julian, R. R. (2012) Dissociation energies of X–H bonds in amino acids. *Phys. Chem. Chem. Phys. 14*, 3148.
[40] Tao, Y., Quebbemann, N. R., and Julian, R. R. (2012) Discriminating d -Amino Acid-Containing Peptide Epimers by Radical-Directed Dissociation Mass Spectrometry. *Anal. Chem. 84*, 6814–6820.
[41] Tao, Y., and Julian, R. R. (2014) Identification of Amino Acid Epimerization and Isomerization in Crystallin Proteins by Tandem LC-MS. *Anal. Chem. 86*, 9733–9741.
[42] Lyon, Y. A., Sabbah, G. M., and Julian, R. R. (2017) Identification of Sequence Similarities among Isomerization Hotspots in Crystallin Proteins. *J. Proteome Res. 16*, 1797–1805.
[43] Robinson, N. E., Robinson, Z. W., Robinson, B. R., Robinson, A. L., Robinson, J. A., Robinson, M. L., and Robinson, A. B. (2004) Structure-dependent nonenzymatic deamidation of glutaminyl and asparaginyl pentapeptides. *J. Pept. Res. 63*, 426–436.
[44] Serra, A., Gallart-Palau, X., Wei, J., and Sze, S. K. (2016) Characterization of Glutamine Deamidation by Long-Length Electrostatic Repulsion-Hydrophilic Interaction Chromatography-Tandem Mass Spectrometry (LERLIC-MS/MS) in Shotgun Proteomics. *Anal. Chem. 88*, 10573–10582.

CHAPTER 2

Sequence and Solution Effects on the Prevalence of D-isomers Produced by Deamidation

Abstract

Deamidation of asparagine is a spontaneous and irreversible post-translational modification associated with a growing list of human diseases. While pervasive, deamidation is often overlooked because it represents a relatively minor chemical change. Structural and functional characterization of this modification is complicated because deamidation of asparagine yields four isomeric forms of Asp. Herein, radical directed dissociation (RDD), in conjunction with mass spectrometry, is used to identify and quantify all four isomers in a series of model peptides that were subjected to various deamidation conditions. Although primary sequence significantly influences the rate of deamidation, it has little impact on the relative proportions of the product isomers. Furthermore, the addition of ammonia can be used to increase the rate of deamidation without significantly perturbing isomer populations. Conversely, external factors such as buffer conditions and temperature alter product distributions but exhibit less dramatic effects on the deamidation rate. Strikingly, the common laboratory and biologically significant bicarbonate buffer is found to strongly promote racemization, yielding increased amounts of D-Asp and D-isoAsp. These outcomes following deamidation have broad implications in human aging and should be considered during the development of protein-based therapeutics.

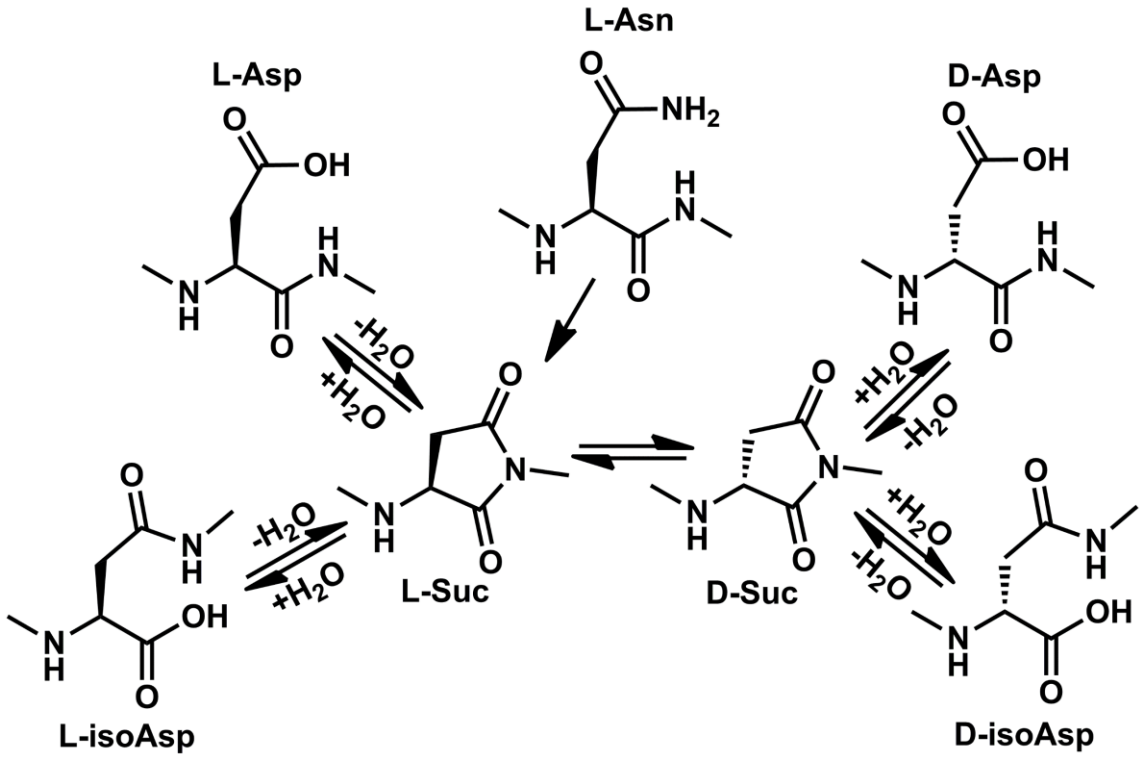
Introduction

Post-translational modifications (PTMs) behave as powerful regulatory elements that often define protein function through covalent modification. Because PTMs profoundly affect protein behavior, they are typically tightly controlled, and dysregulation is strongly associated with numerous diseases.[1] Deamidation is an exception among PTMs due to its non-enzymatic and spontaneous nature. Consequently, it is both unregulated and irreversible. The inevitable nature and predictable rate of this modification prompted the idea that deamidation may have been engineered as a molecular clock that signals protein aging and turnover.[2] Although there are select examples of deamidation serving functional roles, it is generally considered to represent degradation[3] and is associated with a growing list of age-related, neurological, aggregation-prone, and autoimmune diseases.[4]

While the implications of asparagine deamidation are becoming increasingly apparent, the specific structural outcomes have not been well-defined because the reaction yields four isomeric products: L-Asp, D-Asp, L-isoAsp, and D-isoAsp (see **Scheme 2.1**). All deamidation pathways introduce a negatively charged side chain under physiological conditions, but each of the four isomers is additionally likely to impact structure and function in distinct ways. Although L-Asp and D-Asp differ by only a single chiral center (thus making peptide epimers when present in pairs that otherwise have the same sequence), D-amino acids are known to have a profound impact on peptide function. For instance, the D-Ala epimer of the peptide opioid

dermorphin (YAFGYPS) is ~1,000 times more potent than morphine while the L-Ala epimer is biologically inactive.[5] Unlike dermorphin, which is enzymatically modified in tree frogs (*Phyllomedusa sauvagei*), D-amino acids accumulate in human proteins due to aging. Recently, racemization was reported to be the most abundant PTM in aged lens proteins.[6] Furthermore, D-Ser epimers have been shown to contribute to the toxicity of ß-amyloid *in vivo.*[7] In addition to yielding racemized Asp, deamidation produces isoAsp, which incorporates a methylene group into the peptide backbone (**Scheme 2.1**). Isoaspartic acids are associated with protein aggregation,[8,9] most notably in the lens crystallin proteins where isomerization may induce aggregation[10] and promote cataract formation[11,12] and in ß-amyloid peptides linked to Alzheimer's Disease.[8,13,14] Additionally, isoaspartic acids are known to trigger autoimmune responses,[15] which may be involved in diseases such as multiple sclerosis.[16] Importantly, the B cell response triggered by isoAsp peptides from cytochrome c has been reported to be cross-reactive, targeting both the isoAsp peptides and the native protein in mice.[17] Low levels of deamidation may, therefore, trigger the removal of the entire protein population.

Scheme 2.1 Deamidation proceeds via an intramolecular nucleophilic attack. The succinimide (Suc) intermediate is prone to racemization and hydrolysis yielding four Asp isomers.



As a spontaneous process, the importance of deamidation extends outside the cell. Deamidation is a source of heterogeneity in recombinant protein production.[18] These structural changes have been shown to alter protein folding transitions[19] and may contribute to protein aggregation.[20] Degradation can continue during storage,[4] making deamidation an important consideration for therapeutic formulations.[21] Protein-based bioconjugates represent an increasingly popular choice for pharmaceuticals,[22,23] accentuating the need for better understanding of effective protein lifetime and degradation byproducts. For example, perturbations to complementarity-determining regions (CDRs) of antibodies can drastically reduce antigen binding efficacy,[24] as demonstrated by deactivation of monoclonal

immunoglobulin γ2 following a single isomerization.[25] Furthermore, deamidation is facilitated by the solvent exposed and flexible nature of CDRs.

Under physiological conditions, deamidation proceeds through a succinimide intermediate that is formed as the backbone nitrogen (on the C-terminal side) attacks the side chain carboxyl group of asparagine[26] (**Scheme 2.1**). Therefore, peptide flexibility[27] and the amino acid on the C-terminal side of asparagine strongly influence the deamidation rate.[2] Catalysis by ammonia and increased temperature can accelerate deamidation.[28] Once formed, the asymmetric succinimide ring can be hydrolyzed to yield Asp or isoAsp with the isoAsp products being favored at a ratio of about 3:1.[26] The ring is prone to racemization prior to hydrolysis, and the resulting products include both the L- and D- stereoisomers of Asp and isoAsp. Asp residues are also susceptible to succinimide ring formation, ultimately yielding the same four products, but this reaction is significantly slower.

Despite the ease in detecting deamidation (which results in a 0.984 Da mass shift) via mass spectrometry (MS), high-throughput analysis of the four isomers generated by deamidation remains challenging.[29] Most chromatographic methods are limited to reporting the ratio of the Asp/isoAsp, although some isomeric peptides have been fully resolved via ultra-performance liquid chromatography (UPLC).[12] Additionally, ion mobility MS (IMS-MS) has been successfully employed to characterize peptide epimers[30] and ß-amyloid isomers.[31] Mass spectrometry-based dissociation techniques can be used to detect isoAsp residues based on diagnostic c- and z- fragment mass shifts generated by electron capture

dissociation (ECD),[32] and electron transfer dissociation (ETD).[33] Like ECD and ETD, radical directed dissociation (RDD) is a gas-phase radical-based dissociation technique, but RDD differs by generating radicals via site-specific photodissociation.[34] Radical migration facilitates cleavage of the peptide backbone as well as amino acid sidechains.[35] Importantly, these radical migration pathways are structurally sensitive.[36] RDD has previously been used to distinguish peptide epimers based on fragmentation patterns[37] and identify Asp isomers via tandem LC-MS.[38] Despite these recent advances, there are few cases where deamidation has been systematically studied, likely due to time constraints and difficulty with distinguishing and quantifying the products.

Presented herein is a systematic evaluation of the four Asp isomers generated via deamidation, with a focus on the influence of primary peptide sequence, side chain chemistry, and solvent conditions. Short peptides were selected to test deamidation in the absence of confounding factors such as higher-order structure. We analyzed a series of hexapeptides that vary in sequence by a single amino acid C-terminal to asparagine. The sequence variations include glycine which is side-chain deficient, alanine which is apolar and has a minimal steric footprint, serine which is polar but sterically unimposing, and finally histidine which is polar, basic, and bulky. The peptides were deamidated under physiological, catalytic buffer, and high heat conditions. By detecting and quantitating all four isomers, the trends associated with these deamidation conditions can be identified. Importantly, D-asp and D-isoAsp were detected in appreciable amounts under all conditions

and were found to significantly increase when deamidation was carried out in bicarbonate buffer or under elevated temperature conditions.

Results and Discussion

Mass Spectrometry. All experiments were conducted with a series of synthetic peptides that varied in sequence by a single amino acid, 4IB-VKLNXG (where X = G, H, S, A and 4IB = para-iodobenzoic acid). The half-lives for these peptides at 37 °C in 50 mM Tris at pH 7.8 are 1.4, 6.8, 41, and 73 days, respectively. Based on related experiments conducted previously,[2] these sequences should be among the fastest to deamidate while still offering a variety in side chain chemistry. In addition, 4IB was added to each N-terminus to provide a photocleavable radical precursor for RDD experiments (see supporting information for structures). Tris has long been among the buffers of choice to study deamidation *in vitro* due to its modest influence on deamidation.[39]

Following sufficient incubation to deamidate a majority of the sample, peptides were separated and analyzed via LCMS. Typical results are shown in Figure 2.1a, which depicts the ion chromatogram (extracted from full MS scans for the m/z of the intact deamidated peptide) for 4IB-VKLDSG after incubation for 6 months. Only two peaks are apparent, though four Asp isomers could be generated by the pathways illustrated in Scheme 1. Given that the structural differences between L-Asp and D-Asp or L-isoAsp and D-isoAsp are quite subtle, it is likely that isomers have co-eluted. Tandem MS can provide additional information on this possibility in the form of fragmentation profiles.

RDD tandem MS results are shown in Figures 2.1b–1d for locations denoted 1–4, respectively, in Figure 2.1a. In RDD experiments, a radical is generated photolytically, and the product is subjected to collisional activation. Migration of the radical is structurally sensitive, leading to observation of distinct fragmentation patterns for isomers. Indeed, close examination of the spectra shown in Figures 2.1b–1d reveals notable differences. The standard method for quantifying such differences is to calculate an R_{isomer} value, which is obtained from the ratios of product ions changing the most between two spectra (see supporting information for additional details).[40] R_{isomer} values for all possible pairs from Figures 2.1b and 1c were calculated, which identified the x_5 and c_4-56L-CO_2 fragments as those undergoing the greatest relative changes. These fragments yield an R_{isomer} value of 6.3, which is well above the threshold for isomer identification (> 2.4). A similar analysis revealed an R_{isomer} value of 6.1 for the diagnostic fragments labeled in Figures 2.1d and 1e. Fully annotated spectra are available in the supporting information.

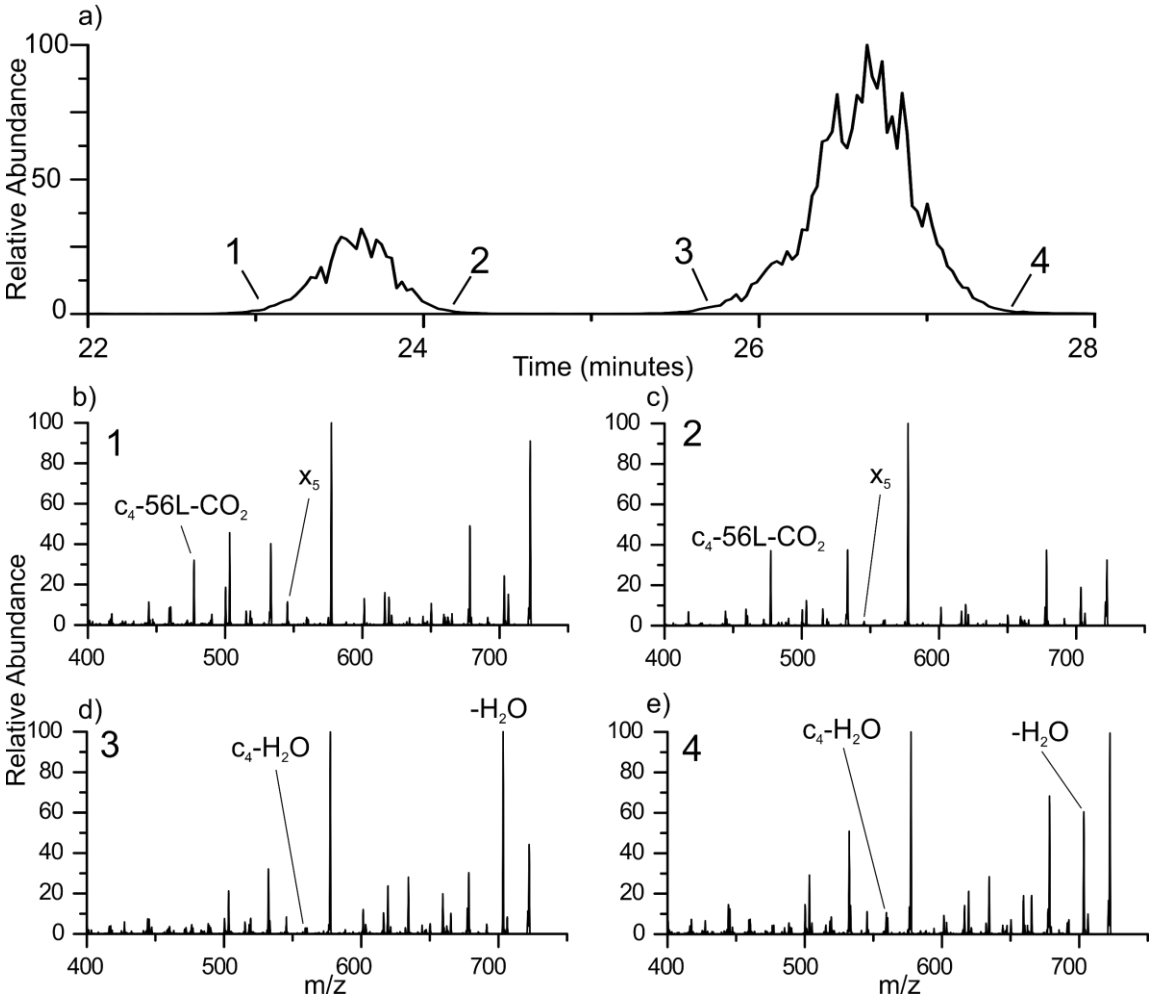


Figure 2.1. (a) LCMS chromatogram of the 4IB-VKLDSG deamidation products which yields only two chromatographic peaks. **(b–e)** RDD tandem mass spectra on [4IB-VKLDSG+H]1+ from the leading and trailing edges of the peaks in **(a)** show differences in fragment intensities that reveal co-eluting isomers within both peaks.

The c_4-56L-CO_2 fragment corresponds to a c_4 backbone fragment which has additionally lost CO_2 and 56 Da from the leucine sidechain (i.e. 56L). Such multiple

dissociation events and side chain losses are common in RDD.35 Predicting the fragments that will yield the maximum R_{isomer} values is not trivial because changes in fragment abundances are due to differences in three-dimensional structure (not sequence). Collision-induced dissociation (CID) was also used to generate fragmentation spectra for all peptides examined herein, but as a less structurally sensitive method, CID usually yields lower R_{isomer} values. The identity of each isomer in Figure 2.1a can be determined by comparing spectra from the leading and trailing edge of each peak to synthetic standards. The spectrum in Figure 2.1b best matches that from D-Asp, while Figure 2.1c matches L-Asp. The dominant isomer represented at each labeled point in Figure 2.1a therefore correspond to 1 = D-Asp, 2 = L-Asp, 3 = D-isoAsp, and 4 = L-isoAsp. A full list of diagnostic peaks and R_{isomer} values used for quantitation is provided in the supporting information.

Although the R_{isomer} values readily confirm co-elution, they do not allow quantitation of each isomer. Results from the synthetic standards for L-Asp and D-Asp isomers of 4IB-VKLDSG are shown in Figures 2.2a and 2b, where again the x_5 and c_4-56L-CO_2 abundances differ significantly between the spectra. By plotting the difference/sum of the two diagnostic peaks as a function of the percent L-Asp, the calibration curve shown in Figure 2.2d was generated. Figure 2.2c shows the averaged fragmentation pattern obtained from the entire LC peak at ~23.5 minutes in Figure 2.1. While similar to both standards, Figure 2.2c is not a perfect match to either Figure 2.2a or 2b because it contains contributions from both isomers. In order to quantify the amount of each isomer present, the data from Figure 2.2c

was mapped onto the calibration curve shown in Figure 2.2d. The LC peak at 23.5 minutes in Figure 2.1 contains 88% L-Asp and 12% D-Asp.

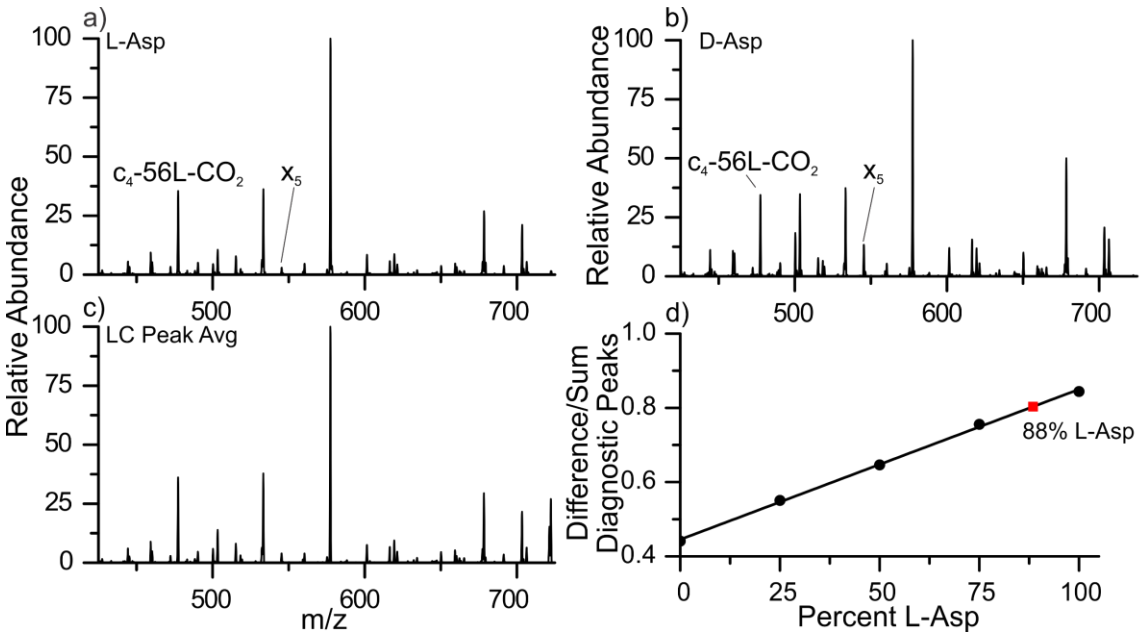


Figure 2.2. (a,b) Standards of the L-Asp and D-Asp isomers of 4IB-VKLDSG are distinguished by diagnostic fragment peaks. **c)** The average fragmentation abundance for the LC peak that contains both isomers. **d)** A calibration curve generated from purified standards reveals the LC peak contains 88% L-Asp (red square).

This methodology was repeated using both RDD and CID for each peptide. Calibration curves were generated using the fragmentation method that afforded the best isomer discrimination. Interestingly, 4IB-VKLNAG separates into three fully resolved peaks, representing D-Asp, L-Asp, and L-/D- isoAsp. Quantitation via peak area matched the results derived from the RDD calibration curve for the

ratio of L-Asp and D-Asp to within 0.5% agreement (see supporting information), offering independent confirmation of our approach.

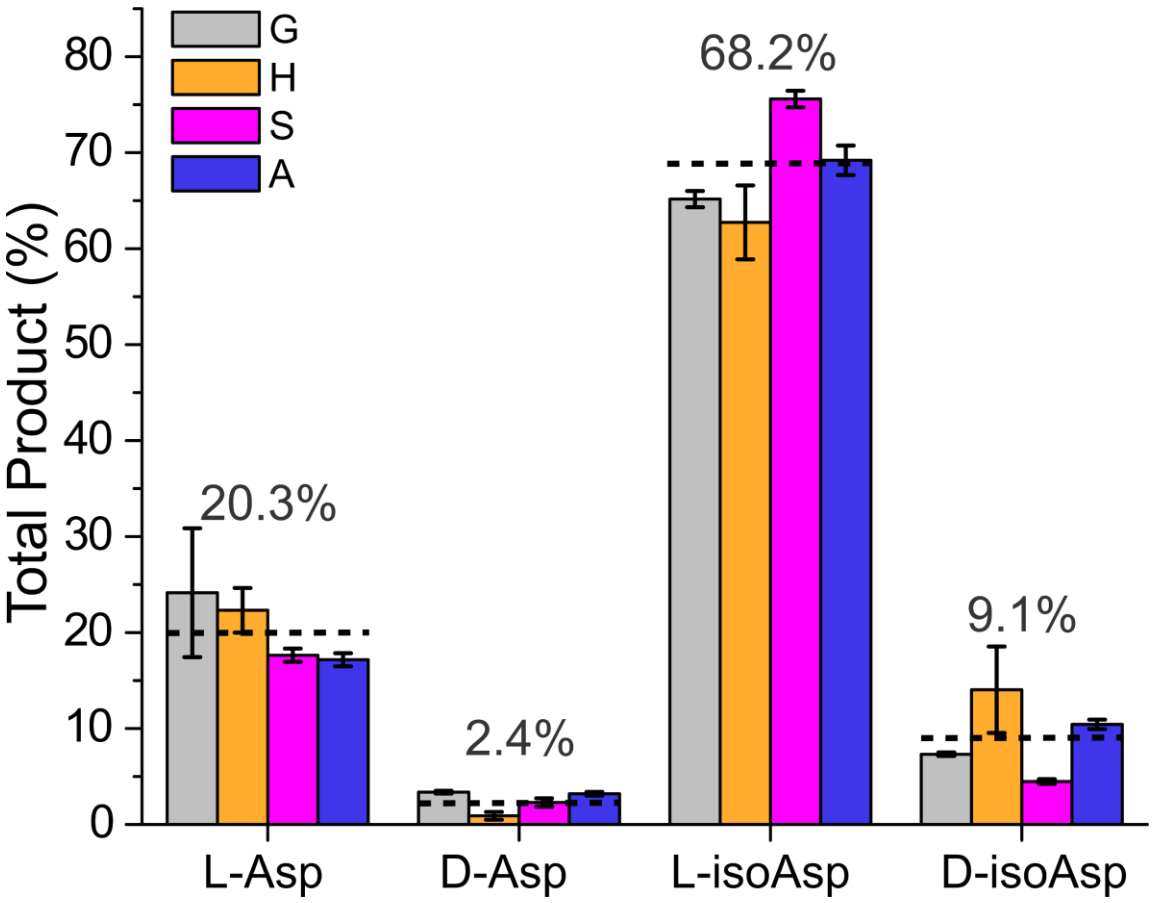


Figure 2.3. Deamidation products at 37 °C in 50mM Tris at pH 7.8. Averages are indicated with dashed lines.

The results obtained in Tris at pH 7.8 for all four peptide sequences are summarized in Figure 2.3. Despite having considerably different half-lives and side chain chemistry, the results from all four peptides exhibit similar trends. L-isoAsp

is generated at 68.2% on average and is confirmed to be the dominant product for each peptide sequence, in agreement with previous findings.[41] In comparison, L-Asp represents only 20.3% of the total deamidation products on average. A similar preference for the iso-variant is observed between D-isoAsp and D-Asp, with D-isoAsp present in at 9.1% on average compared to 2.4% for D-Asp. The paucity of D-Asp is easily rationalized because its formation requires both racemization of the succinimide ring and hydrolysis to open the ring via the less favorable mechanism. On the other hand, the abundance of D-isoAsp is significant and illustrates that racemization during deamidation should not be ignored.

Although deamidation always results in the mutation of asparagine to aspartic acid, the observation that ~80% of the total deamidation products represent unnatural amino acids may explain the widespread immunogenic responses associated with deamidation *in vivo.*[42–44] For the same reason, deamidation in protein therapeutics is likely to have undesirable effects ranging from loss of activity to altered immunogenicity.44,[45] The high abundance of D-isoAsp may also be cause for a re-examination of previous conclusions discounting its importance. For example, the presumed-to-be negligible abundance of D-isoAsp was previously used to rationalize the substrate specificity of the repair enzyme, protein L-isoaspartate methyltransferase (PIMT), which can bind L-isoAsp and D-Asp but not D-isoAsp residues.[9] Given that D-Asp is significantly less abundant than D-isoAsp, PIMT specificity is likely due to other factors. One potential explanation is that PIMT requires the α-nitrogen atom to be positioned ‘behind’ the methyl-

accepting carboxyl group as is the case for L-isoAsp and D-Asp, but not L-Asp or D-isoAsp.41

Buffer Conditions.

The influence of buffer, pH, temperature, and ammonia (a known catalyst)28 were explored, and the results are summarized in Figure 2.4. To simplify comparison, the relative abundances of each isomer are plotted together for all conditions that were examined. The first grouping for each isomer represents data acquired at 37 °C in 50mM Tris at pH 7.8 and is derived from the results shown in Figure 2.3 (replotted for easy comparison). For all remaining experiments ammonia was added (indicated by the green box) as ammonium hydroxide to a final concentration of 100mM ammonia in Tris and CHES at each pH or as 100mM ammonium bicarbonate for the carbonate buffer. Ammonia has been reported to accelerate deamidation through general base catalysis,28 but its influence on deamidation products has not been previously determined. Although rates are accelerated (as discussed in greater detail below), the addition of ammonia and a slight increase to pH 8.8 do not significantly influence the distribution of Asp isomers generated by deamidation in Tris buffer (compare Figure 2.4, Tris 7.8 and Tris 8.8).

Results obtained in CHES buffer at pH 8.8 and 9.8 are also shown in Figure 2.4. Although CHES buffer and Tris are chemically dissimilar and have pK_a's of 9.5 and 8.1 respectively,[46] there is little influence on the overall distribution of Asp

isomers produced by the two buffers. Furthermore, increasing the pH to 9.8 does not have a significant impact on the results, though minor variations in product abundance are noted. Intracellular pH can vary by over 3 pH units, ranging from 4.7 to 8.0.[47] Our data indicate that even the most basic intracellular conditions are unlikely to strongly influence succinimide-mediated deamidation products.

The effect of heating was explored by raising the temperature to 80 °C in CHES at pH 9.8 (these results are marked with an asterisk in Figure 2.4). Increasing the temperature impacts the resulting isomer distribution by decreasing the amounts of L-Asp and L-isoAsp while increasing the proportion D-Asp and D-isoAsp at the same time. These results suggest that higher temperature leads to greater racemization of the intermediate succinimide ring, leading to greater production of the D-isomers. After racemization, ring-opening again favors formation of the iso-products. Consequently, elevated temperatures used previously to accelerate deamidation *in vitro* may have been biased towards the production of the D-isomers.[28,48] Interestingly, the increase in D-Asp observed at elevated temperatures may have prompted an evolutionary adaptation in the hyperthermophile *Pyrococcus furiosus*, which expresses a variant of PIMT that repairs D-Asp containing peptides with 120-fold higher affinity relative to the human enzyme.[49] D-isoAsp, however, is still not a substrate for this variant.

Experiments with bicarbonate buffer at pH 7.7 also revealed altered isomer product distributions. Although the amount of L-Asp generated is comparable to

other buffers, the production of L-isoAsp is significantly reduced. This observation suggests that ring-opening to produce L-isoAsp is specifically disfavored by the presence of bicarbonate. The isomer abundance that would typically end up in the L-isoAsp channel is redirected to the D-isomers, which are both more abundant relative to other buffer conditions. These results from bicarbonate differ from those obtained at 80 °C because the increased yield of D-isomers derives from only L-isoAsp rather than from a diminution of both L-isomers. The enhanced racemization observed in bicarbonate is of particular interest because bicarbonate has been the buffer of choice for many deamidation and digestion studies.[50–52] Bicarbonate also plays a prominent role in the native blood-buffer system.[53] It is surprising that bicarbonate buffer yields the highest amount of racemization despite having the lowest initial pH (7.7). Carboxylic acids have been reported to stabilize deamidation transition states. We attribute the increase in racemization to noncovalent interactions with bicarbonate that disfavor formation of L-isoAsp.

Though spontaneous and pervasive, deamidation is typically slow under physiological conditions, which has deterred large-scale comprehensive studies and favored examination of sequences that deamidate quickly. Evaluation of deamidation can also be complicated by spontaneous degradation for certain peptide sequences.28,[54] These limitations could potentially be overcome if the rate of deamidation could be accelerated without biasing the products or favoring other degradation pathways. Our results suggest that both Tris and CHES based buffer systems can be used with ammonia as a catalyst over a wide pH range without

significantly influencing the distribution of Asp isomers that will be formed. However, caution should be taken if elevated temperature, bicarbonate or other untested buffers are to be used. Although the ratio of (total isoaspartic acids)/(total aspartic acids) has long been the metric of choice for deamidation studies due to the relative ease of analysis, our data reveal that the overall product distribution can shift significantly without changing the isoAsp/Asp ratio. For instance, the isoAsp/Asp ratio for CHES 9.8 at 37 °C is 3.3 while the ratio at 80 °C is 3.4 despite the noted increase in D-isomers. Chirally distinct isomers are likely to have different biological impacts, emphasizing the importance of quantitating all four isomers.

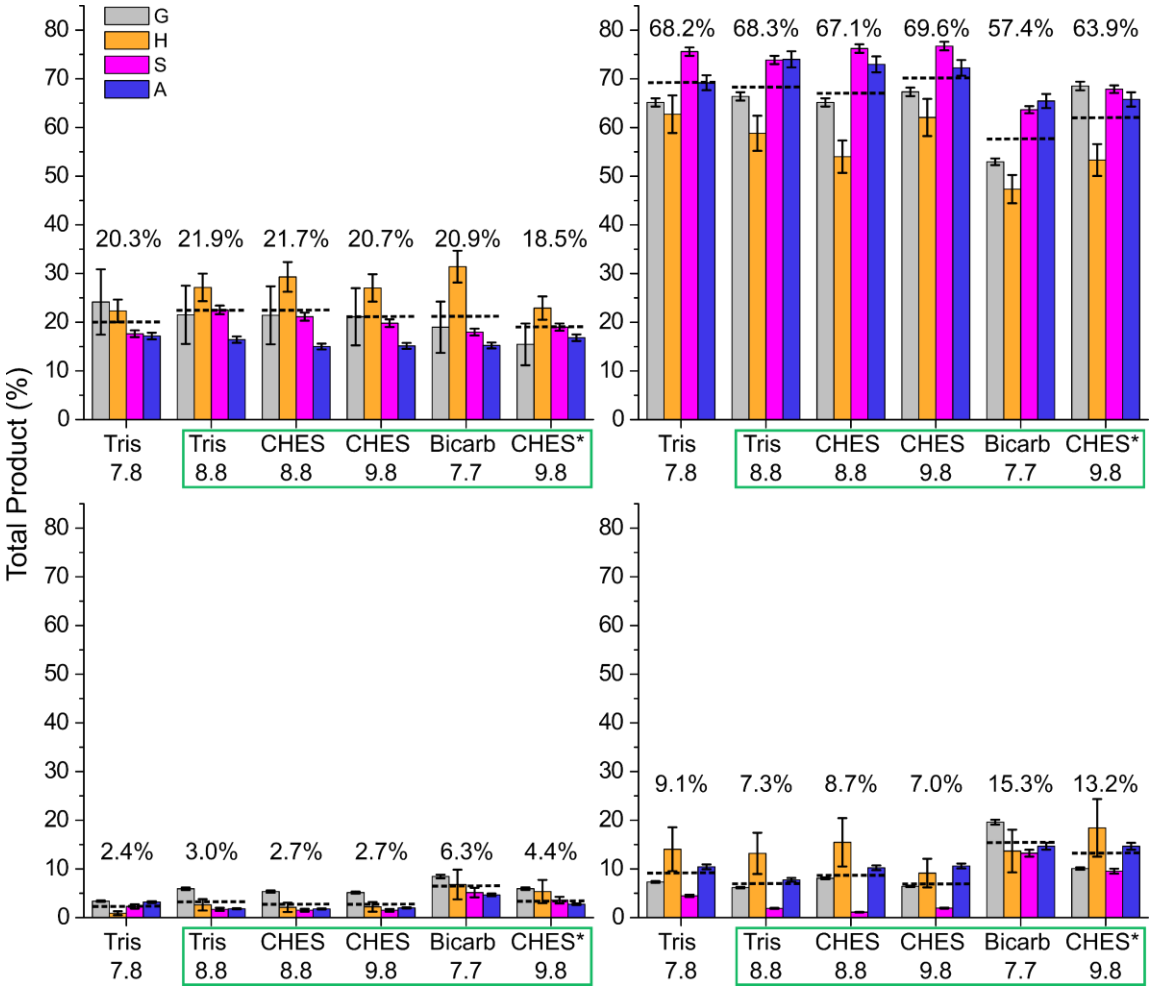


Figure 2.4. Total product outcomes from 6 different buffer conditions. Clockwise from top left: L-Asp, L-isoAsp, D-isoAsp, D-Asp). Averages are depicted with horizontal bars. Green box indicates ammonia catalysis. The asterisk indicates 80 °C.

Rate Enhancement.

The degree of rate enhancement relative to Tris (pH 7.8, 37 °C) is shown as a fold increase in Figure 2.5 for each buffer and peptide sequence. On average, Tris

8.8 was found to be 17x faster than Tris 7.8, where potential acceleration derives from both increased pH and ammonia catalysis. Comparing CHES at 8.8 and 9.8 reveals that increasing the buffer by one pH unit yields only a modest increase in the rate, from 1.2x to 2.0x faster (relative to Tris 7.8). This observation suggests that the bulk of the acceleration between the Tris samples is due to the presence of ammonia. This rate enhancement, along with the modest effect on isomer product distributions shown above, suggests ammonia is a viable catalyst for deamidation studies.

The most dramatic average rate increase (81x) was observed by increasing the temperature to 80 °C in CHES. At this temperature, over three-quarters of 4IB-VKLNGG was deamidated within 30 minutes, making it difficult to determine the deamidation rate accurately. The reported rate increase for this peptide in Figure 2.5 is likely accompanied by significant uncertainty, though deamidation clearly occurs rapidly. Although elevated temperature appears to provide the greatest increase in deamidation rate, higher temperature also influences the product distribution, and yields increased racemization.

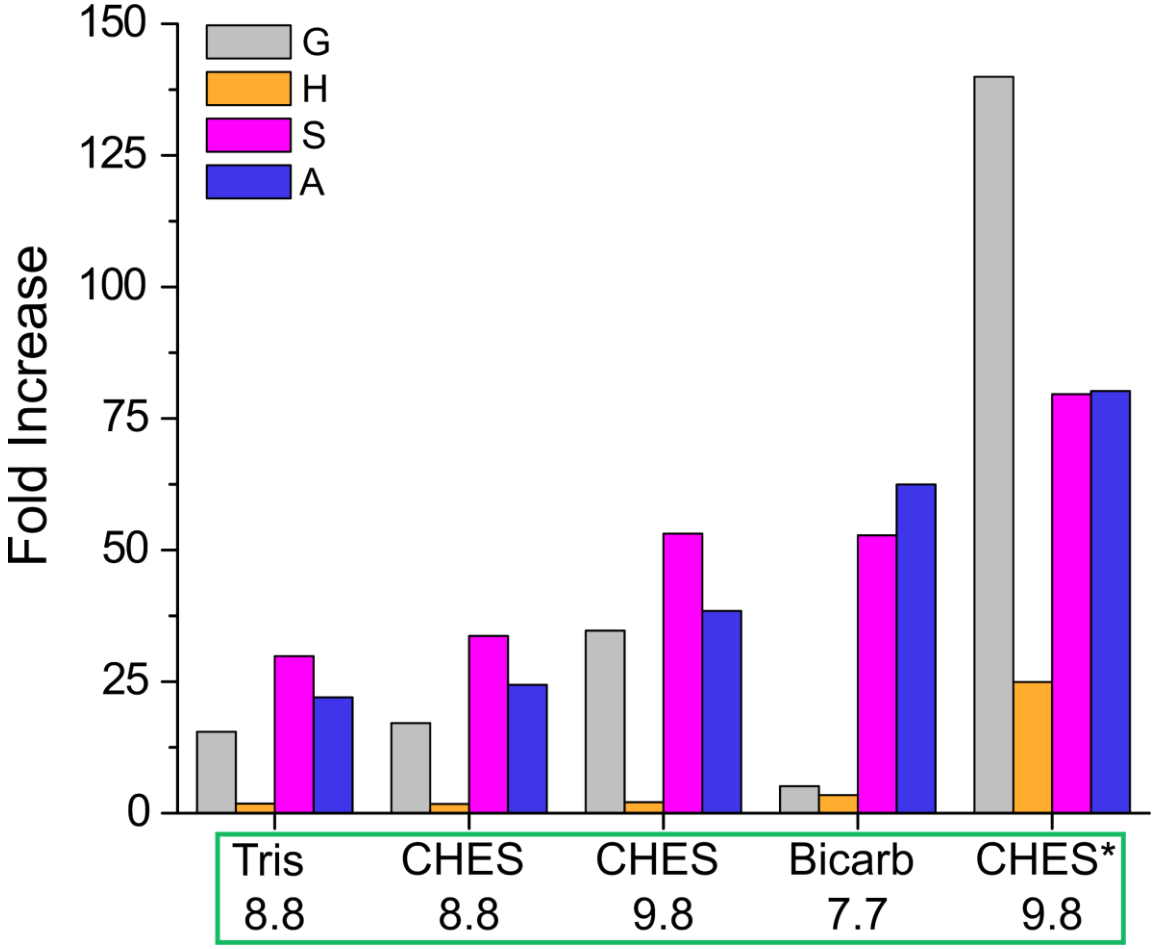


Figure 2.5. Deamidation rate increases relative to Tris 7.8 for each accelerated buffer condition. The green box indicates ammonia catalysis. The asterisk indicates 80 °C.

Strikingly, the rate of acceleration varies greatly among different peptide sequences. Most notably, the histidine-containing peptide is least affected by the addition of ammonia. This result is particularly interesting because peptides with asparagine followed by histidine have previously attracted attention due to their rapid deamidation rates despite the bulky nature of histidine. Ultimately, this rate

enhancement has been attributed to the ability of the histidine sidechain to act as a general base.[54] If true, the addition of ammonia, which also enhances the rate of deamidation by general base catalysis,28 would not be expected to have a significant effect, in agreement with the results in Figure 2.5. Importantly, the histidine peptide does deamidate more quickly at increased temperature, which has little effect on acid/base chemistry but increases the energy available to the system. This observation further confirms that the basic side chain of histidine likely explains the lack of response to the addition of ammonia.

Curiously, the rate of deamidation does not increase as significantly with ammonia for 4IB-VKLNGG in bicarbonate, though increases for this peptide were comparable to the other non-histidine residues in all other conditions. This result may suggest that carbonate interacts strongly with the backbone, inhibiting access to the required transition state. The absence of a sidechain for the glycine peptide may increase interactions between carbonate and the backbone.

Methods

Materials. Organic solvents and reagents were purchased from Fisher Scientific or Acros Organics and used without further purification. Water was purified to 18.2 MΩ using a Millipore 147. FMOC-protected amino acids and Wang resins were purchased from Anaspec, Inc or Chem-Impex International.

Peptide Synthesis. Peptides were synthesized manually following an accelerated FMOC-protected solid-phase peptide synthesis protocol.[55] *Para*-

iodobenzoic acid (4IB) was attached to the N-terminus of the nascent polypeptide during synthesis. Following synthesis, peptides were purified using a Phenomenex Jupiter Proteo C12 4 µm 90 Å 250 mm x 4.6 mm column. Purified peptides were stored frozen in 50/50 acetonitrile/water (v/v).

Deamidation. Peptides were deamidated in various conditions to explore the impact of buffer, pH, ammonia catalysis, and temperature. 50 mM Tris pH 7.8 at 37 °C results are quantitated and used as reference values for subsequent comparisons which include 50 mM Tris pH 8.8 + 100 mM NH_3 at 37 °C, 50 mM CHES pH 8.8 + 100 mM NH_3 at 37 °C, 50 mM CHES pH 9.8 + 100 mM NH_3 at 37 °C, 100 mM ammonium bicarbonate at 37 °C, and 50 mM CHES + 100 mM at 80 °C. For the Tris and CHES buffers at pH 8.8 and 9.8, ammonium hydroxide was added at 381 mM and 128 mM respectively to yield a final concentration of 100 mM unprotonated ammonia.

Analysis. Following deamidation, peptides were analyzed via LCMS. An Agilent 1100 binary pump was used with a 5 µm 100 Å C5 50 mm x 2 mm column (Phenomenex) interfaced to a Thermo Fisher Scientific LTQ mass spectrometer with a standard electrospray ionization source. Deamidated peptides were eluted with an isocratic mixture of 18% acetonitrile and 82% water with 0.1% formic acid (v/v). Synthetic standards were prepared as ~10µM samples in 49.5/49.5/1 methanol/water/acetic acid (v/v) and infused into a modified LTQ linear ion trap using the standard electrospray ionization source. The LTQ was modified with a quartz window to allow fourth harmonic (266 nm) laser pulses from a Nd:YAG laser

to irradiate the trapped ion cloud which allows for photoinitiated radical directed dissociation.[56]

R_{isomer} Values. Isomers are distinguished based on MS^n fragmentation spectra. To quantitatively evaluate the differences between two spectra, R_{isomer} values are calculated using Equation 1 where R_1 and R_2 refer to the ratios of a pair of fragment ions that vary the most in abundance between two different MS^n scans.

$$R_{isomer} = \frac{R_1}{R_2} \qquad (1)$$

Identical fragmentation patterns result in R_{isomer} values of 1 thereby indicating no discrimination, while larger values reflect a higher degree of discrimination. Values >2.4 obtained via LCMS are indicative of analytes that can be distinguished based on fragmentation patterns.[38]

Conclusions

Although deamidation yields similar product trends among vastly different buffer conditions and peptide sequences, it is noticeably sensitive to specific extrinsic factors. Both bicarbonate buffer and high temperature strongly promote racemization, while increasing pH from 7.8–9.8 has a surprisingly limited effect. Importantly, we have observed that D-isomers, which are often ignored, are generated in appreciable quantities under all conditions. These results establish the innate deamidation tendencies for all four Asp products and demonstrate that isomer distributions are susceptible to external influences. Furthermore, this

methodology should be able to quantify isomers for many classes of peptides, including isomers identified in proteomics experiments targeting long-lived proteins relevant to aging, human disease, and biological therapeutics. Understanding the products of deamidation, and the factors that alter them should facilitate more reliable prediction and control of degradation products. This information may become integral for engineering proteins that minimize the detrimental effects of deamidation and isomerization.

Acknowledgements

We gratefully acknowledge funding from the National Institutes of Health (Grant R01GM107099)

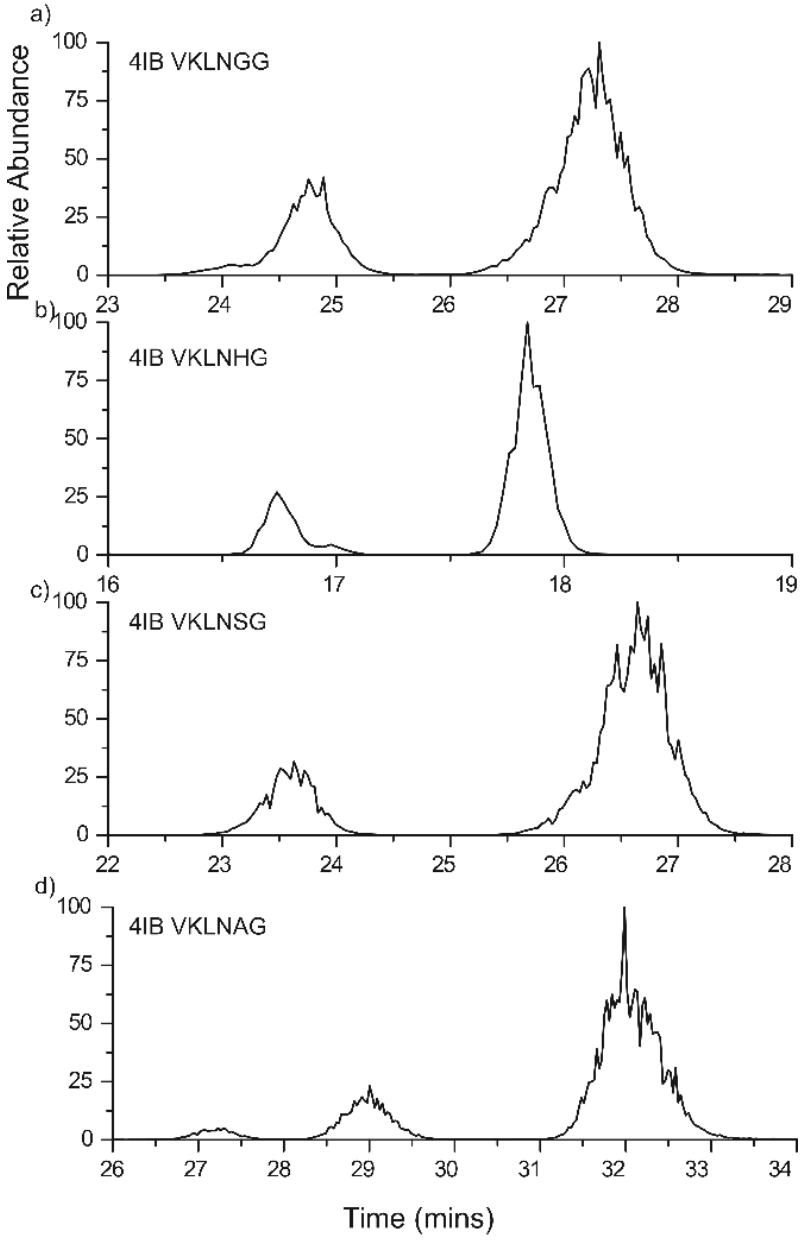


Figure 2.S1. Typical LCMS chromatograms depicting the separation of the deamidation products for all four peptide sequences in Tris 7.8 at 37 °C.

In the case of 4IB VKLNAG (d), the LC yields three fully resolved peaks representing D-Asp, L-Asp and L-/D- isoAsp. The results from calculating peak area are in good agreement with the results obtained via RDD:

Peak area: 15.1% L-Asp, 2.7% D-Asp

RDD: 15.6% L-Asp, 2.3% D-Asp

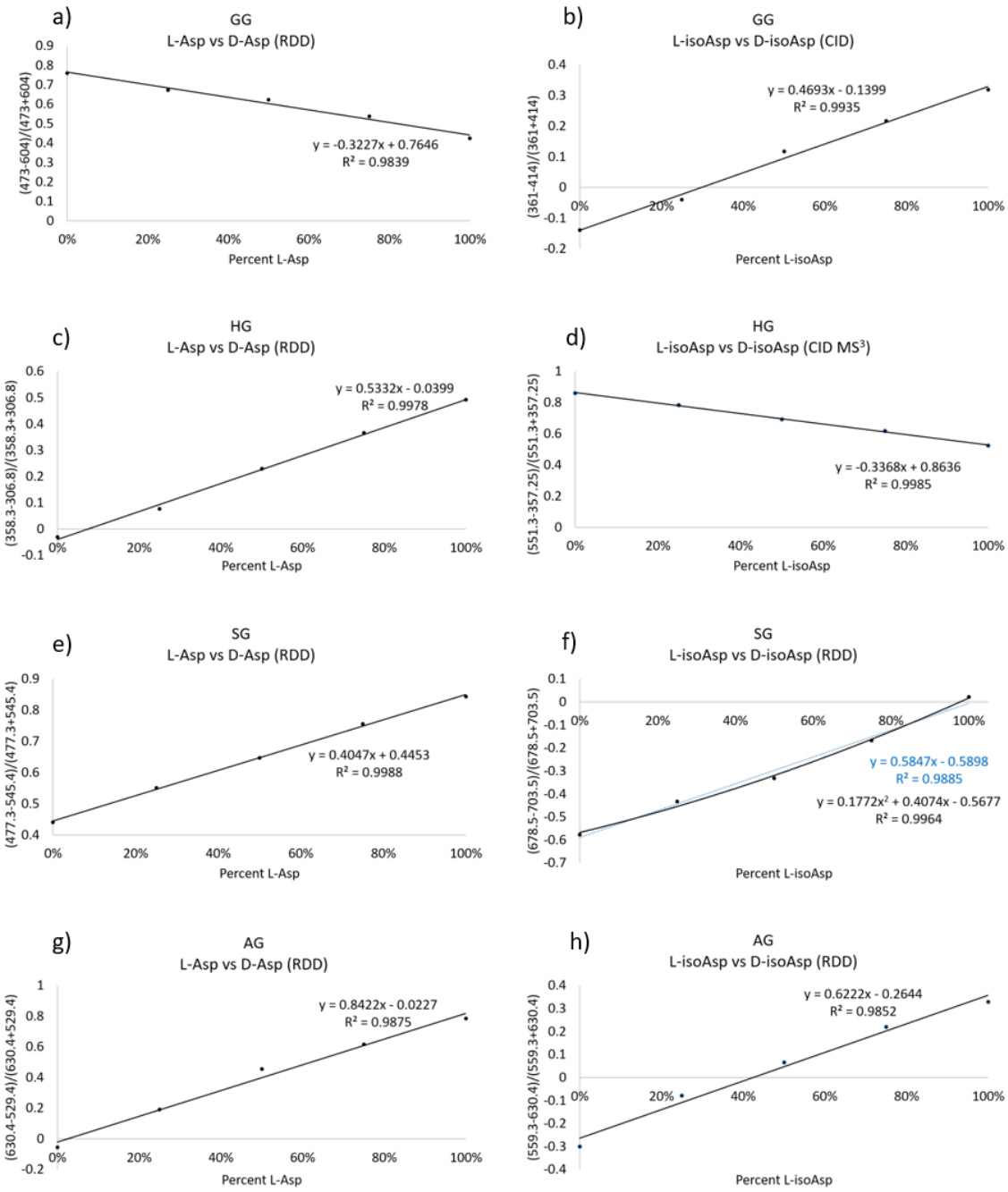

a)
GG
L-Asp vs D-Asp (RDD)
(473-604)/(473+604)
y = -0.3227x + 0.7646
R² = 0.9839
Percent L-Asp
b)
GG
L-isoAsp vs D-isoAsp (CID)
(361-414)/(361+414)
y = 0.4693x - 0.1399
R² = 0.9935
Percent L-isoAsp
c)
HG
L-Asp vs D-Asp (RDD)
(358.3-306.8)/(358.3+306.8)
y = 0.5332x - 0.0399
R² = 0.9978
Percent L-Asp
d)
HG
L-isoAsp vs D-isoAsp (CID MS³)
(551.3-357.25)/(551.3+357.25)
y = -0.3368x + 0.8636
R² = 0.9985
Percent L-isoAsp
e)
SG
L-Asp vs D-Asp (RDD)
(477.3-545.4)/(477.3+545.4)
y = 0.4047x + 0.4453
R² = 0.9988
Percent L-Asp
f)
SG
L-isoAsp vs D-isoAsp (RDD)
(678.5-703.5)/(678.5+703.5)
y = 0.5847x - 0.5898
R² = 0.9885
y = 0.1772x² + 0.4074x - 0.5677
R² = 0.9964
Percent L-isoAsp
g)
AG
L-Asp vs D-Asp (RDD)
(630.4-529.4)/(630.4+529.4)
y = 0.8422x - 0.0227
R² = 0.9875
Percent L-Asp
h)
AG
L-isoAsp vs D-isoAsp (RDD)
(559.3-630.4)/(559.3+630.4)
y = 0.6222x - 0.2644
R² = 0.9852
Percent L-isoAsp

Figure 2.S2. Calibration curves for each of the coeluting isomers with the dissociation method indicated. The y-axis is labeled as the difference/sum of the two diagnostic fragment ions observed m/z. For the SG peptide L-isoAsp vs D-isoAsp comparison (f) the linear fit (blue) exhibits systematic error of a few percent at high L-isoAsp percentages – to reduce this error, a polynomial fit (black) is used instead.

Fragment assignments for calibration curves are as follows:

4IB VKLDGG:

L-Asp vs D-Asp: 473.3 = z_5^+, 604.3 = $[M^{\cdot}+H]$-43L-CO_2^+. R_{isomer} = 2.97

L-isoAsp vs D-isoAsp: 361.2 = y_4^+, 414.3 = $KLDG^+$. R_{isomer} = 2.56

4IB VKLDHG:

L-Asp vs D-Asp: 306.8 = a_5-$56L^{2+}$, 358.3 = $[M^{\cdot}+2H]$-$56L^{2+}$. R_{isomer} = 3.12

L-isoAsp vs D-isoAsp: 357.3 = KLD^+, 551.3 = y_5-$H2O^+$. R_{isomer} = 4.18

4IB VKLDSG:

L-Asp vs D-Asp: 477.3 = c_4-56L-CO_2^+, 545.4 = x_5^+. R_{isomer} = 4.59

L-isoAsp vs D-isoAsp: 678.5 = $[M^{\cdot}+H]$-$43L^+$, 703.5 = $[M^{\cdot}+H]$-H_2O^+. R_{isomer} = 3.88

4IB VKLDAG: 559.3 = a_5-CO_2^+.

L-Asp vs D-Asp: 529.4 = a_5-56L-H_2O^+, 630.4 = c_5-H_2O^+. R_{isomer} = 9.25

L-isoAsp vs D-isoAsp: 559.3 = a_5-CO_2^+, 630.4 = c_5-H_2O^+. R_{isomer} = 3.68

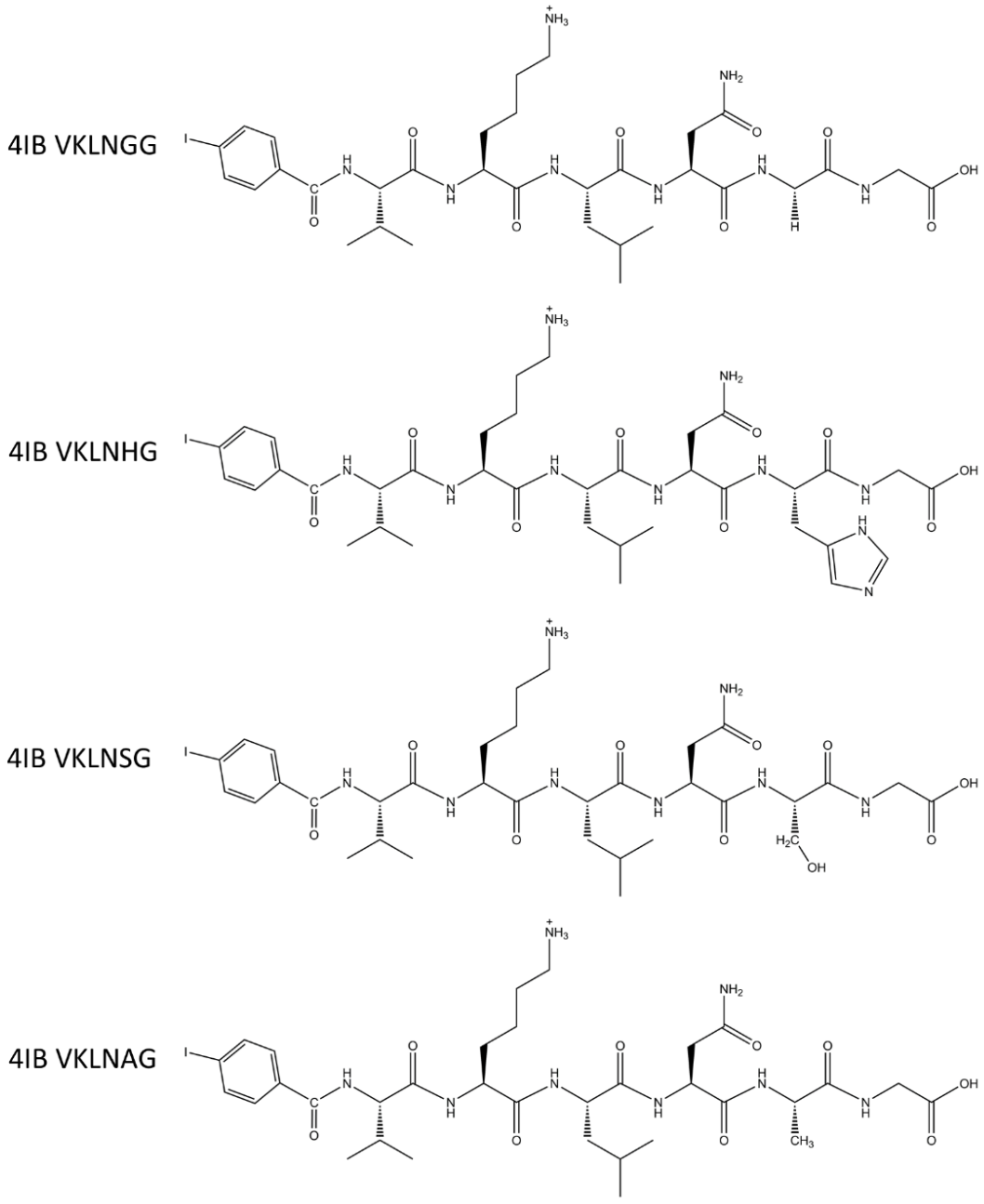


Figure 2.S3. Peptide structures following 4IB attachment. Note only the histidine sidechain can participate in typical acid/base chemistry in the examined pH range.

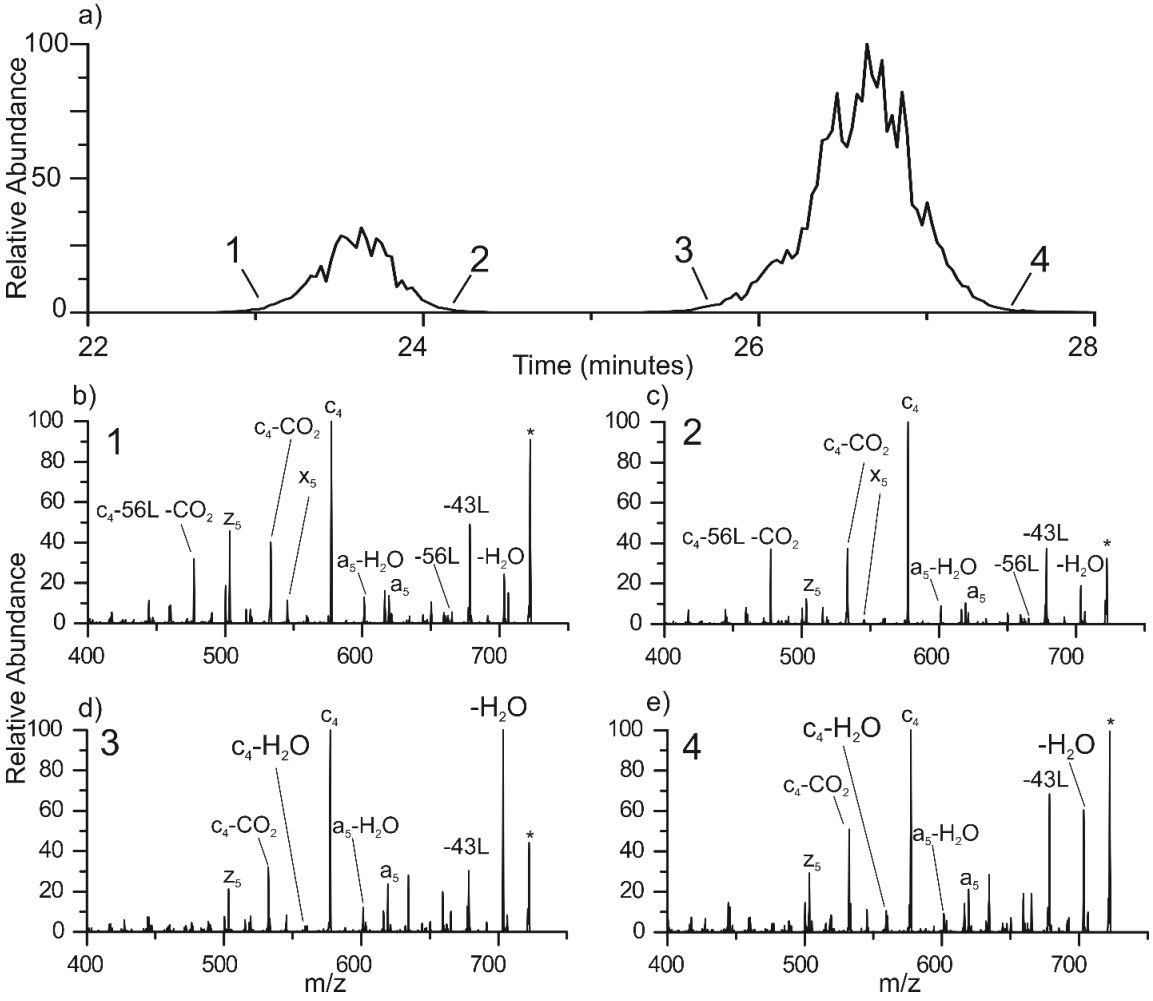


Figure 2.S4. Figure 2.1 from the main text with additional fragment labels. * indicates the radical precursor produced via photodissociation. The same Figure 2.labels apply: **(a)** LCMS chromatogram of the 4IB-VKLDSG deamidation products which yields only two chromatographic peaks. **(b-e)** RDD tandem mass spectra on $[4IB\text{-}VKLDSG+H]^{1+}$ from the leading and trailing edges of the peaks in (a) show differences in fragment intensities that reveal co-eluting isomers within both peaks.

R_{isomer} scores

To quantitatively determine the difference between two fragmentation spectra, R values are used. R values refer to the ratio of relative abundances for two fragment peaks within an MS^n spectrum. In this study, R values are always reported for pairs of isomers, thus R_{isomer} values are presented as defined by Equation 1 where R_1 represents the R value for isomer 1, and R_2 represents the R value of isomer 2. Thus, R_{isomer} scores represent a ratio of R values between two spectra.

$$R_{isomer} = \frac{R_1}{R_2} \qquad (1)$$

Identical fragmentation patterns result in result in R_{isomer} values of 1 thereby indicating no discrimination, while larger values reflect a higher degree of discrimination. The relevant threshold for R_{isomer} values has previously been determined for LCMS analysis by use of a standard *t* test.[40] R_{isomer} values > 1.9 for CID and >2.4 for RDD are outside the 99% confidence interval and therefore positively identify isomers. See reference 39, and 40 for additional examples.

References

[1] Huang, Q., Chang, J., Cheung, M. K., Nong, W., Li, L., Lee, M., Kwan, H. S. (2014**)** Human Proteins with Target Sites of Multiple Post-Translational Modification Types Are More Prone to Be Involved in Disease, *J. Proteome Res. 13*, 2735–2748.

[2] Robinson, N. E., Robinson, A. B. (2001) Molecular Clocks, *Proc. Natl. Acad. Sci. 98*, 944–949.

[3] Curnis, F., Longhi, R., Crippa, L., Cattaneo, A., Dondossola, E., Bachi, A., Corti, A. (2006**)** Spontaneous Formation of L-Isoaspartate and Gain of Function in Fibronectin, *J. Biol. Chem. 281*, 36466–36476.

[4] Lindner, H., Helliger, W. (2001) Age-Dependent Deamidation of Asparagine Residues in Proteins, *Exp. Gerontol. 36*, 1551–1563.

[5] Jilek, A., Kreil, G. (2008) D-Amino Acids in Animal Peptides, *Monatshefte für Chemie*. *139*, 1–5.

[6] Truscott, R. J. W., Friedrich, M. G. (2014) Old Proteins and the Achilles Heel of Mass Spectrometry, The Role of Proteomics in the Etiology of Human Cataract. *Proteomics - Clin. Appl. 8*, 195–203.

[7] Kaneko, I., Morimoto, K., Kubo, T. (2002) Drastic Neuronal Loss in Vivo by β-Amyloid Racemized at Ser^{26} Residue: Conversion of Non-Toxic [$D\text{-}Ser^{26}$] β-Amyloid 1-40 to Toxic and Proteinase-Resistant Fragments presence in the brains of Alzheimer's patients, *Neuroscience 104*, 1003–1011.

[8] Shimizu, T., Matsuoka, Y., (2005) Shirasawa, T. D-Amino Acid Biosystem Biological Significance of Isoaspartate and Its Repair System, *Biol. Pharm. Bull. 28*, 1590–1596.

[9] Hooi, M. Y. S., Raftery, M. J., Truscott, R. J. W. (2013) Interconversion of the Peptide Isoforms of Aspartate: Stability of Isoaspartates, *Mech. Ageing Dev. 134*, 103–109.

[10] Lyon, Y. A.; Sabbah, G. M., Julian, R. R. (2017) Identification of Sequence Similarities among Isomerization Hotspots in Crystallin Proteins, *J. Proteome Res. 16*, 1797–1805.

[11] Truscott, R. J. W., Friedrich, M. G. (2016) The Etiology of Human Age-Related Cataract. Proteins Don't Last Forever, *Biochim. Biophys. Acta. 1860*, 192–198.

[12] 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.

[13] Kozin, S. A., Mitkevich, V. A., Makarov, A. A. (2016) Amyloid-β Containing Isoaspartate 7 as Potential Biomarker and Drug Target in Alzheimer's Disease, *Mendeleev Commun. 26*, 269–275.

[14] Roher, A. E., Lowenson, J. D., Clarke, S., Wolcow, C., Wang, R., Cotter, R. J., Reardon, I. M., Zürcher-Neely, H. A., Heinrikson, R. L., Ball, M. J., and Greenberg

B. D. (1993) Structural Alteration in the Peptide Backbone of β-Amyloid Core Protein May Account for Its Deposition and Stability in Alzheimerr's Disease, *J. Biol. Chem. 268*, 3072–3083.
[15] Doyle, H. A., Mamula, M. J. (2001) Post-Translational Protein Modifications in Antigen Recognition and Autoimmunity, *Trends Immunol. 22*, 443–449.
[16] Friedrich, M. G., Hancock, S. E., Raftery, M. J., Truscott, R. J. W. (2016) Isoaspartic Acid Is Present at Specific Sites in Myelin Basic Protein from Multiple Sclerosis Patients: Could This Represent a Trigger for Disease Onset?, *Acta Neuropathol. Commun. 4:83*.
[17] Mamula, M. J., Gee, R. J., Elliott, J. I., Sette, A., Southwood, S., Jones, P. J., Blier, P. R. (1999) Isoaspartyl Post-Translational Modification Triggers Autoimmune Responses to Self-Proteins, *J. Biol. Chem. 274*, 22321–22327.
[18] Liu, H., Nowak, C., Shao, M., Ponniah, G., Neill, A. (2016) Impact of Cell Culture on Recombinant Monoclonal Antibody Product Heterogeneity, *Biotechnol. Prog. 32*, 1103–1112.
[19] Jha, S. K., Deepalakshmi, P. D., Udgaonkar, J. B. (2012) Characterization of Deamidation of Barstar Using Electrospray Ionization Quadrupole Time-of-Flight Mass Spectrometry, Which Stabilizes an Equilibrium Unfolding Intermediate, *Protein Sci. 21*, 633–646.
[20] Takata, T., Oxford, J. T., Demeler, B., Lampi, K. J. (2008) Deamidation Destabilizes and Triggers Aggregation of a Lens Protein, βA3-Crystallin, 17, 1565–1575.
[21] Wakankar, A. A., Borchardt, R. T. (2006) Formulation Considerations for Proteins Susceptible to Asparagine Deamidation and Aspartate Isomerization, *J. Pharm. Sci. 95*, 2321–2336.
[22] Ecker, D. M., Jones, S. D., Levine, H. L. (2015) The Therapeutic Monoclonal Antibody Market, *MAbs*. *7*, 9–14.
[23] Carter, P. J. (2011) Introduction to Current and Future Protein Therapeutics: A Protein Engineering Perspective, *Exp. Cell Res. 317*, 1261–1269.
[24] Eakin, C. M., Miller, A., Kerr, J., Kung, J., Wallace, A. (2014) Assessing Analytical Methods to Monitor isoAsp Formation in Monoclonal Antibodies, *Front. Pharmacol. 5*, 1–9.
[25] 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. (2008) Isomerization of a Single Aspartyl Residue of Anti-Epidermal Growth Factor Receptor Immunoglobulin γ2 Antibody Highlights the Role Avidity Plays in Antibody Activity, *Biochemistry* 47, 2518–2530.
[26] Geiger, T., Clarke, S. (1987) Deamidation, Isomerization, and Racemization at Asparaginyl and Aspartyl Residues in Peptides, Succinimide-Linked Reactions That Contribute to Protein Degradation, *J. Biol. Chem. 262*, 785–794.
[27] Radkiewicz, J. L., Zipse, H., Clarke, S., Houk, K. N. (2001) Neighboring Side Chain Effects on Asparaginyl and Aspartyl Degradation: An Ab Initio Study of the

Relationship between Peptide Conformation and Backbone NH Acidity, *J. Am. Chem. Soc. 123*, 3499–3506.
[28] Tyler-Cross, R., Schirch, V. (1991) Effects of Amino Acid Sequence, Buffers, and Ionic Strength on the Rate and Mechanism of Deamidation of Asparagine Residues in Small Peptides, *J. Biol. Chem. 266*, 22549–22556.
[29] Yang, H., Zubarev, R. A. (2010) Mass Spectrometric Analysis of Asparagine Deamidation and Aspartate Isomerization in Polypeptides, *Electrophoresis 31*, 1764–1772.
[30] Jia, C., Lietz, C. B., Yu, Q., Li, L. (2013) Site-Specific Characterization of D-Amino Acid Containing Peptide Epimers by Ion Mobility Spectrometry, *Anal. Chem. 86*, 2972–2981.
[31] Zheng, X., Deng, L., Baker, E. S., Ibrahim, Y. M., Petyuk, V. A., Smith, R. D. (2017) Distinguishing D- and L-Aspartic and Isoaspartic Acids in Amyloid β Peptides with Ultrahigh Resolution Ion Mobility Spectrometry, *Chem. Commun. 56,* 7913–7916.
[32] Yang, H., Fung, E. Y. M., Zubarev, A. R., Zubarev, R. A. (2009) Toward Proteome-Scale Identification and Quantification of Isoaspartyl Residues in Biological Samples, *J. Proteome Res. 8*, 4615–4621.
[33] 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.
[34] Ly, T., Zhang, X., Sun, Q. Y., Moore, B., Tao, Y. Q., Julian, R. R. (2011) Rapid, Quantitative, and Site Specific Synthesis of Biomolecular Radicals from a Simple Photocaged Precursor, *Chem. Commun. 47*, 2835–2837.
[35] Sun, Q., Nelson, H., Ly, T., Stoltz, B. M., Julian, R. R. (2008) Side Chain Chemistry Mediates Backbone Fragmentation in Hydrogen Deficient Peptide Radicals, *J. Proteome Res. 8,* 1–7.
[36] Zhang, X., Julian, R. R. (2013) Exploring Radical Migration Pathways in Peptides with Positional Isomers, Deuterium Labeling, and Molecular Dynamics Simulations, *J. Am. Soc. Mass Spectrom. 24*, 524–533.
[37] 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.
[38] 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.
[39] Robinson, N. E., Robinson, A. B. (2001) Deamidation of Human Proteins, *Proc. Natl. Acad. Sci. 98*, 12409–12413.
[40] Tao, W. A., Zhang, D., Nikolaev, E. N., and 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.*

[41] Clarke, S. (2003) Aging as War between Chemical and Biochemical Processes: Protein Methylation and the Recognition of Age-Damaged Proteins for Repair, *Ageing Res. Rev. 2*, 263–285.
[42] Moss, C. X., Matthews, S. P., Lamont, D. J., Watts, C. (2005) Asparagine Deamidation Perturbs Antigen Presentation on Class II Major Histocompatibility Complex Molecules, *J. Biol. Chem. 280*, 18498–18503.
[43] Verma, A., Ngundi, M. M., Burns, D. L. (2016) Mechanistic Analysis of the Effect of Deamidation on the Immunogenicity of Anthrax Protective Antigen, *Clin. Vaccine Immunol. 23*, 396–402.
[44] Doyle, H. A., Gee, R. J., Mamula, M. J. (2007) Altered Immunogenicity of Isoaspartate Containing Proteins, *Autoimmunity 40*, 131–137.
[45] Liu, Y. D., van Enk, J. Z., Flynn, G. C. (2009) Human Antibody Fc Deamidation in Vivo, *Biologicals 37*, 313–322.
[46] Fuguet, E., Reta, M., Gibert, C., Rosés, M., Bosch, E., Ràfols, C. (2008) Critical Evaluation of Buffering Solutions for pK_a Determination by Capillary Electrophoresis, *Electrophoresis 29*, 2841–2851.
[47] Chen, S., Hong, Y., Liu, Y., Liu, J., Leung, C. W. T., Li, M., Kwok, R. T. K., Zhao, E., Lam, J. W. Y., Yu, Y., Tang, B. Z. (2013) Full-Range Intracellular pH Sensing by an Aggregation-Induced Emission-Active Two-Channel Ratiometric Fluorogen, *J. Am. Chem. Soc. 135*, 4926–4929.
[48] Conrad, U., Fahr, A., and Scriba, G. K. E. (2010) Kinetics of aspartic acid isomerization and enantiomerization in model aspartyl tripeptides under forced conditions. *J. Pharm. Sci. 99*, 4162–4173.
[49] Thapar, N., Griffith, S. C., Yeates, T. O., Clarke, S. (2002) Protein Repair Methyltransferase from the Hyperthermophilic Archaeon *Pyrococcus Furiosus*: Unusual Methyl-Accepting Affinity for D-Aspartyl and N-Succinyl-Containing Peptides, *J. Biol. Chem. 277*, 1058–1065.
[50] Li, X., Cournoyer, J. J., Lin, C., O'Connor, P. B. (2008) Use of ^{18}O Labels to Monitor Deamidation during Protein and Peptide Sample Processing, *J. Am. Soc. Mass Spectrom. 19*, 855–864.
[51] Krokhin, O. V., Antonovici, M., Ens, W., Wilkins, J. A., Standing, K. G. (2006) Deamidation of -Asn-Gly- Sequences during Sample Preparation for Proteomics: Consequences for MALDI and HPLC-MALDI Analysis, *Anal. Chem*. 78, 6645–6650.
[52] Liu, S., Moulton, K. R., Auclair, J. R., Zhou, Z. S. (2016) Mildly Acidic Conditions Eliminate Deamidation Artifact during Proteolysis: Digestion with Endoprotease Glu-C at pH 4.5, *Amino Acids 48*, 1059–1067.
[53] Siggaard-Andersen, O., Fogh-Andersen, N. (1995) Base Excess or Buffer Base (Strong Ion Difference) as Measure of a Non-Respiratory Acid-Base Disturbance, *Acta Anaesthesiol. Scand. 39*, 123–128.
[54] Goolcharran, C., Stauffer, L. L., Cleland, J. L., Borchardt, R. T. (2000) The Effects of a Histidine Residue on the C-Terminal Side of an Asparaginyl Residue

on the Rate of Deamidation Using Model Pentapeptides, *J. Pharm. Sci. 89*, 818–825.

[55] 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.

[56] Ly, T., Julian, R. R. (2008) Residue-Specific Radical-Directed Dissociation of Whole Proteins in the Gas Phase, *J. Am. Chem. Soc. 130*, 351–358.

CHAPTER 3

Ultraviolet Photodissociation Initiated Radical Chemistry Enables Facile Glycan Isomer Identification

Abstract

Glycans are fundamental biological macromolecules, yet despite their prevalence and recognized importance, a number of unique challenges hinder routine characterization. The multiplicity of OH groups in glycan monomers easily afford branched structures and alternate linkage sites, which can result in isomeric structures that differ by minute details. Herein, radical chemistry is employed in conjunction with mass spectrometry to enable rapid, accurate, and high throughput identification of a challenging series of closely related glycan isomers. The results are compared with analysis by collision-induced dissociation, higher-energy collisional dissociation, and ultraviolet photodissociation (UVPD) at 213 nm. In general, collision-based activation struggles to produce characteristic fragmentation patterns, while UVPD and radical-directed dissociation (RDD) could distinguish all isomers. In the case of RDD, structural differentiation derives from radical mobility and subsequent fragmentation. For glycans, the energetic landscape for radical migration is flat, increasing the importance of three-dimensional structure. RDD is therefore a powerful and straightforward method for characterizing glycan isomers.

Introduction

Glycans are ubiquitous biomolecules that serve vital functions across the entire spectrum of life on Earth.[1] Despite their simple origins, arising from carbon dioxide and water to form simple sugars *via* photosynthesis, assembled glycans can be exceptionally diverse and complex. Unlike DNA and proteins, which assemble in a consistent fashion into linear polymers formed from chemically distinct monomers, glycans can adopt branched structures with a diverse regio- and stereochemistry. For example, a standard hexose has five highly similar connectivity sites, each with isomeric linkage configurations.[2] As a result of these extra dimensions of structural variability, glycan complexity scales far more rapidly than that of DNA or proteins.

Changes to any stage of glycan assembly can impact the assembled polymer properties. As an example, the most abundant organic compound on Earth, cellulose, is an inedible component of cell walls,[3] while starch, which provides the bulk of human energy needs,[4] differs only by the configuration of the glycosidic bonds linking their glucose monomers. Although cellulose and starch neatly demonstrate the importance of glycan polymorphism, these isolated carbohydrate repeats represent only a small subset of biologically relevant glycans.

Commonly, glycans are attached to cell surfaces or biomolecules to enable or modulate a certain function.[5] Perhaps the most well-known example of this is the ABO blood group system, which originates from the glycan structure of antigens on the surface of red blood cells; these glycans govern an individual's ability to

receive a blood transfusion. Arguably, the ABO system was the first form of personalized medicine, predating the human genome project by a full century.[6] Although these extra dimensions of structural variability enable remarkable diversity in chemical characteristics, this comes with the cost of complicating analysis. Consequently, advances in the field of glycomics, which comprises the study of isolated and conjugated glycan biomolecules, have not kept pace with the closely related fields of proteomics and genomics due to paucity of analytical tools capable of handling this extra complexity.1

Glycan structure is defined by the monomeric sugar building blocks in the same manner that DNA and proteins are dependent on nucleic acid and amino acid content, respectively. However, glycans can form branched structures with a more a complex regio- and stereochemistry. **Figure 3.1** depicts a succinct family of oligomers with systematic variations in connectivity, configuration, and composition. Glycan analysis is convoluted by the isomeric nature of the building blocks, which results in compositional isomers (*i.e.*, isomers **1** and **2** as well as **4** and **5**). Further complication arises because there are two possible configurations at each anomeric center, α and ß (see isomers **2** and **3** as well as **5** and **6**). Finally, the monomers can be linked with a different regiochemistry for example by (1→3) or (1→4) glycosidic bonds (compare isomers **1-3** with **4-6**). This collection of oligomers represents an extreme set of isomers and serves as a rigorous benchmark for evaluating new analytical methods.[7,8]

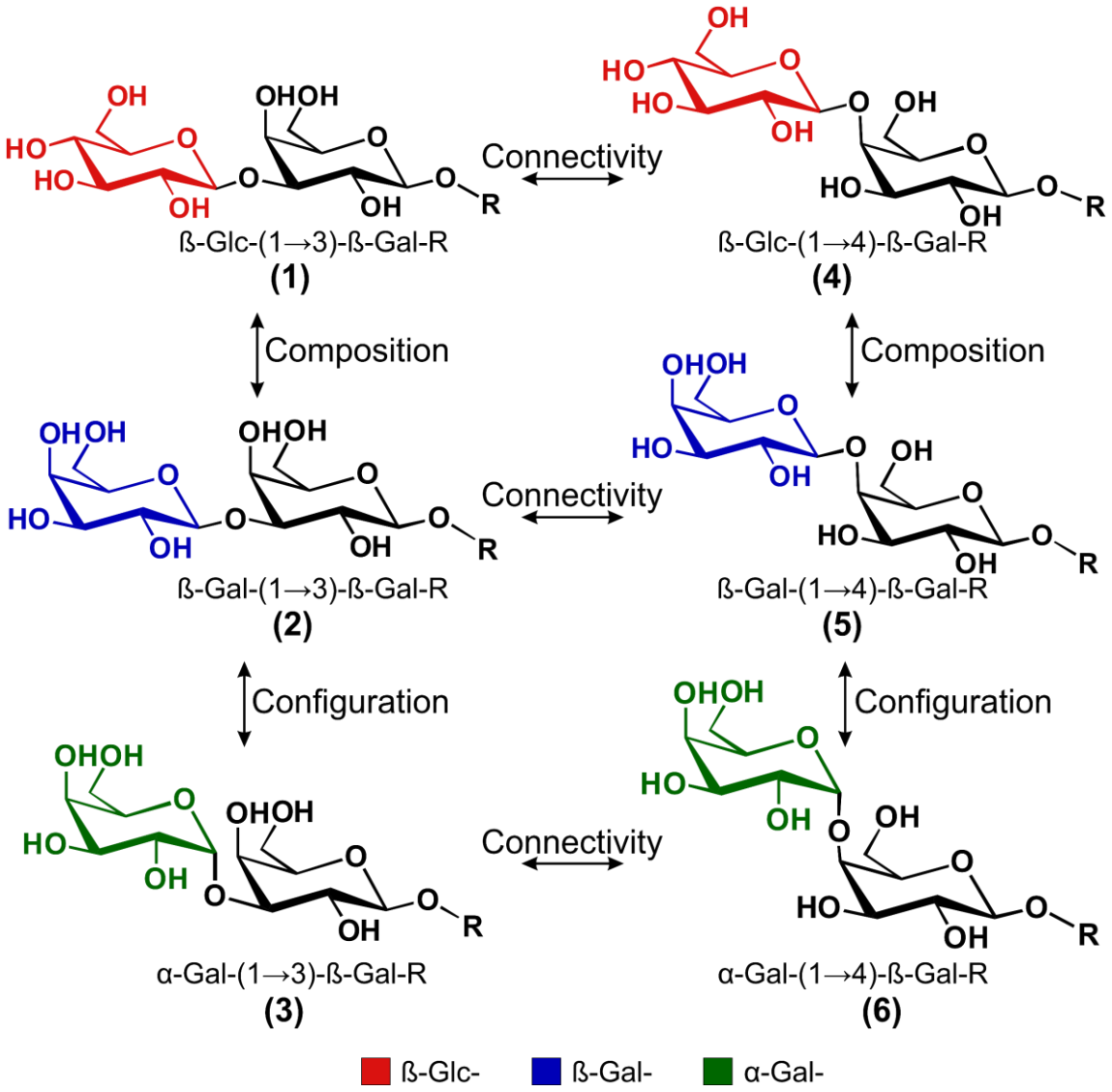


Figure 3.1. A succinct family of glycan isomers that differ in one or more of three structural aspects: connectivity, configuration, and composition. To simplify the structural comparisons, the R- group containing a ß-(1→4) glucose and C_5H_{10}-NH_3 linker that remains consistent between the oligomers is not shown.

An ideal analytical technique should be able to differentiate all isomers shown in Figure 3.1 in a rapid, robust, and high-throughput manner. Although, nuclear magnetic resonance (NMR) has been successfully employed for glycan structural analysis, the characterization is often incomplete, relatively slow, and requires large amounts of purified material.[9,10,11] Mass spectrometry (MS) offers advantages in speed and sensitivity, and tandem mass spectrometry-based fragmentation (MS^n) excels at sequence determination in peptides.[12] MS has also shown tremendous promise in glycan structural elucidation,[13,14] but extending characterization beyond sequence information often requires odd-electron dissociation such as electron capture dissociation,[15] electron transfer dissociation,[16,17] electronic excitation dissociation,[18] or electron detachment dissociation.[19] Furthermore, glycans exhibit poor adsorption behavior on conventional reversed phase LC columns[20] which limits the potential for meaningful separation. Recently, by utilizing porous graphitized carbon and nanoflow liquid chromatography, the Lebrilla group has achieved quantitative analysis of glycopeptides, including several isomers.2 Alternatively, coupling ion mobility to mass spectrometry has been recognized as a powerful tool for differentiating isomers without the need for solution phase separation.[21-24] Of particular relevance to the current study, the recent ion mobility report by Pagel *et al.* revealed that differences in connectivity and configuration result in changes to three-dimensional shape, enabling isomeric glycan separation,8 however compositional isomers remained indistinguishable.

In the absence of separation, a variety of fragmentation-based methods have proven valuable in the structural characterization of glycans. Ultraviolet photodissociation (UVPD) has been demonstrated to improve sequence coverage of linear[25] and various branched glycans,[26] in addition to identifying native glycan isomers from ovalbumin.[27] Similarly, tunable Infrared multiple photon dissociation (IRMPD) has been utilized to determine disaccharide linkage configuration,[28] as well as *de novo* sequencing.[29] Improvements to IRMPD in regard to spectral resolution can be made by utilizing cold-ion infrared spectroscopy vibrational fingerprinting,[30,31] which can distinguish even the most similar compositional isomers, but the method also requires specialized equipment that is not widely available.7 Finally, the Beauchamp group has pioneered a dual reagent approach that enables acid-catalyzed glycan sequencing and radical-based discrimination of several glycan isomers.[32,33] These advances, and others, have positioned MS at the center of glycan profiling efforts even though complete characterization remains challenging.[34]

The present work seeks to explore the utility of novel fragmentation methods for discriminating glycan isomers. Radical directed dissociation (RDD), which generates a radical upon UV laser photodissociation inside a mass spectrometer,[35] has been previously demonstrated to have excellent structural sensitivity.[36-38] In RDD, a radical is generated in a site-specific manner, making subsequent radical migration *via* hydrogen abstraction highly sensitive to three-dimensional structure. RDD can readily distinguish all isomers in Figure 3.1. Comparisons are made

between 213 nm UVPD, RDD, and traditional dissociation techniques, such as collision-induced dissociation (CID) and higher-energy collisional dissociation (HCD). Finally, the mechanistic and energetic basis for differences in fragmentation are elucidated.

Experimental Section

Synthesis of oligosaccharides **1-6** was performed using automated solid-phase glycan assembly.[39] Briefly, different compositions are generated by the sequential assembly of building blocks. Each assembly step consists of a defined procedure involving coupling, washing and deprotection. The connectivity is controlled by the choice of protecting groups that exclusively enable the formation of a glycosidic bond at a specific position. Configurational selectivity is aided by participating protecting groups. After the synthetic route is completed, the crude molecules are cleaved from the resin, confirmed by MS- and NMR-analysis and additionally validated via HPLC. Before and after deprotection, the oligosaccharides were purified by preparative HPLC and again confirmed using the analytical techniques described above.

A simple nanoelectrospray source was built in-house to accommodate small volumes of glycan analytes. A ~10 cm length of silica (360 µm outer diameter x 100 µm inner diameter) is pulled to a ~5 µm emitter using a Sutter P-2000 laser tip puller. Empty emitters are loaded by applying vacuum to one end of the emitter while the opposite end is submerged in sample. The emitter is housed in a stainless-steel union and 1-2 kV is applied to the conductive union. A second capillary is joined in the union to provide a reservoir of solvent which prevent air from entering the emitter. Upon applying voltage, the sample electrosprays spontaneously (i.e., without a mechanical pump) at a rate of ~50 nL/min. Underivatized oligomers were electrosprayed from a solution of 50/50/1

methanol/water/acetic acid (v/v) and analyzed as $[M+H]^+$ cations. Modified oligomers were analyzed as $[M+Na]^+$ cations.

Oligomers were modified with *para*-iodobenzoate (4IB) *via* amide bond formation between the linker amine to the 4IB carboxylic acid, to provide a photocleavable C-I bond for radical directed dissociation. Succinimidyl-4IB was synthesized following previously established protocols.35 Oligomer modification was carried out in 1.5 mM borate buffer at pH 8.6 in a 37 °C incubator. Succinimidyl-4IB was added from a 2.5 mM stock at a molar ratio of ~25:1. The majority of the oligomer is found to be modified within one hour. The modified oligomers were diluted into 50/50/1 methanol/water/acetic acid (v/v) and electrosprayed without further work-up.

RDD was performed using a Thermo Fisher LTQ linear ion trap. The LTQ has been modified with a quartz window to allow fourth harmonic (266 nm) pulses from a Continuum Minilite II Nd:YAG laser to irradiate the trapped ion cloud. The modified analyte is isolated in an MS^2 step and irradiated with a single pulse from the laser. The radical analyte generated *via* photodissociation is subsequently isolated in an MS^3 step and activated *via* CID with a relative energy of 15. UVPD was performed on a Thermo Fisher Orbitrap Velos Pro. The orbitrap has been modified with a quartz window to allow photodissociation experiments on ions trapped in the HCD cell using a 213 nm laser from CryLas GmbH operating at 1 kHz. UVPD was performed with 100 laser pulses. Traditional CID and HCD was also performed on

the Orbitrap with relative collision energies of 25 and 30 respectively. During CID and HCD operation, the laser is not in use.

Oligomers are distinguished using MS^n fragmentation patterns. To quantitatively determine the differences between two spectra, R_{isomer} values are calculated using Equation 1 where R_1 and R_2 refer to the ratios of two pairs of fragment ions that vary the most in relative abundance between different MS^n spectra. Identical fragmentation patterns result in R_{isomer} values of 1 thereby indicating no discrimination, while larger values reflect a higher degree of discrimination.

$$R_{isomer} = \frac{R_1}{R_2} \qquad (1)$$

Bond dissociation energy calculations were performed using Gaussian 09 Rev. B.01.[40] Structural optimizations and energy calculations were performed with hybrid density functional theory B3LYP at the 6-31G(d) basis set. All BDEs were calculated by the isodesmic reaction method using experimentally determined BDEs.[41,42] Galactose was modelled as shown in Figure 3.5, with a simple ether representing the glyosidic linkage.

Results and Discussion

We begin by examining the most common MS/MS method, CID. Collisional activation of protonated glycans **1**-**6** yields a limited number of fragments that are found in similar abundances, (**Figure 3.2**, left panel) in agreement with previous results.[43] Disappointingly, the fragmentation patterns for all six oligomers are quite similar. The spectra are dominated by the loss of one and two galactose rings, observed at 428 and 266 *m/z* respectively. Fragmentation sites and the resulting fragment-ion *m/z* values for each isomer are shown on the right side of Figure 3.2. Less abundant water losses are also noted, but the origins of these fragments are ambiguous due to the high number of hydroxyl groups in any given saccharide. Higher-energy collisional dissociation (HCD) yields similar fragmentation pathways, with a preference for the loss of two rings (**Figure 3.2**, middle). Because glucose and galactose are isobaric, each resulting in 162 Dalton (Da) losses, the fragmentation patterns for **4** and **5** are indistinguishable. Similarly, comparisons between **5** and **6** indicate that the losses are insensitive to the anomeric configuration.

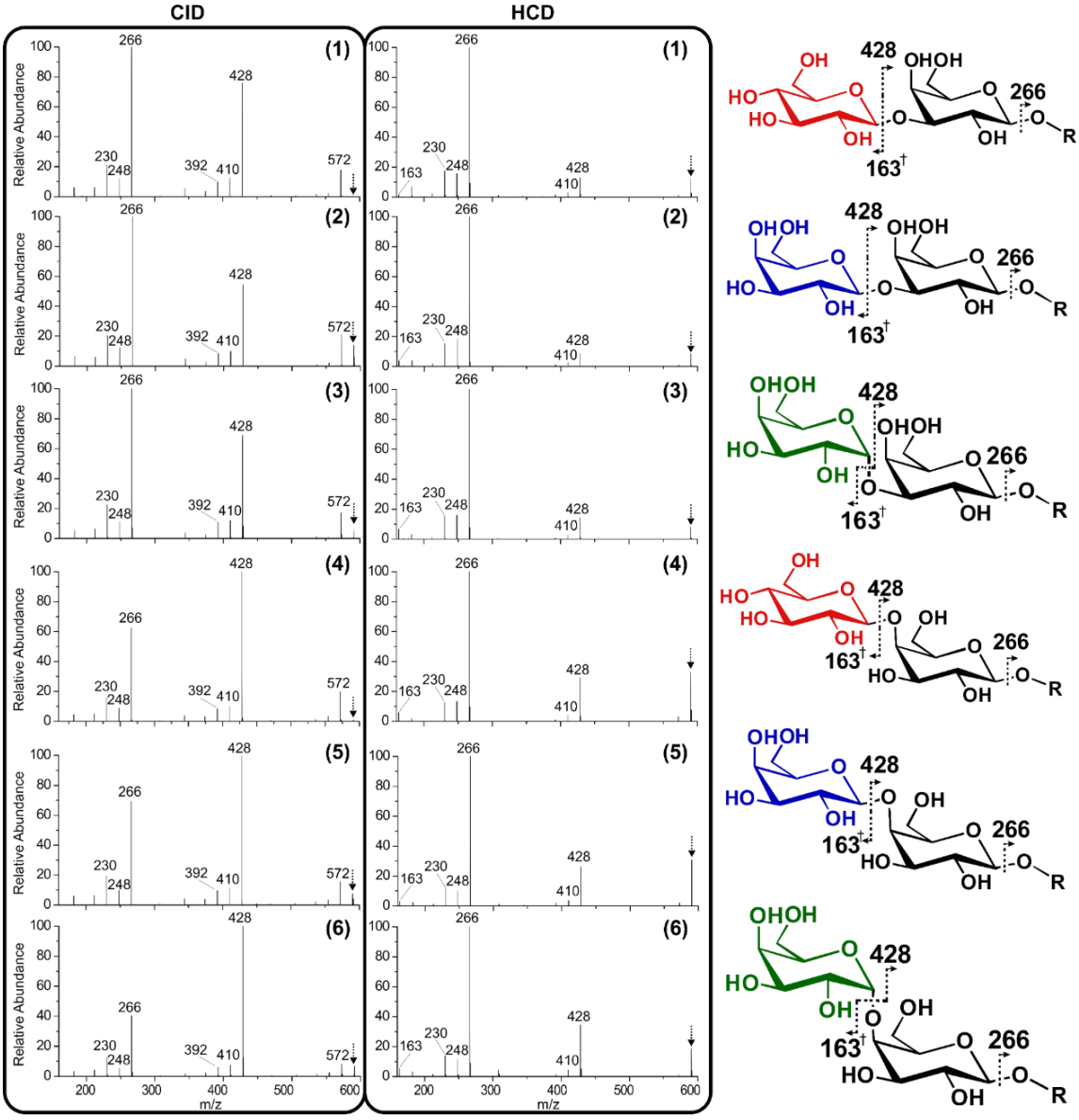


Figure 3.2. Collisional activation fragmentation spectra (CID and HCD) of isomers **1**-**6** and peak identifications for each oligomer. † indicates a fragment containing a double bond. The $[M+H]^{+1}$ precursor observed at 590 *m/z* is indicated with an arrow.

Despite the similarity between spectra in Figure 3.2, there are subtle differences in the relative intensities for some fragments in both CID and HCD. For example, within the CID spectra, the ratio of 428 *m/z* to 266 **m/z** is inverted between **2** and **6**. Changes such as this can be used to quantitate the difference between similar spectra. Commonly, the ratio of fragment intensities that change most between spectra are used to discriminate isomers using R-values.[44] R-values reflect the degree of difference for the fragment ions that vary the most in intensity between a pair of independently collected fragmentation spectra (the equation is provided in the Experimental section). Higher R-values indicate greater differences between spectra. We have previously established the statistically significant R thresholds for CID fragmentation.[36] Given these thresholds, some of the isomers in Figure 3.2 should be distinguishable by CID or HCD. For instance, comparing the 428 *m/z* and 266 *m/z* peaks for **2** and **6** yields a R-value of 4.5, which is well above the threshold of 1.9 for CID. Conversely, the largest difference between **2** and **3** results in an R-value of 1.6, which is insignificant (see the SI for a complete list of R-values). To simplify comparison between different fragmentation methods, we report the minimum R-value after cross-comparison of all isomers, which indicates whether all isomers can be distinguished, and the median value which reflects the magnitude of differences typically observed. The minimum R-value for CID is 1.6, and the median is 2.1. These values indicate minimal capability for distinguishing isomers with CID. Interestingly, HCD fares slightly better at discriminating the oligomers even though similar fragment ions are produced. Some of the difference

is attributed to production of a protonated sugar monomer, observed at 163 *m/z*, which is not observed during CID. For HCD, the minimum R-value is 1.4, while the median is 3.1. Neither technique was able to discern all six isomers. Furthermore, quantitation in mixtures would likely not be possible with such low R-values.[45]

The scarcity of fragments and overall similarity between CID and HCD point to a fundamental limitation of collisional activation, which often relies on a mobile proton to facilitate fragmentation.[46] The limited number of fragmentation sites observed for the oligomers suggest that the glycosidic bonds are susceptible to proton mediated fragmentation, while the remainder of the molecule is relatively inert. This may be due, in part, to the polycyclic nature of glycan oligomers, which requires a minimum of two cleavages to observe cross-ring fragmentation. In effect, the oligomers are resistant to double fragmentation by CID and HCD.

UVPD is an emerging dissociation technique[47,48] that can access both vibrational and excited-state electronic dissociation.[49] In order to utilize UVPD and radical dissociation, the oligomers were modified with *para*-iodobenzoate (4IB), which favors loss of iodine upon UV excitation.[35] The 213 nm UVPD spectra for 4IB modified oligomers is presented in the left panel of Figure 3.3. UVPD was carried out with 100 shots of 213 nm light at ~1.5 μJ/pulse, resulting in ~50% photodissociation yield. Several differences between the oligomers are visually apparent, with the α- linked isomers, **3** and **6** yielding intense -I• peaks at 715 *m/z* while all other oligomers preferentially generate the 554 *m/z* fragment. Less abundant cross-ring cleavage fragments are also observed at 582 and 420 *m/z* for

several isomers, along with the less favorable glycosidic bond cleavage products at 536 and 537 *m/z*. Importantly, the relative abundance of -I•, and radical directed fragmentation pathways varies greatly among the different isomers resulting in an R-value minimum of 5.9, and a median of 149.7. This represents a substantial improvement in isomer discrimination for every comparison relative to CID and HCD (see SI for complete score comparisons).

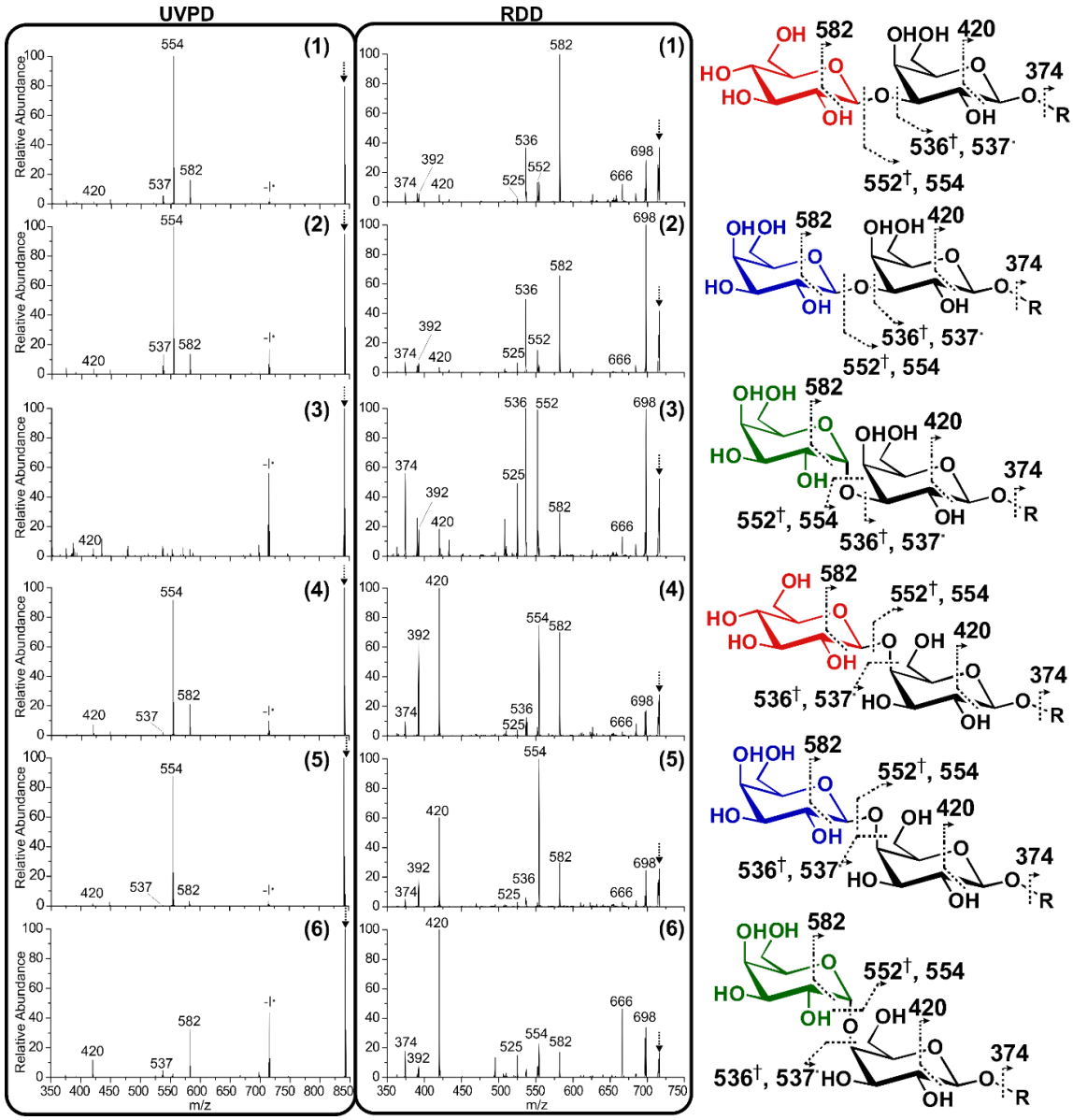


Figure 3.3. Radical fragmentation spectra afforded by UVPD and RDD. Peak identifications for each oligomer. The dot indicates radical products, the dagger indicates that the product forms a double bond. The [4IB M+Na]$^{+1}$ precursor observed at 842 *m/z* during UVPD, and the -I$^{\bullet}$ radical precursor observed at 715 *m/z* during RDD, is indicated with an arrow.

213 nm irradiation preferentially yields homolytic C-I bond cleavage (labeled -I•), but several oligomers undergo spontaneous subsequent fragmentation, indicating that the radical is formed close to the cleavage site, enabling facile migration. Radical migration occurs *via* hydrogen abstraction, and even subtle differences in the isomer structure lead to pronounced differences in the UVPD spectra. For instance, the abundance of the 554 *m/z* fragment, which arises from loss of one galactose ring and the glycosidic bond oxygen, is found to be very favorable for ß-, but not α-, glycosidic bonds (note the low abundance of 554 for α-oligomers **3** and **6**. These data indicate that the linker configuration has a considerable impact on the radical mobility, which is governed by the three-dimensional structure and proximity of abstractable hydrogen atoms. Although additional cleavage products from either side of the glycosidic linkage are also present, the abundances are generally low.

Collisional activation of the radical generated by loss of iodine following UVPD yields an RDD spectrum. Even a cursory examination of the spectra afforded by RDD (**Figure 3.3** middle panel) reveals striking differences in fragment intensities and obvious isomer discrimination. RDD yields far greater abundance and variety of radical-derived fragments relative to UVPD alone. Collisional activation gradually heats the radical ion, thereby enabling migration and exploration of the proximal structural features. Importantly, the relative abundances of fragment ions vary significantly between isomers, often by more than an order of magnitude. These remarkable differences enable unambiguous discrimination of each isomer.

Importantly, RDD yields a minimum R-value of 10.9 and a median of 170.9. Again, these values represent significant improvements over collisional activation, and notable increases over UVPD.

Furthermore, all six isomers can be fully distinguished by examining just four RDD peaks, as shown in Figure 3.4. Many isomers exhibit quite distinct 552:554 peak ratios, with (1→4) linked oligomers preferentially forming the 554 Da fragment, while the (1→3) linkage provides a competitive pathway for generating the 552 Da fragment. Unsurprisingly, comparisons between the compositional isomers (**1** vs. **2**) and (**4** vs. **5**), which differ only by the orientation of a single hydroxyl group, are the most difficult to discern. However, isomers **1** and **2** are readily distinguished by the relative abundance of the 582 and 698 *m/z* fragments. When comparing **4** vs. **5**, the nearby 582 ring cleavage fragment can be monitored to supplement the 552:554 ratio. Oligomer **4** is characterized by a 554:582 ratio of ~1, while oligomer **5** produces a ratio of 3.6. Together, these three fragments readily identify and discriminate each of the oligomers, and further hint at the mechanistic basis of isomeric discrimination. Proposed mechanisms for each of these diagnostic fragments are provided in the SI.

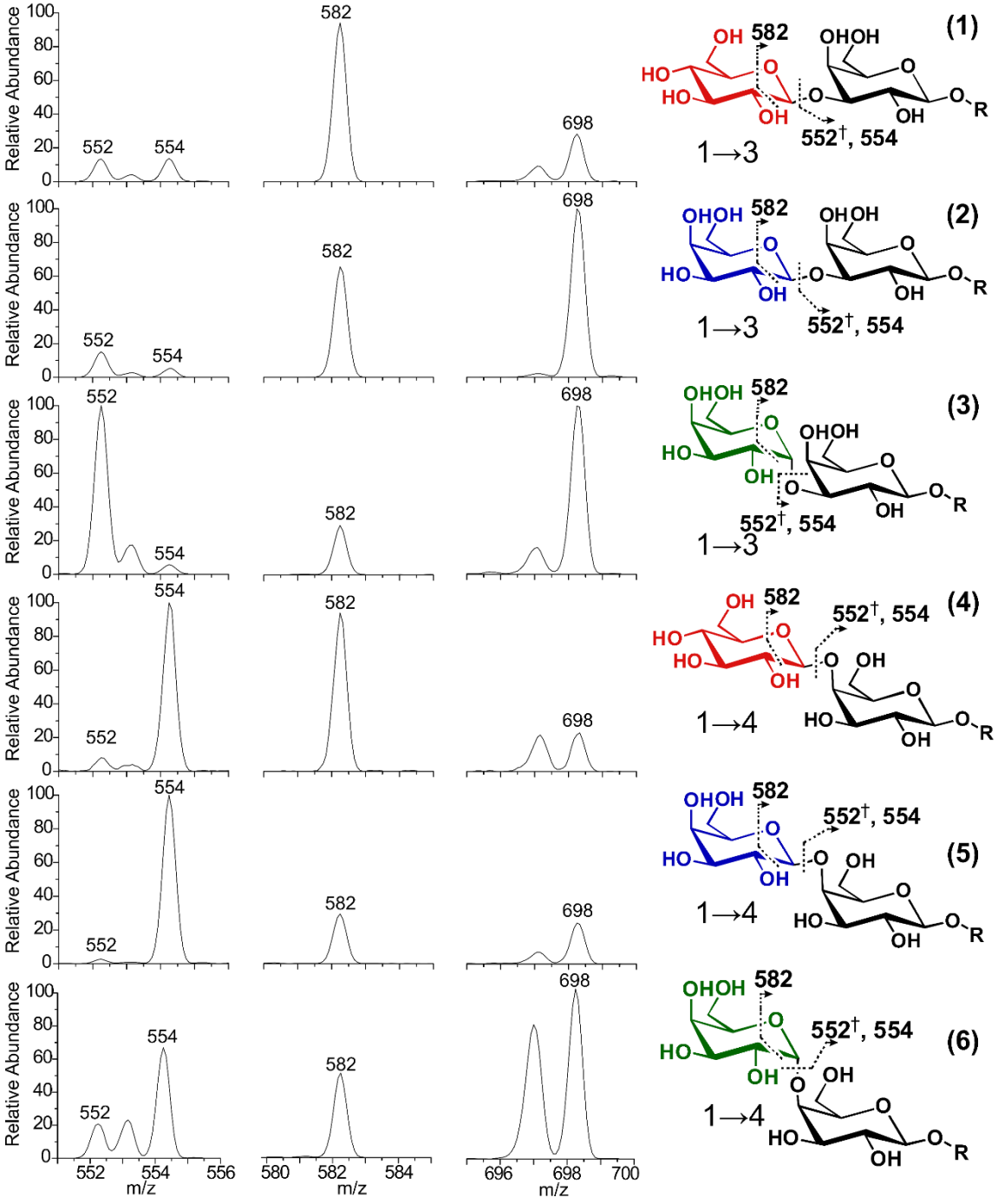


Figure 3.4. RDD fragmentation discriminates oligomers based on fragmentation abundances. The dagger indicates that the observed fragment contains a double bond.

Many of the R-values produced by UVPD and RDD represent some of the largest we have observed, and are far greater than typical values obtained for peptide isomers.[44,45,50,51] Although initially perplexing, we posit that the relative chemical uniformity exhibited by glycans (relative to peptides) impacts hydrogen abstraction and therefore radical migration. Two of the most important parameters that govern hydrogen abstraction are bond dissociation energy and sterics. For migration to occur, the radical and hydrogen must be in close proximity, and the bond dissociation energy (BDE) must be sufficiently low. Distance and sterics are primarily functions of three-dimensional structure, while the BDE is influenced by local chemical environment (*i.e.*, the functional group). In contrast to peptides, which are composed of diverse chemical side chains, glycans are made up of relatively few unique functional groups.

To explore potential radical migration pathways, we calculated the C–H BDEs for every abstractable hydrogen in galactose, and the results are shown in Figure 3.5. In short, isodesmic reactions were used with quantum mechanical calculations at the B3LYP/6-31G(d) level of theory, as described previously.[52] In contrast to peptides, where BDEs vary from 319 kJ mol^{-1} to 527 kJ mol^{-1}, the range for galactose is only 372 kJ $mol^{-1}$to 396 kJ mol^{-1}. This indicates that with respect to hydrogen abstraction, the energetic landscape is relatively flat for glycans. Therefore, every C-H bond represents a competitive site for migration and offers an energetically favorable hydrogen for abstraction when accessible by the 4IB radical, which has a BDE of 474 kJ mol^{-1}.[52] This effectively diminishes the

energetic considerations for radical migration within glycans, thereby elevating the importance of three-dimensional structure.

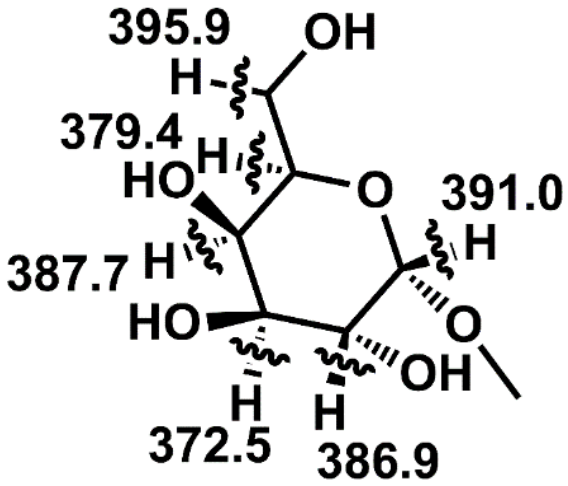


Figure 3.5. Bond dissociation energies (kJ mol^{-1}) for C-H bonds in galactose.

Collectively, these results demonstrate the remarkably utility of RDD in discriminating a full suite of isomeric oligomers. In contrast to traditional dissociation techniques, which exhibit a limited selection of favorable dissociation channels, radical migration proceeds to numerous sites, with hydrogen abstraction being dictated by oligomer structure. This structural sensitivity leads to unparalleled isomer discrimination while retaining the rapid and high throughput nature of mass spectrometry analysis. Importantly, the radical chemistry depicted can be carried out using 213 nm solid state lasers that are currently available on commercial mass spectrometers.

Conclusions

Both radical directed dissociation and UVPD exhibit excellent isomer discrimination for a complete family of glycan oligomers. The radical migration pathways and bond fragmentation tendencies reveal specific structural alterations that emerge as a function of glycan connectivity, configuration, and composition. Importantly, isomeric discrimination is accomplished *via* mass-spectrometry based fragmentation which is rapid, robust, and requires minimal amounts of sample. This work directly addresses the long-term roadmap goal of glycan structural determination established by the National Academy of Sciences1 by providing a high-throughput methodology that is accessible to non-experts and can be implemented in a wide variety of labs. Furthermore, demonstrating that this established proteomic approach is amenable to glycomics, and can be implemented on commercially available instruments, may ease the union of these two closely related but often isolated fields.

Acknowledgements

The authors are grateful for funding from the NIH (NIGMS grant R01GM107099 to RRJ). JH, HSH, PHS and KP thank the Max Planck Society and the Freie Universität Berlin for financial support.

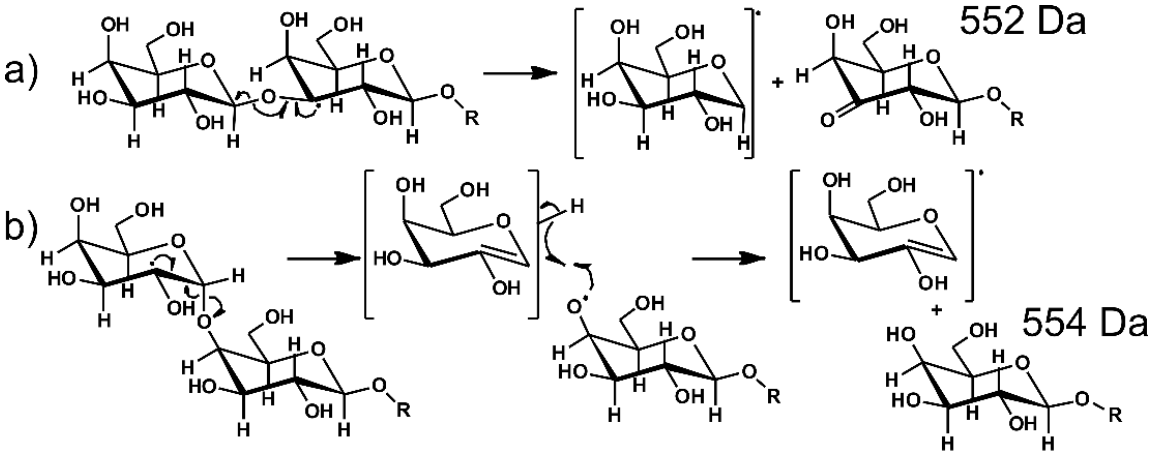
Supporting Information

Radical Fragmentation Mechanisms:

Radical directed dissociation (RDD) is highly dependent on three-dimensional structure. Many potential dissociation pathways exist for molecules with the structural complexity of the glycans studied herein. Below we provide a few pathways that are likely important in the present experiments to serve as illustrative examples of the type of chemistry that yields structural selectivity. **Scheme 3.1a** depicts one likely mechanism that results in glycosidic bond cleavage yielding the 552 Da fragment, which is preferentially generated for (1→3) linked isomers (**2** is shown). Radical migration begins as hydrogen is abstracted from the internal galactose ring adjacent to the glycosidic linkage. Subsequently, the linkage is broken as the C-4 carbon of the internal sugar forms a double bond with the linker oxygen resulting in the observed 552 Da product.

The mechanism to produce the 554 Da fragment, which is the dominant product for (1→4) linked isomers, is illustrated in **Scheme 3.1b** where **6** is shown. There are redundant pathways to produce the 554 Da product, so we have depicted just one possible mechanism. In this case, the first hydrogen is abstracted from the terminal sugar at the C-2 position. The rings are cleaved from one another as a double bond forms between C-1 and C-2 of the terminal sugar. Finally, a second hydrogen is abstracted from the terminal sugar, resulting in the 554 Da fragment. Although it is apparent that the second hydrogen is abstracted from the terminal

sugar ring, the exact location is ambiguous and is therefore drawn in a generic manner.



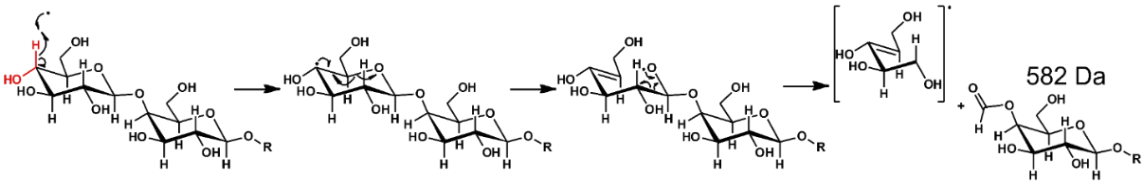
Scheme 3.S1. Radical directed dissociation mechanisms that cleave the glycosidic linker to produce 552 and 554 Da fragments.

Although both pathways are initiated by hydrogen atom abstraction, the origin of the hydrogen atom is distinct. In the case of the 554 Da fragment, a hydrogen is first abstracted from the terminal carbohydrate ring (furthest in sequence from the 4IB chromophore where the radical is generated *via* photodissociation). In contrast, the 552 fragment is produced following hydrogen abstraction from the internal carbohydrate ring. Considering that both (1→3) and (1→4) isomers generate 552 and 554 *m/z* products (albeit in vastly different abundances), it is apparent that the discrepancies in fragmentation abundance are a function of three-dimensional structure rather than bond-strength chemistry. This notion is

supported by the quantum mechanical calculations, which indicate the C-H bond dissociation energies are comparable at each position on the ring (see Figure 3.5 in the main text). Observing that (1→3) linked oligomers abstract hydrogen from the terminal ring less frequently than their (1→4) counterparts suggests the final ring may be too distant from the 4IB radical starting location. Conversely, the (1→4) linked oligomers preferentially abstract from the terminal carbohydrate ring, indicating that the terminal ring must be close in space to 4IB (i.e., the structure is compact). This radical migration preference is in good agreement with previous ion mobility data, which showed the (1→4) linked oligomers are more compact than the (1→3) isomers. (Hofmann, J., Hahm, H. S., Seeberger, P. H., and Pagel, K. (2015) Identification of carbohydrate anomers using ion mobility-mass spectrometry. *Nature 526*, 241–244.)

A similar discrepancy in the availability of a given hydrogen for abstraction is witnessed when comparing the compositional isomers, **4** and **5**. While these oligomers are known to be exceptionally difficult to discern, resulting in overlapping collisional cross sections and indistinguishable collisionally activated fragmentation spectra, the structural sensitivity of radical migration is able to tease out the minute structural difference. The two isomers differ only in the relative orientation of the hydrogen and hydroxyl group on C-4 of the monosaccharide at the non-reducing end. **Scheme 3.2** depicts **4** and illustrates how the relative orientation of a single hydrogen impacts production of the cross-ring cleavage products observed at 582 *m/z*. Oligomer **4** terminates with glucose, which has an

axial hydrogen (indicated by an ‘up’ bond in red) while oligomer **5** terminates in galactose, where the hydrogen would be equatorial. The RDD fragmentation behavior indicates that the hydrogen is more easily abstracted in the axial position, leading to formation of a double bond between C4 and C5 with concurrent homolytic cleavage of the C5-O bond. The radical oxygen formed by this bond cleavage subsequently forms a double bond with C1 as the C1-C2 bond breaks. The net result is a fractured glucose ring, with C1 and O remaining bound to the oligomer, while the remainder of the ring is lost as a radical fragment.



Scheme 3.S2. The radical directed cross ring cleavage mechanism that produces the 582 Da fragment is affected by the relative orientation of an axial versus equatorial hydrogen.

Due to the nature of the problem in discriminating highly similar isomers, the results are, of necessity, relevant to particular structures. Therefore, we expect that the discrimination afforded by RDD and UVPD must be founded on empirical data rather than categorical assignments (ie fragmentation trends are unlikely to be generalizable across distinct sets of isomers). While we have observed consistent trends within this family of isomers, it is likely that these trends reflect the high degree of structural similarity within this glycan family, rather than a more universal

structural constraint imposed by any given isomerism. For instance, we would not suggest that all (1→4) linked oligomers are more compact than their (1→3) isomers. The methodology presented herein may best lend itself to direct comparisons between analytes to positively identify isomers. Finally, based on the magnitude of differences between the fragmentation spectra, and the high signal-to-noise, we expect that a spectral library may prove fruitful for isomer identification.

R-Values:

CID R-Values						
Oligomer #	**1**	**2**	**3**	**4**	**5**	**6**
1		1.65	1.66	2.11	1.90	3.27
2	1.65		1.57	2.93	2.64	4.64
3	1.66	1.57		2.31	2.09	3.59
4	2.11	2.93	2.31		1.60	2.40
5	1.90	2.64	2.09	1.67		1.85
6	3.27	4.64	3.59	2.36	1.85	

HCD R-Values						
Oligomer #	**1**	**2**	**3**	**4**	**5**	**6**
1		2.02	2.34	3.11	3.22	3.69
2	2.02		2.10	4.68	5.63	6.46
3	2.34	2.10		2.49	4.02	3.37
4	3.11	4.68	2.49		1.89	1.38
5	3.22	5.63	4.02	1.89		1.74
6	3.69	6.46	3.37	1.38	1.74	

UVPD R-Values						
Oligomer #	**1**	**2**	**3**	**4**	**5**	**6**
1		5.90	346.56	105.38	7.40	337.00
2	5.90		167.04	48.13	37.17	149.72
3	346.56	167.04		2954.00	5039.00	37.28
4	105.38	48.13	2954.00		6.91	244.87
5	7.40	37.17	5039.00	6.91		454.96
6	337.00	149.72	37.28	244.87	454.96	

RDD R-Values						
Oligomer #	**1**	**2**	**3**	**4**	**5**	**6**
1		11.90	75.29	419.45	82.11	256.80
2	11.90		17.51	828.75	170.93	507.40
3	75.29	17.51		849.50	606.52	184.68
4	419.45	828.75	849.50		10.90	118.63
5	82.11	170.93	606.52	10.90		60.85
6	256.80	507.40	184.68	118.63	60.85	

Table 3.S1. R-values for each oligomer as obtained by each dissociation method.

Note: For demonstrative purposes, the main body of the text references an R-value of 4.5 which is obtained by comparing the peaks at 266 and 428 *m/z* when discussing CID for isomers **2** and **6**. However, comparing the peaks at 428 and 572 *m/z* results in a marginally higher R-value of 4.6 which is reported here.

References

[1] National Research Council. (2012) Transforming Glycoscience: A Roadmap for the Future.

[2] Hua, S., Nwosu, C. C., Strum, J. S., Seipert, R. R., An, H. J., Zivkovic, A. M., German, J. B., and Lebrilla, C. B. (2012) Site-specific protein glycosylation analysis with glycan isomer differentiation. *Anal. Bioanal. Chem. 403*, 1291–1302.

[3] Li, R., Fei, J., Cai, Y., Li, Y., Feng, J., and Yao, J. (2009) Cellulose whiskers extracted from mulberry: A novel biomass production. *Carbohydr. Polym. 76*, 94–99.

[4] Flatt, J.-P. (1995) of Carbohydrate and Fat3. *Am J Clin Nutr 61*, 952S–9S.

[5] Service, R. F. (2012) Looking for a sugar rush. *Science. 338*, 321–323.

[6] Schwarz, H. P. (2003) Historical review: Karl Landsteiner and his his major contributions to haematology. *Br. J. Haematol. 121*, 556–565.

[7] Mucha, E., González Flórez, A. I., Marianski, M., Thomas, D. A., Hoffmann, W., Struwe, W. B., Hahm, H. S., Gewinner, S., Schöllkopf, W., Seeberger, P. H., von Helden, G., and Pagel, K. (2017) Glycan Fingerprinting via Cold-Ion Infrared Spectroscopy. *Angew. Chemie - Int. Ed. 56*, 11248–11251.

[8] Hofmann, J., Hahm, H. S., Seeberger, P. H., and Pagel, K. (2015) Identification of carbohydrate anomers using ion mobility-mass spectrometry. *Nature 526*, 241–244.

[9] Duus, J., Gotfredsen, C. H., and Bock, K. (2000) Carbohydrate structural determination by NMR spectroscopy: modern methods and limitations. *Chem. Rev. 100*, 4589–4614.

[10] Pabst, M., and Altmann, F. (2011) Glycan analysis by modern instrumental methods. *Proteomics 11*, 631–643.

[11] Harvey, D. J. (2001) Identification of protein-bound carbohydrates by mass spectrometry. *Proteomics 1*, 311–328.

[12] Han, X., Aslanian, A., and Yates, J. (2008) Mass Spectrometry for Proteomics. *Curr Opin Chem Biol. 12*, 483–490.

[13] Hofmeister, G. E., Zhou, Z., and Leary, J. A. (1991) Linkage Position Determination in Lithium-Cationized Disaccharides: Tandem Mass Spectrometry and Semiempirical Calculations. *J. Am. Chem. Soc. 113*, 5964–5970.

[14] An, H. J., and Lebrilla, C. B. (2011) Structure elucidation of native N- and O-linked glycans by tandem mass spectrometry (tutorial). *Mass Spectrom. Rev.*

[15] Adamson, J. T., and Håkansson, K. (2007) Electron Capture Dissociation of Oligosaccharides Ionized with Alkali, Alkaline Earth, and Transition Metals. *Anal. Chem. 79*, 2901–2910.

[16] Wolff, J. J., Leach 3rd, F. E., Laremore, T. N., Kaplan, D. A., Easterling, M. L., Linhardt, R. J., and Amster, I. J. (2010) Negative electron transfer dissociation of glycosaminoglycans. *Anal Chem 82*, 3460–3466.

[17] Han, L., and Costello, C. E. (2011) Electron transfer dissociation of milk oligosaccharides. *J. Am. Soc. Mass Spectrom. 22*, 997–1013.
[18] Yu, X., Jiang, Y., Chen, Y., Huang, Y., Costello, C. E., and Lin, C. (2013) Detailed glycan structural characterization by electronic excitation dissociation. *Anal. Chem. 85*, 10017–10021.
[19] Wolff, J. J., Amster, I. J., Chi, L., and Linhardt, R. J. (2007) Electron Detachment Dissociation of Glycosaminoglycan Tetrasaccharides. *J. Am. Soc. Mass Spectrom. 18*, 234–244.
[20] Pabst, M., and Altmann, F. (2011) Glycan analysis by modern instrumental methods. *Proteomics 11*, 631–643.
[21] Mookherjee, A.; Guttman, M. *Curr. Opin. Chem. Biol.* **2018**, *42*, 86–92.
[22] Morrison, K. A.; Clowers, B. H. *Curr. Opin. Chem. Biol.* **2018**, *42*, 119–129.
[23] Chen, Z.; Glover, M. S.; Li, L. *Curr. Opin. Chem. Biol.* **2018**, *42*, 1–8.
[24] Hofmann, J.; Pagel, K. *Angew. Chemie - Int. Ed.* **2017**, *56* (29), 8342–8349.
[25] Devakumar, A.; Thompson, M. S.; Reilly, J. P. *Rapid Communications in Mass Spectrometry*. **2005**, 2313–2320.
[26] Devakumar, A.; Mechref, Y.; Kang, P.; Novotny, M. V.; Reilly, J. P. *Rapid Communications in Mass Spectrometry*. **2007**, 1452–1460.
[27] Devakumar, A.; Mechref, Y.; Kang, P.; Novotny, M. V.; Reilly, J. P. *J. Am. Soc. Mass Spectrom.* **2008**, *19* (7), 1027–1040.
[28] Polfer, N. C.; Valle, J. J.; Moore, D. T.; Oomens, J.; Eyler, J. R.; Bendiak, B. *Analytical Chemistry*. 2006, pp 670–679.
[29] Schindler, B.; Barnes, L.; Renois, G.; Gray, C.; Chambert, S.; Fort, S.; Flitsch, S.; Loison, C.; Allouche, A. R.; Compagnon, I. *Nat. Commun.* **2017**, *8* (1).
[30] Masellis, C.; Khanal, N.; Kamrath, M. Z.; Clemmer, D. E.; Rizzo, T. R. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (10), 2217–2222.
[31] Kamrath, M. Z.; Rizzo, T. R. *Acc. Chem. Res.* **2018**, *51*, 1487–1495.
[32] Gao, J.; Thomas, D. A.; Sohn, C. H.; Beauchamp, J. L. *J. Am. Chem. Soc.* **2013**, *135* (29), 10684–10692.
[33] Desai, N.; Thomas, D. A.; Lee, J.; Gao, J.; Beauchamp, J. L. *Chem. Sci.* **2016**, *7* (8), 5390–5397.
[34] Ruhaak, L. R., Xu, G., Li, Q., Goonatilleke, E., and Lebrilla, C. B. (2018) Mass Spectrometry Approaches to Glycomic and Glycoproteomic Analyses. *Chem. Rev.*
[35] Ly, T., Zhang, X., Sun, Q. Y., Moore, B., Tao, Y. Q., and Julian, R. R. (2011) Rapid, quantitative, and site specific synthesis of biomolecular radicals from a simple photocaged precursor. *Chem. Commun. 47*, 2835–2837.
[36] Tao, Y., Quebbemann, N. R., and Julian, R. R. (2012) Discriminating D -Amino Acid-Containing Peptide Epimers by Radical-Directed Dissociation Mass Spectrometry. *Anal. Chem. 84*, 6814–6820.
[37] Zhang, X., and Julian, R. R. (2014) Radical mediated dissection of oligosaccharides. *Int. J. Mass Spectrom. 372*, 22–28.

[38] Zhang, X., and Julian, R. R. (2013) Exploring radical migration pathways in peptides with positional isomers, deuterium labeling, and molecular dynamics simulations. *J. Am. Soc. Mass Spectrom. 24*, 524–533.
[39] Hahm, H. S., Schlegel, M. K., Hurevich, M., Eller, S., Schuhmacher, F., Hofmann, J., … Seeberger, P. H. (2017). Automated glycan assembly using the Glyconeer 2.1 synthesizer. Proceedings of the National Academy of Sciences of the United States of America, 114(17), E3385–E3389.
[40] Gaussian 09, Revision B.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, T. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2016.
[41] Hehre, W. J., Ditchfield, R., Radom, L., and Pople, J. A. (1970) Molecular Orbital Theory of the Electronic Structure of Organic Compounds. V. Molecular Theory of Bond Separation. *J. Am. Chem. Soc. 92*, 4796–4801.
[42] Blanksby, S. J., and Ellison, G. B. (2003) Bond dissociation energies of organic molecules. *Acc. Chem. Res. 36*, 255–263.
[43] Leymarie, N., and Zaia, J. (2012) Effective use of mass spectrometry for glycan and glycopeptide structural analysis. *Anal. Chem. 84*, 3040–3048.
[44] Tao, W. A., Zhang, D., Nikolaev, E. N., and 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.*
[45]Riggs, D. L., Gomez, S. V., and Julian, R. R. (2017) Sequence and Solution Effects on the Prevalence of D-Isomers Produced by Deamidation. *ACS Chem. Biol. 12*, 2875–2882.
[46] Wysocki, V. H., Tsaprailis, G., Smith, L. L., and Breci, L. A. (2000) Mobile and localized protons: A framework for understanding peptide dissociation. *J. Mass Spectrom. 35*, 1399–1406.
[47] Brodbelt, J. S. (2014) Photodissociation mass spectrometry: new tools for characterization of biological molecules. *Chem. Soc. Rev. 43*, 2757–83.
[48] Halim, M. A., Girod, M., MacAleese, L., Lemoine, J., Antoine, R., and Dugourd, P. (2016) 213 nm Ultraviolet Photodissociation on Peptide Anions: Radical-Directed Fragmentation Patterns. *J. Am. Soc. Mass Spectrom. 27*, 474–486.

[49] R. Julian, R. (2017). The Mechanism Behind Top-Down UVPD Experiments: Making Sense of Apparent Contradictions. Journal of the American Society for Mass Spectrometry, 28(9), 1823–1826.
[50] Lyon, Y. A., Sabbah, G. M., and Julian, R. R. (2017) Identification of Sequence Similarities among Isomerization Hotspots in Crystallin Proteins. *J. Proteome Res. 16*, 1797–1805.
[51] Lyon, Y. A., Sabbah, G. M., and Julian, R. R. (2018) Differences in α-Crystallin isomerization reveal the activity of protein isoaspartyl methyltransferase (PIMT) in the nucleus and cortex of human lenses. *Exp. Eye Res.*
[52] Moore, B. N., and Julian, R. R. (2012) Dissociation energies of X–H bonds in amino acids. *Phys. Chem. Chem. Phys. 14*, 3148.

CHAPTER 4

Exploration of Glutamine Deamidation: Products, Pathways, and Kinetics

Abstract

Spontaneous chemical modifications play an important role in human disease and molecular aging. Deamidation and isomerization are known to be among the most prevalent chemical modifications in long-lived human proteins and are implicated in a growing list of human pathologies, but the relatively minor chemical change associated with these processes has presented a longstanding analytical challenge. Although the adoption of high-resolution mass spectrometry has greatly aided the identification of deamidation sites in proteomic studies, isomerization (and the isomeric products of deamidation) remain exceptionally challenging to characterize. Herein, we present a liquid chromatography/mass spectrometry-based approach to rapidly characterize the isomeric products of glutamine deamidation using diagnostic fragments that are abundantly produced and capable of unambiguously identifying both Glu and isoGlu. Importantly, the informative fragment ions are produced through orthogonal fragmentation pathways, thereby enabling the simultaneous detection of both isomeric forms within a mixture and making this method amenable to shotgun proteomics. Furthermore, the diagnostic fragments associated with isoglutamic acid pinpoint the location of the modified residue. The utility of this technique is demonstrated by characterizing the isomeric products produced during *in vitro* aging of a series

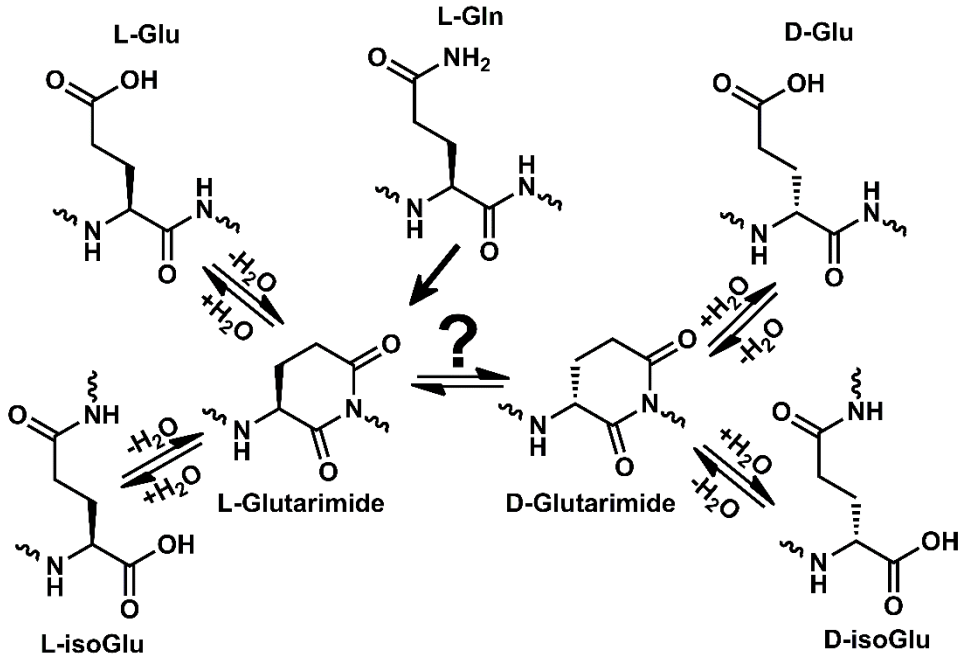
of glutamine-containing peptides. Sequence-dependent product profiles are obtained, and deamidation-linked racemization is confirmed by observation of D-isoGlutamic acid. Finally, comparisons are made between glutamine deamidation which is relatively poorly understood, and asparagine deamidation which has been more thoroughly studied.

Introduction

Asparagine and glutamine are uniquely unstable amino acids that spontaneously deamidate under physiological conditions.[1,2] Deamidation can modulate protein function,[3] and in some cases may be beneficial.[4] Nonetheless, it is more commonly perceived as an unwanted chemical modification that is disruptive to normal cellular function,[5] and is potentially a critical quality attribute in biological therapeutics.[6] Regardless of the origin, deamidation represents an important spontaneous degradation that can alter protein behavior and is associated with numerous age-related diseases,[7] including Alzheimer's,[8,9,10] multiple sclerosis,[11] prion-related encephalopathies,[12] and cataracts.[13,14]

The mechanism of asparagine deamidation has been thoroughly studied,[15] but glutamine deamidation is likely to follow an analogous chemical trajectory.[1,16] Both amino acids can deamidate through direct hydrolysis of the side chain, yielding the corresponding acidic amino acid (aspartic acid or glutamic acid). However, asparagine deamidation more commonly proceeds through a succinimide intermediate which forms as the backbone nitrogen of the C-terminal neighboring residue attacks the amide side chain. The resulting cyclic intermediate is susceptible to hydrolysis at two locations, yielding the isomeric products aspartic acid and isoaspartic acid (Asp and isoAsp). Additionally, the succinimide ring is racemization prone,[17] therefore four potential products are possible: L-Asp, L-isoAsp, D-Asp, and D-isoAsp. The corresponding mechanism for glutamine deamidation could yield a glutarimide intermediate and, in theory, could produce

the corresponding four isomers of glutamic acid (L-Glu, D-Glu, L-isoGlu, and D-isoGlu) as depicted in **Scheme 4.1**. However, experimental evidence for all four isomers is lacking, leaving open the question as to whether D-isomers form during glutamine deamidation.



Scheme 4.1. Deamidation of glutamine and isomerization of glutamic acid can both proceed through the glutarimide intermediate, potentially yielding four isomers of glutamic acid.

Deamidation causes a mass shift of 0.984 Dalton that can be distinguished from isotopic shifts with high-resolution mass spectrometry.[18] Additionally, mass spectrometry techniques employing electron-transfer dissociation and electron-

capture dissociation have been utilized to identify isoAsp[19] and isoGlu[20] residues by producing peptide fragments with characteristic mass-shifts. However, deamidation and isomerization are still considered to be among the most challenging chemical modifications to study.[21] In particular, the heterogeneous nature and chemical similarity of deamidation products has contributed to the difficulty in identifying the specific structural perturbations responsible for protein dysfunction. One of the most illustrative examples of the impact of deamidation arises in RNase U2, which has long been known to lose activity as it deamidates.[22] More recently, crystal structures of the native and deamidated enzymes have revealed that the native α-helix is partially unfolded when Asn32 deamidates to produce L-isoAsp.[23] An even more pronounced loss of function has been described in a therapeutic monoclonal antibody wherein a single isoAsp deactivates the entire antigen binding region.[24] Recently, *in vitro* aging has proven to be a valuable approach for predicting *in vivo* degradation.[25] For these reasons, deamidation and isomerization are important modifications in biological aging and merit consideration when designing biological therapeutics.[26,27]

Despite the growing interest in spontaneous degradation and protein aging, glutamine deamidation studies remain sparse. Although the first report of glutarimide formation during *in vitro* aging was reported nearly 3 decades ago,[16] many of the subsequent aging studies have focused on asparagine deamidation and/or aspartic acid isomerization. For instance, studies on asparagine have elucidated the role of pH, temperature, and secondary structure, while glutamine

remains relatively unexplored.[28,29] This is likely due to the timescale associated with glutamine deamidation[30] and the difficulty associated with detecting isomers via conventional proteomics. While the most rapid asparagine deamidation reactions have half-lives of ~1 day, the corresponding glutamine deamidation has an associated half-life of nearly 2 years.[31] Still, glutamine warrants further attention because it is known to be biologically significant, especially in extremely long-lived proteins such as the crystallins found in the eye lens,[32,33] and in brain proteins where Glu and isoGlu have recently been detected.[34] Interestingly, within the lens crystallin proteins, deamidation at specific glutamine sites has been reported to be more abundant than nearby asparagine deamidation sites,[35,36] although the overall extent of asparagine deamidation remains higher.[33] Unlike asparagine, some portion of the total deamidation observed at glutamine may be due to enzymatic activity, as tissue transglutaminase (TTG) has been associated with deamidation and crosslinking in ß-crystallins *in vitro*.[37] TTG likely plays a role in lens aging,[38] and is associated with cataractogenesis however a detailed understanding of this enzyme is still emerging. Nonetheless, glutamine deamidation in the lens increases with age, and it is associated with aggregation, loss of function, and insolubility.

We have previously utilized radical-directed dissociation (RDD) in conjunction with reverse-phase liquid chromatography to identify and characterize isomerization and deamidation.[39,40,41] In a recent study on cataractous human lens samples, we detected glutamine deamidation and were intrigued by the

observation of tandem MS peptide fragments arising from the Glu sidechain that appeared to diminish across a LC peak, hinting at the possibility of Glu/isoGlu coelution.[42] In this study, we develop a liquid-chromatography/mass spectrometry approach to rapidly identify and characterize the products of glutamine deamidation and glutamic acid isomerization. In particular, we improve on existing techniques by providing a means for positive identification of canonical Glu as well as isoGlu, while interfacing with a common reverse phase liquid chromatography setup. After validating the approach using synthetic glutamic acid isomers, we apply the technique to a series of model peptides subjected to *in vitro* aging. We discuss similarities to asparagine deamidation, and aspartic acid isomerization, and identify several important differences.

Experimental Section

Materials. Organic solvents and reagents were purchased from Fisher Scientific, Sigma-Aldrich, or Acros Organics and used without further purification. FMOC-protected amino acids and Wang resins were purchased from Anaspec, Inc or Chem-Impex International. Glutamic acid isomers were purchased from CarboSynth.

Peptide Synthesis. Peptides were synthesized manually following an accelerated FMOC-protected solid-phase peptide synthesis protocol.[43] *Para*-iodobenzoic acid (4IB) was attached to the N-terminus during synthesis. Following synthesis, peptides were purified using a Phenomenex Jupiter Proteo

C12 4 µm 90 Å 250 mm x 4.6 mm column. Purified peptides were stored frozen in 50/50 acetonitrile/water (v/v).

Aging Studies. Crystallin proteins were extracted from human lens tissue and modified with 4IB as previously described.[42] Briefly, human lens cortices and nuclei were separated using a trephine before homogenization. The water soluble and water insoluble proteins were subsequently fractionated by centrifugation, digested with trypsin, and analyzed by LCMS. Aliquots of each digest were modified with 4IB to facilitate RDD. Synthetic peptides were deamidated at 37 °C in microcentrifuge tubes submerged in a sand-bath style incubator. The initial deamidation studies were performed in 0.1M Tris buffer at pH 7.7. Accelerated studies were conducted in 0.2M Tris with the addition of 300mM NH_3 at pH 8.7. The isomerization study was carried out in 0.2M ammonium acetate at pH 5. Aliquots were collected at various time-points and stored frozen as lyophilized powder until analysis.

Analysis. Following aging studies, peptides were analyzed via LCMS. An Agilent 1100 binary pump was used with a Thermo BetaBasic-18 3 µm C18 150 mm x 2.1 mm column interfaced to a Thermo Fisher Scientific LTQ mass spectrometer with a standard electrospray ionization source. Samples were eluted using water with 0.1% formic acid as mobile phase A, and acetonitrile with 0.1% formic acid as mobile phase B. Synthetic standards were prepared as ~10 µM samples in 49.5/49.5/1 methanol/water/acetic acid (v/v) and infused into a modified LTQ linear ion trap using the standard electrospray ionization source.

The LTQ was modified with a quartz window to allow fourth harmonic (266 nm) laser pulses from a Nd:YAG laser to irradiate the trapped ion cloud which allows for photoinitiated RDD.[44] Following photodissociation, the radical peptide population was isolated and activated by CID to promote radical migration.

R_{isomer} scores. To distinguish isomers based on fragmentation patterns, we utilize an approach first developed by Tao *et al*.[45] Briefly, R_{isomer} scores were calculated using equation 1 where R_1 and R_2 refer to the ratios of a pair of fragment ions that change the most between two spectra. Identical fragmentation patterns would yield an R_{isomer} score of 1, indicating no discrimination, while the magnitude of the R_{isomer} score increases with larger differences between the spectra.

$$R_{isomer} = \frac{R_1}{R_2} \qquad (1)$$

We have previously determined that values >2.4 for RDD are statistically significant and are indicative of isomers that can be distinguished based on fragmentation patterns.[39]

Results and Discussion

We previously identified numerous sites of isomerization and deamidation within the crystallin proteins of the human lens.[40] Our isomer proteomics approach employs LC/MS^n and RDD to access unique and structurally sensitive fragmentation pathways which are capable of discerning peptide isomers. By evaluating the peptide fragmentation patterns, we are often able to identify isomers even when complete chromatographic separation is not possible. For instance,

during the analysis of a fraction from a 72-year-old water-insoluble cortex, the 4IB modified peptide ^{85}AVHLPSGGQYK^{95} from γS-crystallin was found to be heavily deamidated at Q93, but the deamidated peak is only partially resolved chromatographically (Figure 4.1a). Examination of the RDD data reveals that the fragmentation pattern at the beginning of the LC peak differs from the latter scans within the peak. In particular, the -59E and -72E glutamic acid side chain losses are notably absent in the leading shoulder of the peak but are observed in the trailing edge of the peak (Figure 4.1b). These side chain losses are commonly observed for glutamic acid but cannot be generated when the side chain is modified as in the case of isoglutamic acid (see supporting information Figure 4.S1 and reference [46] for additional mechanistic details.) We hypothesized that systematic differences in the radical fragmentation behavior for glutamic acid and isoglutamic acid could explain these observations.

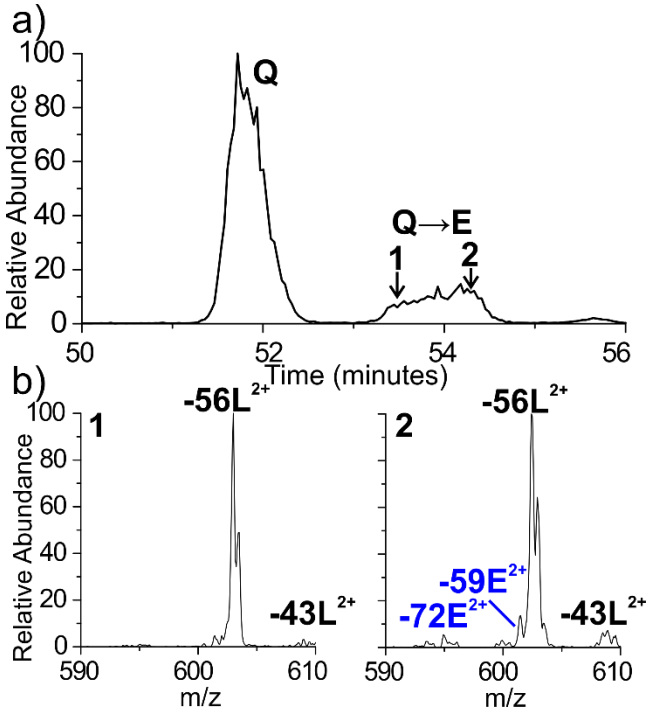


Figure 4.1. (a) The γS-crystallin AVHLPSGGQYK peptide is found to be deamidated in 72-year-old human lens tissue, yielding a broad peak at ~54 minutes. **(b)** The full MS are identical, however the RDD fragmentation patterns differ from the leading and trailing edge of the peak, indicating coeluting isomers. Notably, the -59E and -72E glutamic acid side chain losses are absent in the leading edge of the peak but appear in the trailing edge.

To evaluate the fragmentation patterns of glutamic acid isomers, we synthesized all four potential isomers of 4IB-VHLGGEGYK and performed RDD on the $[M+2H]^{2+}$ ions of each peptide. This synthetic peptide is derived from the γS-crystallin peptide sequence, with modifications to the sequence to facilitate *in vitro* aging (discussed below).[1] Figure 4.2 highlights the key differences and

similarities between the fragmentation patterns. Fragments common to all four isomers are labelled in black, while fragments that positively identify Glu are labelled in blue, and those that are unique to isoGlu are labelled in red. Relative intensities in all four spectra are normalized to the base peak, which is common to all four isomers.

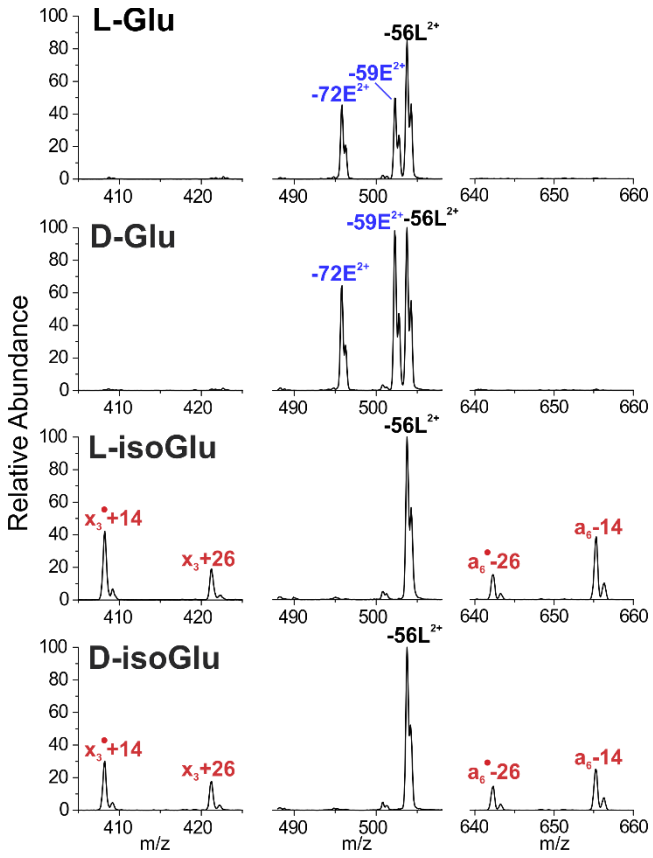
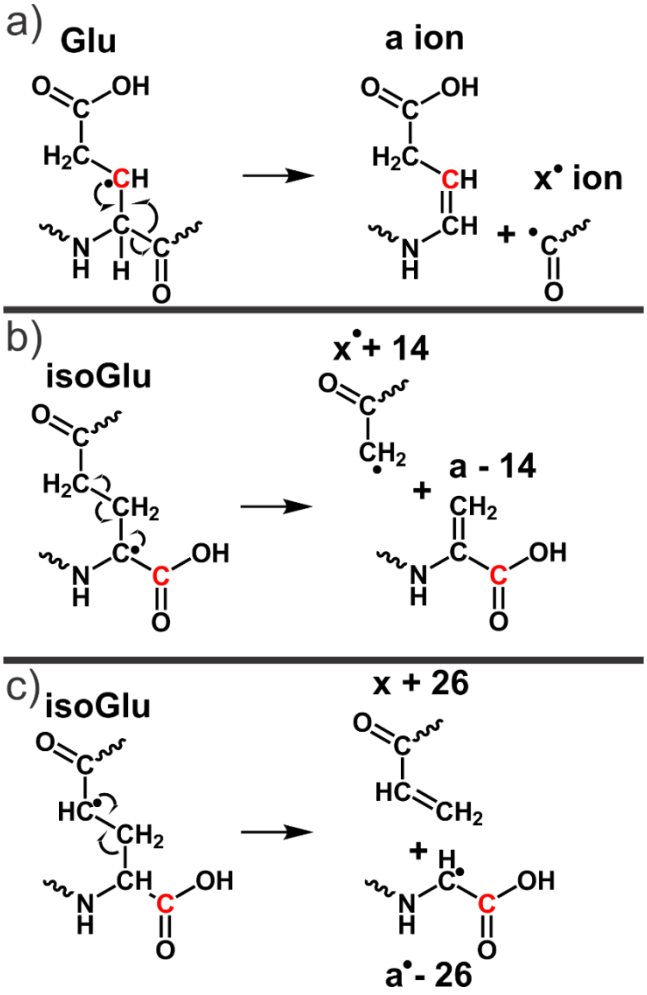


Figure 4.2. RDD fragmentation enables unambiguous identification of Glu and isoGlu containing peptides. Fragments common to all isomers are labelled in black. Fragments that positively identify Glu are labelled in blue, while fragments unique to isoGlu are labelled in red.

Several diagnostic fragments are apparent in Figure 4.2. For instance, L- and D-Glu can be identified by the abundant Glu side chain losses labelled -59E^{2+} and -72E^{2+} that are absent in the isoGlu-containing peptides. Conversely, both isoGlu peptides produce mass-shifted backbone fragments which are characteristic of the

unnatural ethylene group in the peptide backbone. Notably, these backbone mass-shifts are abundantly observed in peptides derived from both N-terminal ('$a_6^{\bullet}$ - 26'and 'a_6 - 14') as well as the C-terminal ('$x_3^{\bullet}$ + 14' and 'x_3 + 26') fragments. This observation is useful because fragment ions may be favored for one terminus in a typical LCMS run (i.e. C-terminal fragments are more commonly observed for tryptic peptides). Full spectra are provided in the supporting information.

The fragments observed in Fig. 2 arise from radical migration pathways localized at the Glu and isoGlu residues and are therefore expected to be largely sequence independent. RDD pathways that would yield the observed fragments are illustrated in Scheme 4.2. Migration of the initially formed radical to the β-carbon position of canonical Glu (depicted in red in Scheme 4.2a) leads to formation of normal a/x ions following β-scission. However, for isoGlu peptides, the β-carbon of the new side chain no longer contains an abstractable hydrogen. Instead, fragmentation at isoGlu is mediated by radical migration to the α-carbon and γ-carbon positions, which produces fragmentation within the unnatural backbone. Abstraction of the α-hydrogen leads to cleavage between the two ethylene carbons, yielding 'a - 14' and 'x$^{\bullet}$ + 14' backbone fragments (Scheme 4.2b). Alternatively, γ-hydrogen abstraction facilitates cleavage between the backbone ethylene and the α-carbon, producing 'x$^{\bullet}$ + 26' and 'a$^{\bullet}$ - 26' fragment ions as depicted in Scheme 4.2c. In both cases, the fragments are unique to isoGlu containing peptides and pinpoint the location of the unnatural backbone modification.



Scheme 4.2. **(a)** Radical fragmentation pathways depict formation of an a-type and x•-ion pair at the canonical Glu residue following β-hydrogen abstraction. **(b,c)** Lacking a β-hydrogen, isoGlu residues yield mass-shifted a- and x-type ions following α- or γ- hydrogen abstraction.

After establishing the mechanism for isomer discrimination, we aimed to test the utility of this approach by aging Gln containing peptides *in vitro*. 4IB-VHLGGQGYK (referred to as GQG) was aged in 100 mM Tris buffer at pH 7.7, 37 °C and monitored by LCMS. The chromatogram in Figure 4.3a depicts the 176-day (~6

month) aging time-point where the glutamine-containing precursor peak at ~44 minutes is followed by two deamidated product peaks. After 176 days, the peptide is 19.6% deamidated, corresponding to a half-life of 559 days, which is comparable to aging studies under similar conditions.[1] Of the deamidation products, the Glu peak represents 5.5% of the total peptide signal, while the isoGlu peak represents 14.1%, yielding a ratio of 2.6 which parallels previous observations for asparagine deamidation where ring cleavage favors the formation of the unnatural isomer (i.e. isoAsp or isoGlu).[15] The aging experiment was repeated after increasing the pH to 8.9 with ammonia, which is known to accelerate asparagine deamidation[47] without significantly altering the product profile.[41] The results are shown in Figure 4.3b, and the deamidation reaction is found to be greatly accelerated for glutamine as well. After just two weeks, the peptide is 9.4% deamidated, representing a 12-fold increase in rate relative to Tris alone. Interestingly, the isoGlu:Glu ratio is slightly elevated to 3.3 in the accelerated sample, suggesting that direct hydrolysis (which cannot form isoGlu) may contribute more to deamidation over the longer 176-day period. To confirm the deamidation product identities, RDD was employed continuously throughout the LCMS run. The characteristic fragmentation pattern of the first deamidated peak is shown in Figure 4.3c where the -59E and -72E fragments confirm the presence of the canonical Glu residue. Conversely, the second deamidated peak generated the diagnostic isoGlu mass-shifted backbone fragments, as shown in Figure 4.3d. Full spectra are available in the SI.

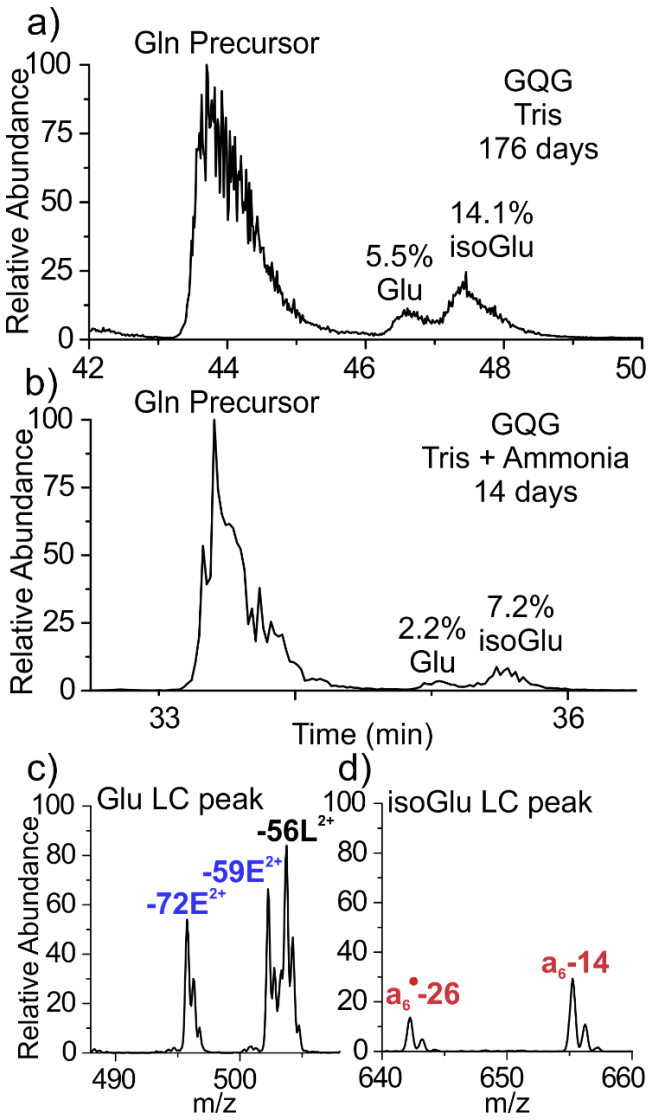


Figure 4.3. **(a)** Incubation of the 4IB-VHLGGQGYK peptide yields 19.6% deamidated products after 176 days. **(b)** Addition of ammonia greatly accelerates the deamidation process, yielding 9.4% deamidation within 14 days. **(c)** RDD confirms that the leading deamidation peak contains the canonical Glu residue, while the latter peak contains the isoGlu product.

The amino acid residue immediately C-terminal to glutamine or asparagine is known to strongly influence the ring formation rate.[1] To expose the impact of ring formation kinetics on the deamidation products, we incubated 4IB VHLGGQAYK and 4IB VHLGGQSYK (referred to as GQA and GQS respectively) in ammonia accelerated tris buffer. The deamidation results after 127 days are shown in Figure 4.4. To simplify comparisons, the accelerated 4IB VHLGGQGYK deamidation results are repeated in Figure 4.4. Interestingly, the product profile is inverted; the GQA peptide preferentially yields the canonical Glu product, while isoGlu is a minor product. A nearly identical product ratio is observed for the GQS peptide. In the case of the GQS peptide, only one product peak was observed *via* LCMS however the RDD spectra contained -59E and -72E side chain losses indicative of Glu in addition to the shifted a-type and x-type backbone fragments unique to isoGlu. The presence of both sets of fragments indicates that the deamidation products are coeluting, which is unsurprising given the close chemical similarity of the peptide isomers and further highlights the challenge in characterizing deamidation products *via* liquid chromatography. The product ratios were deconvoluted based on calibration curves for the diagnostic isoGlu fragments. Additional details and data are available in the SI.

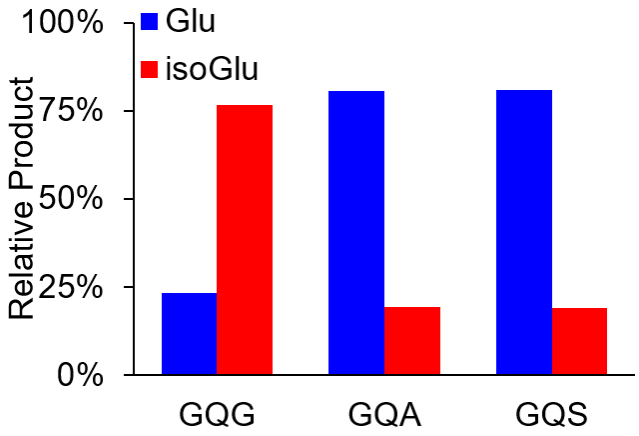


Figure 4.4. Relative deamidation product ratios following ammonia accelerated deamidation for 4IB-VHLGGQGYK (t=14d), 4IB-VHLGGQAYK (t=127d), 4IB-VHLGGQSYK (t=127d).

Interestingly, the GQG peptide which parallels asparagine deamidation results by preferentially generating isoGlu seems to represent an exception among this set of glutamine containing peptides. Both the GQA and GQS peptides, which are considerably slower to deamidate, predominately produce the canonical Glu residue. We attribute this difference primarily to the local flexibility afforded by the GQG peptide which is flanked by glycine on either side, and may therefore facilitate ring formation. While asparagine deamidation is known to proceed primarily through the succinimide intermediate, direct hydrolysis has previously been found to be a competing pathway in glutamine deamidation, especially as the bulkiness of the neighboring amino acids increased.[1] Because alanine and serine are among the least bulky side-chain neighbors, we did not expect to observe significant levels of direct hydrolysis in the GQA and GQS peptides, however the isoGlu products

are clearly disfavored. This reveals that even in the presence of ammonia which greatly accelerates ring formation, direct hydrolysis remains a competitive pathway for glutamine deamidation in all but the most flexible peptides. This finding may help explain why isoGlu has remained so difficult to identify in biological samples, despite the comparatively common observation of glutamine deamidation.

As a final comparison between Glu and Asn deamidation, we were interested in exploring the potential of D- isomer production during glutarimide formation. The succinimide ring formed during asparagine deamidation is known to be racemization prone,[17] yielding D-isomers with ~10% relative abundance independent of peptide sequence and deamidation rate.[41] Clarke and Houk previously rationalized the formation of D- isomers by comparing the proton affinity of the succinimide α-carbon using *ab initio* quantum mechanics calculations (RHF/6-31+G*) and density-functional theory (B3LYP/6-31+G*)[17] where they found the succinimide to be 18 kcal/mol more acidic than its amide counterpart, corresponding to a pKa shift of 10.8 pH units. A comparable effect is expected for the glutarimide ring which is chemically similar. To evaluate this possibility, L-isoGlu and D-isoGlu synthetic standard were used to construct a calibration curve. Plotting the fragmentation pattern observed for the 4IB-VHLGGEGYK isoGlu peak after 37 days under accelerated conditions corresponds with 86% L-isoGlu (see Figure 4.S5 in the SI). The differences in fragmentation between the L-isoGlu and D-isoGlu are subtle, this coupled with the low overall signal observed for the deamidated products makes quantitating the chiral isomers exceptionally

challenging. As such, there is a non-negligible amount of error associated with the measurement. Nonetheless, the fragmentation data suggest that the level of racemization occurring during glutamine deamidation is comparable to previous observations with asparagine deamidation.[41]

Although we only detected racemization for the isoGlu products of the GQG peptide, we cannot rule out the possibility of racemization occurring within the other sequences. The isoGlu products of the GQG sequence are likely to produce the largest quantity of D- isomer products and are therefore easiest to detect. This is based on the mechanistic reasoning that all of the isoGlu products are derived from the ring (where racemization is most likely to take place), while only some of the Glu product arises from the ring (with the remainder arising from direct hydrolysis). Furthermore, the GQG sequence produced the most isoGlu among the sequences we examined; the greater overall signal is beneficial for detecting the small fraction of D-isomers present.

Given the similarities to the succinimide pathway, which can be accessed from both Asn and Asp, we considered the potential of glutamic acid isomerization *via* the glutarimide intermediate. Drawing from the known pH dependence of asparagine deamidation (which is favored under neutral to basic conditions)[47] and aspartic acid isomerization (which is favored under acidic conditions),[48] we incubated 4IB-VHLGGEGYK in ammonium acetate, at pH 5. The aging data following 2 months of incubation (61 days) is shown in Figure 4.5. Chromatographically, only one peak is observed (Figure 4.5a), suggesting that no

isoGlu has formed within this timeframe. During the majority of the peak, only one species is present in the full mass spectrum at 595.2 m/z, matching the doubly protonated L-Glu precursor (Figure 4.5b left panel), however a minor second peak appears at 586.2 m/z (-18Da) toward the trailing end of the peak (Figure 4.5b right panel), consistent with the observation of the glutarimide ring. To verify the identity of the -18 Da peak, we performed RDD (Figure 4.5c). The -18 Da loss is found to be localized on peptide fragments which contain the Glu residue, as expected following formation of a glutarimide ring. Notably, this reaction was found to proceed ~37x slower than the corresponding deamidation reaction at pH 7.7. This confirms that Glu can form the glutarimide ring in an analogous manner to Gln, and establishes another parallel between the two aging processes. It is expected that same combination of isomers would be produced following the re-opening of this glutarimide ring, which is identical to that produced during deamidation.

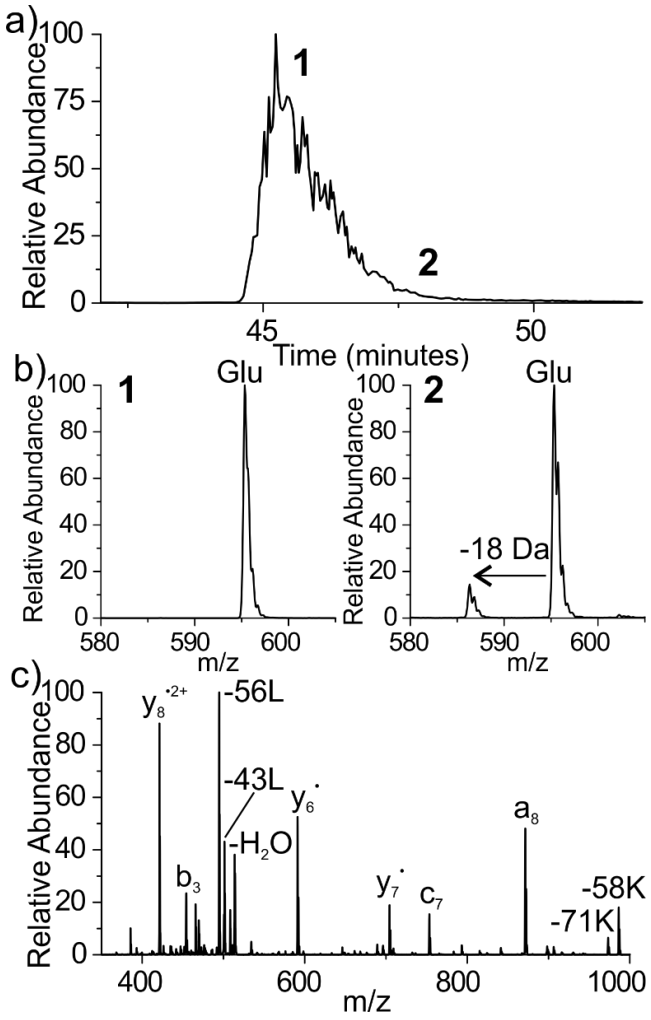


Figure 4.5. Aging results for 4IB-VHLGGEGYK peptide under acidic conditions. **(a)** Only one peak is observed by LCMS, suggesting that no isoGlu has formed within 61 days. **(b)** Trace amounts of a coeluting species at -18 Dalton, consistent with the glutarimide, is observed during the trailing edge of the main peak. **(c)** RDD fragmentation confirms the presence of the glutarimide intermediate.

Conclusion

Asparagine and glutamine deamidation share many similarities and have been grouped together as nearly equivalent aging processes, but our *in vitro* aging

studies have revealed important differences. As previously reported with asparagine, ammonia is effective at accelerating glutamine deamidation. Yet, glutamine peptides are found to be more variable in terms of the product profiles which seem to be more strongly influenced by primary sequence (and likely flexibility). Even under accelerating conditions, direct hydrolysis was found to be a competitive process and may account for a larger percentage of the overall deamidation at glutamine. Nonetheless, it is clear that isoGlu merits consideration for certain sequences (i.e. QG motifs) or otherwise flexible regions in extremely long-lived proteins. Ultimately, RDD was found to excel at discriminating Glu from isoGlu, with positive identifications enabled by the -59E and -72E fragments which are unique to the canonical Glu side chain, and mass shifted backbone fragment pairs at '$x^{\bullet}$ + 14' and 'x + 26', and '$a_6^{\bullet}$ - 26' and 'a - 14' which can be used to pinpoint the location of an isoGlu residue. Additionally, we have demonstrated that the level of isoGlu in an aged sample can be quantitatively assessed *post hoc* by use of a calibration curve. While the combination of ammonia acceleration with RDD-based isomer characterization is poised to enable more comprehensive *in vitro* aging studies, the technique may prove most useful in the analysis of complex aged human tissue such as the eye lens and brain where long-lived proteins are abundant and often highly deamidated. RDD benefits greatly by enabling positive identification of both Glu and isoGlu through orthogonal fragmentation channels, and in a way that is directly applicable to shotgun proteomics.

Acknowledgements

The authors are grateful for funding from the NIH (NIGMS grant R01GM107099).

Supporting Information

Radical Fragmentation Pathways:

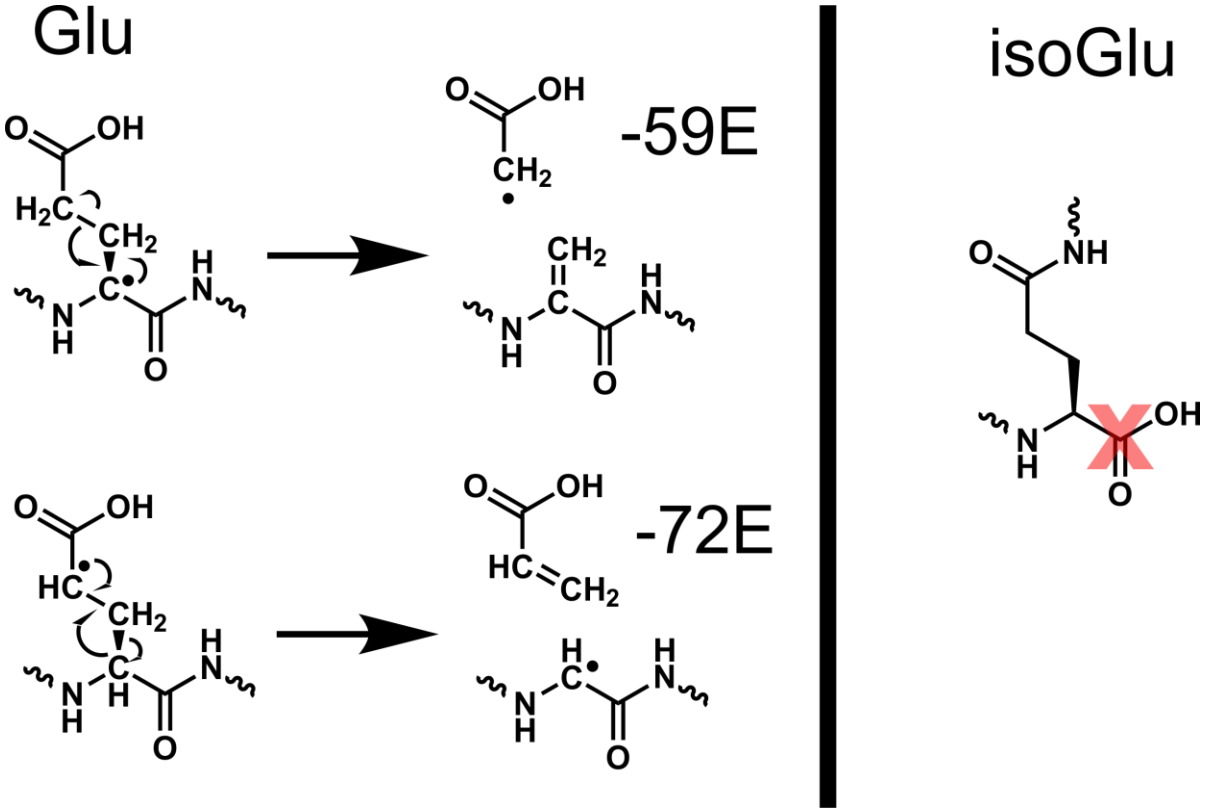


Figure 4.S1. The -59E and -72E side chain losses associated with glutamic acid in radical directed dissociation arise from α-hydrogen and γ-hydrogen abstraction. These side chain losses are not possible in isoglutamic acid.

Full RDD spectra:

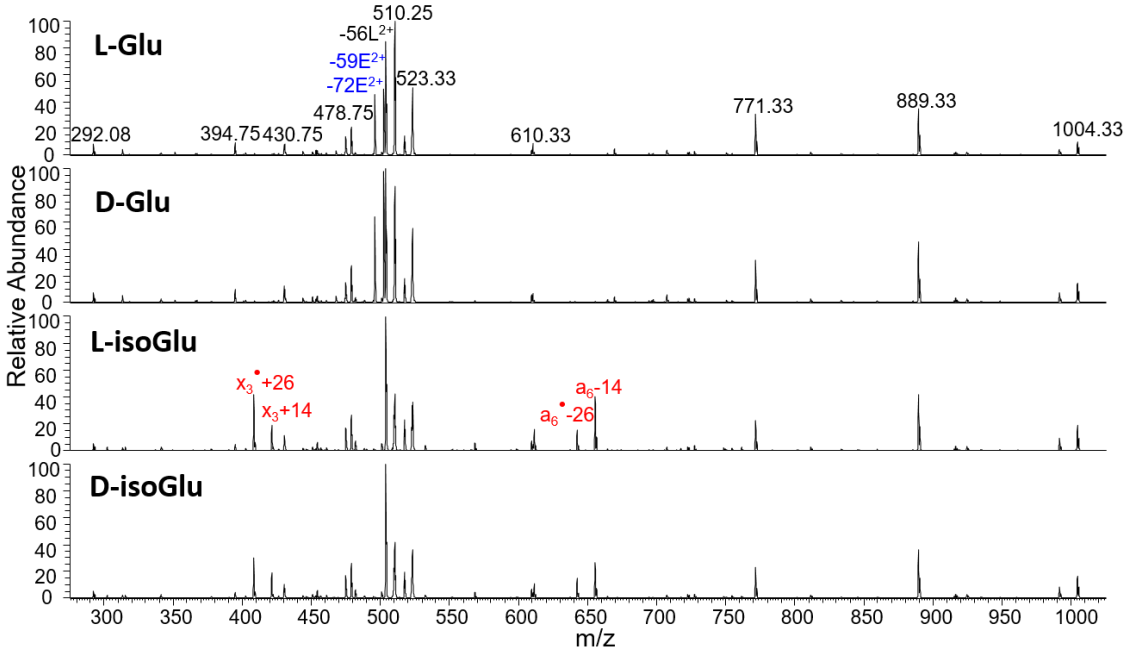


Figure 4.S2. Full RDD spectra for the 4IB-VHLGGEGYK isomer standards. Label colors match Figure 4.2 in the main text.

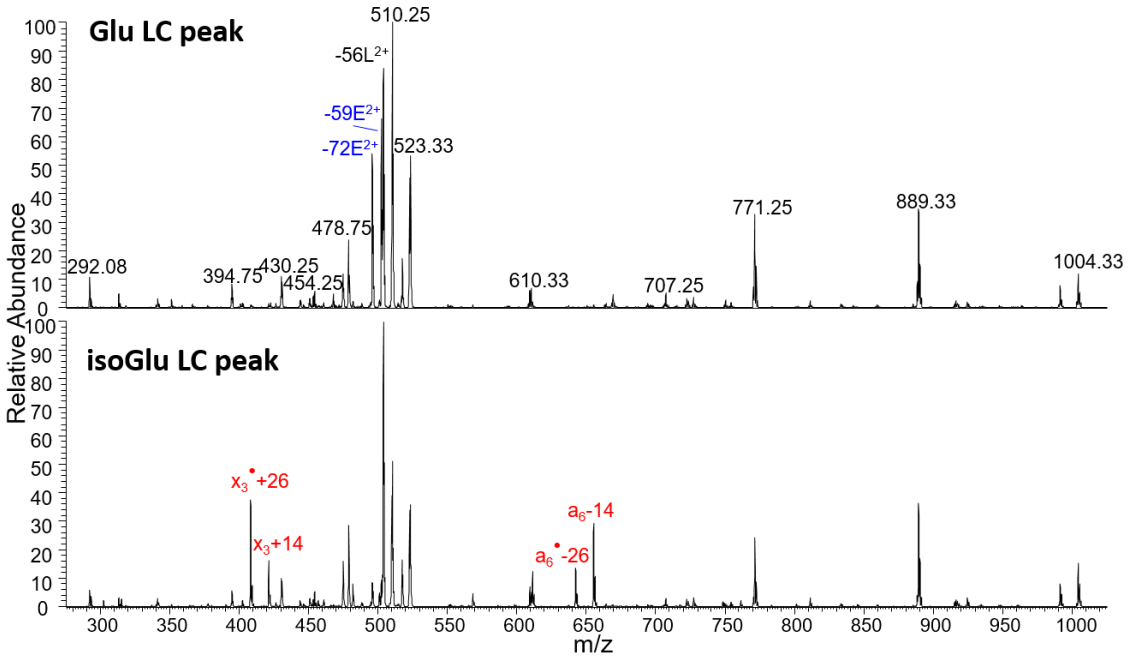


Figure 4.S3. RDD spectra for the leading and trailing edge of the deamidated product peak generated upon incubation of 4IB-VHLGGQGYK for 127 days. The diagnostic fragments confirm the presence of both Glu and isoGlu.

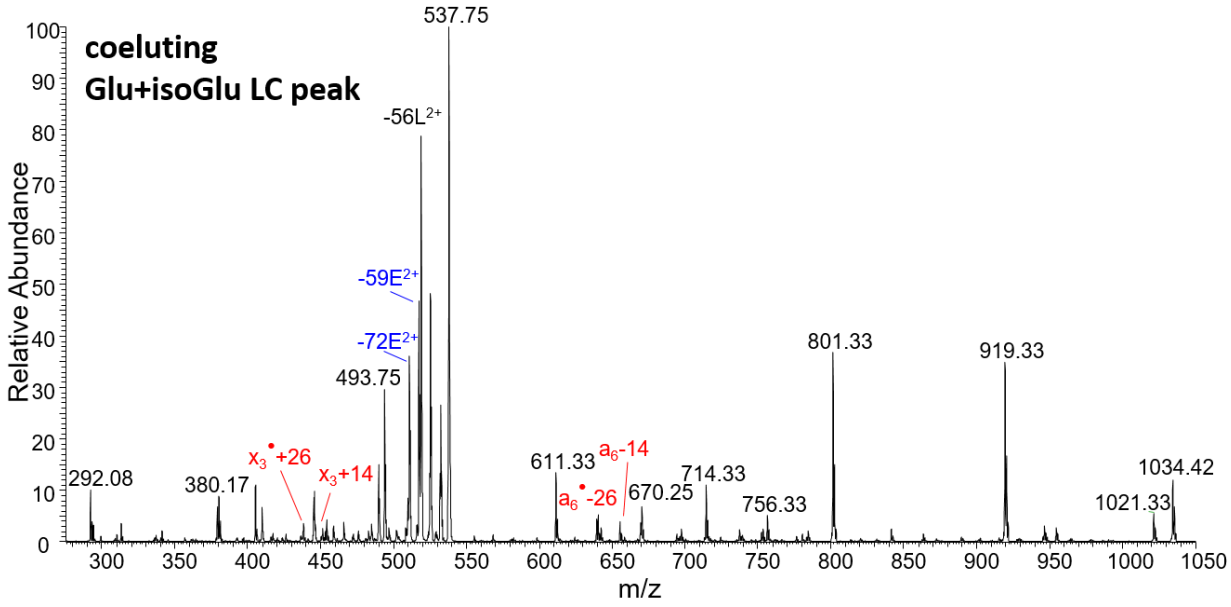


Figure 4.S4. The RDD spectra averaged across the coeluting deamidation product peak for the 4IB-VHLGGQSYK peptide contains both sets of diagnostic fragments, revealing the presence of both Glu and isoGlu.

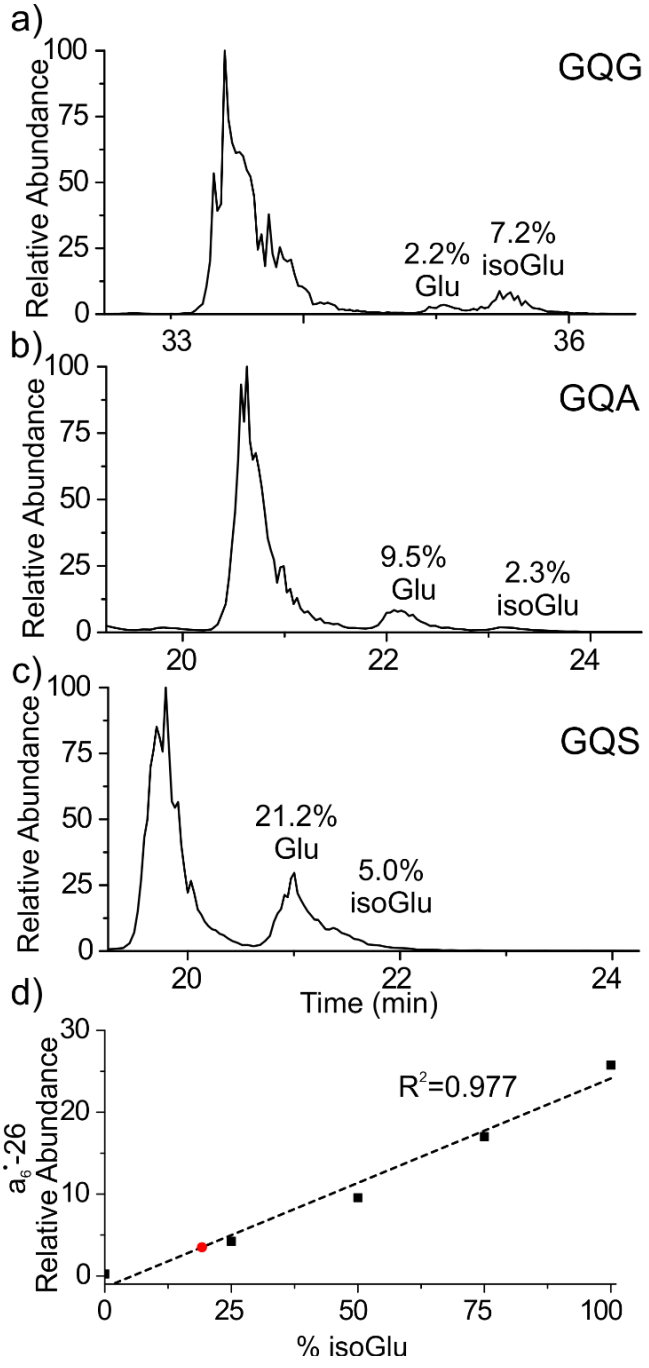
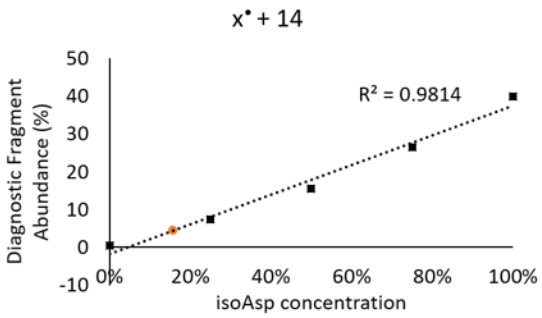
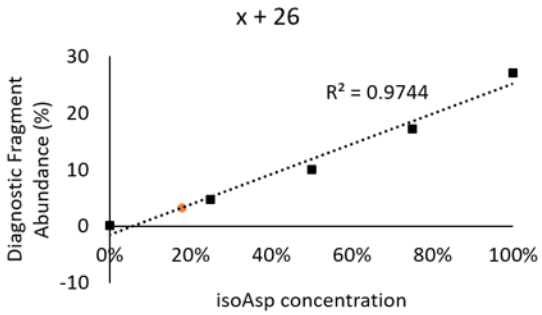
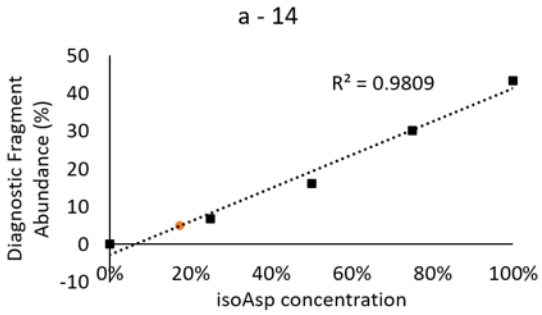

a)
GQG
Relative Abundance
100
75
50
25
0
2.2%
Glu
7.2%
isoGlu
33
36
b)
GQA
Relative Abundance
9.5%
Glu
2.3%
isoGlu
20
22
24
c)
GQS
Relative Abundance
21.2%
Glu
5.0%
isoGlu
20
22
24
Time (min)
d)
$a_6^{\bullet}$-26
Relative Abundance
30
20
10
0
R^2=0.977
0
25
50
75
100
% isoGlu

Figure 4.S5. Ammonia accelerated deamidation results for **(a)** 4IB-VHLGGQGYK (t=14d), **(b)** 4IB-VHLGGQAYK (t=127d), **(c)** 4IB-VHLGGQSYK (t=127d). **(d)** A calibration curve based on RDD fragmentation can be used to characterize the coeluting Glu and isoGlu isomers produced by the GQS peptide (red circle).


x• + 14
Diagnostic Fragment Abundance (%)
R² = 0.9814
50
40
30
20
10
0
-10
0%
20%
40%
60%
80%
100%
isoAsp concentration


x + 26
Diagnostic Fragment Abundance (%)
R² = 0.9744
30
20
10
0
-10
0%
20%
40%
60%
80%
100%
isoAsp concentration


a - 14
Diagnostic Fragment Abundance (%)
R² = 0.9809
50
40
30
20
10
0
-10
0%
20%
40%
60%
80%
100%
isoAsp concentration

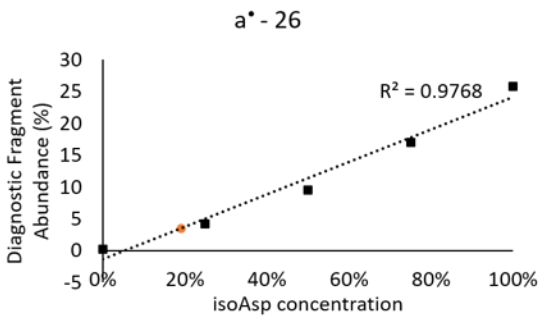

a• - 26
Diagnostic Fragment Abundance (%)
R² = 0.9768
30
25
20
15
10
5
0
-5
0%
20%
40%
60%
80%
100%
isoAsp concentration

Figure 4.S6. Calibration curves for each diagnostic fragment produce consistent results when evaluating the amount of isoAsp in the GQS deamidated product peak. Avg isoAsp = 17.48%, stdev = 1.45%.

To assess the product ratios in the GQS peptide, we generated calibration curves which plot the abundance of each isoGlu diagnostic fragment ion as a function of isoGlu concentration. The curve depicting the '$a_6^{\bullet}$ - 26' fragment is shown in Figure 4.S6. The deamidated product peak produces the diagnostic '$a_6^{\bullet}$ - 26' fragment ion with 3.5% relative abundance, indicating that the peak contains 19.2% isoGlu. The average value obtained from the four diagnostic fragment curves is 17.48% with a standard deviation of 1.45%. Accounting for the relative peak area, this corresponds to a total product profile of 21.2% Glu and 5.0% isoGlu in the 127-day aged GQS sample and an isoGlu:Glu ratio of 0.23, similar to the 0.24 ratio observed for the GQA peptide.

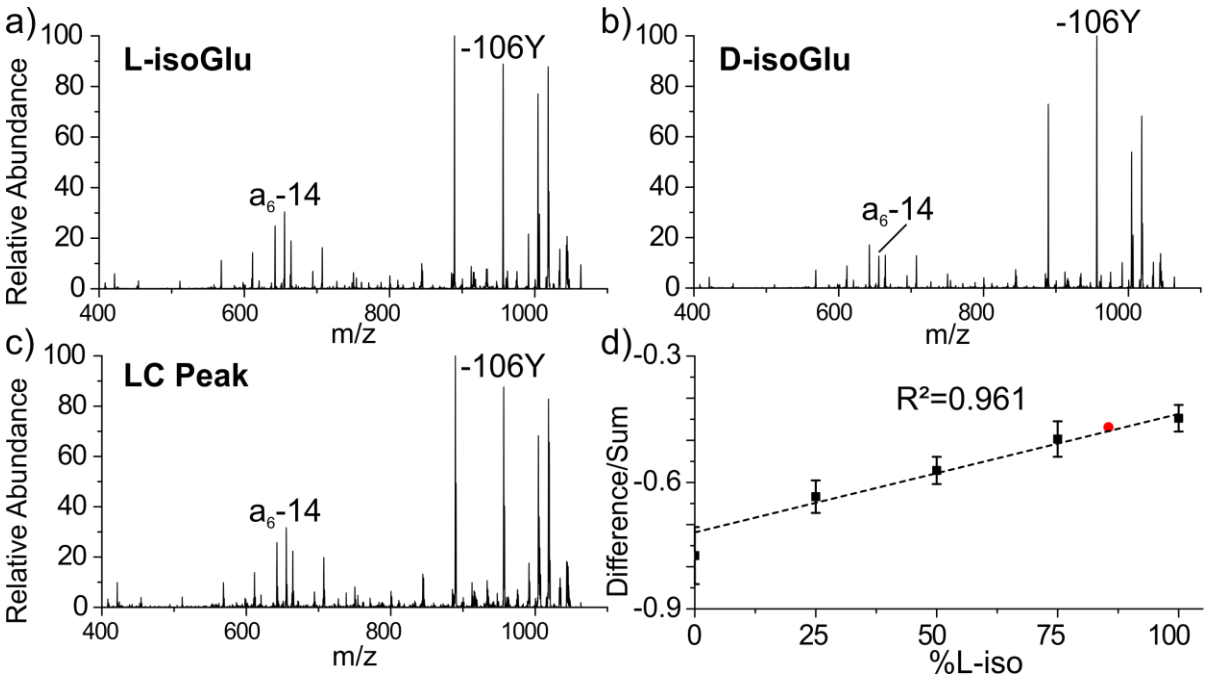


Figure 4.S.5. RDD fragmentation patterns for synthetic forms of 4IB-VHLGGEGYK in the L-isoGlu form **(a)** and D-isoGlu form **(b)** with significant differences in the relative abundances of the a_6-14 and -106Y fragments. The RDD fragmentation for the LC peak corresponding to the isoGlu products following deamidation of the 4IB-VHLGGQGYK peptide **(c)**. A calibration curve quantitates the level of L-isoGlu within the LC peak by plotting the difference/sum of the relative abundance for the two diagnostic peaks **(d)**.

The fragmentation patterns of L-isoGlu synthetic standard were compared with the D-isoGlu standard (Figures S.5a,b) for 4IB-VHLGGEGYK. For this peptide, the a_6-14 and -106Y (tyrosine side chain loss) fragment ion abundances differ at a subtle, but statistically significant level. Comparing the ratio of these two peaks

between the L-iso and D-iso standards yields an R_{isomer} score of 3.0 which is above our threshold of 2.4 to differentiate two isomers. Importantly, the fragmentation pattern obtained for the deamidated isoGlu product LC peak is not an exact match to either pure standard, consistent with isomer coelution (Figure 4.S.5c). To quantify the amount of L-isoGlu and D-isoGlu, we generated a calibration curve by plotting the difference divided by the sum of the two fragment ion abundances that yield the R_{isomer} score (Figure 4.S.5d). The fragmentation pattern reveals that the LC peak contains 86% L-isoGlu, and 14% D-isoGlu which is comparable to the levels of racemization detected in previous studies on asparagine deamidation.

References

[1] Robinson, N. E.; Robinson, Z. W.; Robinson, B. R.; Robinson, A. L.; Robinson, J. A.; Robinson, M. L.; Robinson, A. B. *J. Pept. Res.* **2004**, *63* (5), 426–436.
[2] Robinson, N. E.; Robinson, A. B. *Proc. Natl. Acad. Sci.* **2001**, *98* (3), 944–949.
[3] Curnis, F.; Longhi, R.; Crippa, L.; Cattaneo, A.; Dondossola, E.; Bachi, A.; Corti, A. *J. Biol. Chem.* **2006**, *281* (47), 36466–36476.
[4] Reissner, K. J.; Aswad, D. W. *Cell. Mol. Life Sci.* **2003**, *60* (7), 1281–1295.
[5] Clarke, S. *Ageing Res. Rev.* **2003**, *2* (3), 263–285.
[6] Yan, Q., Huang, M., Lewis, M. J., and Hu, P. (2018) Structure Based Prediction of Asparagine Deamidation Propensity in Monoclonal Antibodies. *MAbs 10*, 901–912.
[7] Lindner, H.; Helliger, W. *Exp. Gerontol.* **2001**, *36* (9), 1551–1563.
[8] Lam, Y. P. Y.; Wootton, C. A.; Hands-Portman, I.; Wei, J.; Chiu, C. K. C.; Romero-Canelon, I.; Lermyte, F.; Barrow, M. P.; O'Connor, P. B. *Chem. Commun.* **2018**, *54* (98), 13853–13856.
[9] Nguyen, P. T.; Zottig, X.; Sebastiao, M.; Bourgault, S. *Biochemistry* **2017**, *56* (29), 3808–3817.
[10] Roher, A. E.; Lowenson, J. D.; Clarke, S.; Wolcow, C.; Wang, R.; Cotter, R. J.; Reardon, I. M.; Zurcher-Neely, H. A.; Heinrikson, R. L.; Ball, M. J.; Greenberg, B. D. *J. Biol. Chem.* **1993**, *268* (5), 3072–3083.
[11] Friedrich, M. G.; Hancock, S. E.; Raftery, M. J.; Truscott, R. J. W. *Acta Neuropathol. Commun.* **2016**, *4* (1), 83.
[12] Sandmeier, E.; Hunziker, P.; Kunz, B.; Sack, R.; Christen, P. *Biochem. Biophys. Res. Commun.* **1999**, *261* (3), 578–583.
[13] Truscott, R. J. W.; Schey, K. L.; Friedrich, M. G. *Trends Biochem. Sci.* **2016**, *41* (8), 654–664.
[14] Takata, T.; Oxford, J. T.; Demeler, B.; Lampi, K. J. **2008**, pp 1565–1575.
[15] Geiger, T.; Clarke, S. *J. Biol. Chem.* **1987**, *262* (2), 785–794.
[16] Capasso, S.; Mazzarella, L.; Sica, F.; Zagari, A. *J. Chem. Soc., Chem. Commun.* **1991**, No. 23, 1667–1668.
[17] Radkiewicz, J. L.; Zipse, H.; Clarke, S.; Houk, K. N. *Journal of the American Chemical Society*. **1996**, 9148–9155.
[18] Yang, H.; Fung, E. Y. M.; Zubarev, A. R.; Zubarev, R. A. *J. Proteome Res.* **2009**, *8* (10), 4615–4621.
[19] Cournoyer, J. J.; Lin, C.; O'Connor, P. B. *Anal. Chem.* **2006**, *78* (4), 1264–1271.
[20] Li, X.; Lin, C.; O'connor, P. B. *Anal. Chem.* **2010**, *82* (9), 3606–3615.
[21] Lampi, K. J.; Wilmarth, P. A.; Murray, M. R.; David, L. L. *Prog. Biophys. Mol. Biol.* **2014**, *115* (1), 21–31.
[22] Uchida, T.; Shibata, Y. *J. Biochem.* **1981**, *90* (2), 463–471.
[23] Noguchi, S. *Biopolymers* **2010**, *93* (11), 1003–1010.
[24] Rehder, D. S.; Chelius, D.; Mcauley, A.; Dillon, T. M.; Xiao, G.; Crouse-zeineddini, J.; Vardanyan, L.; Perico, N.; Mukku, V.; Brems, D. N.; Matsumura, M.; Bondarenko, P. V. *Biochemistry* **2008**, 2518–2530.
[25] Tran, J. C.; Tran, D.; Hilderbrand, A.; Andersen, N.; Huang, T.; Reif, K.; Hotzel, I.; Stefanich, E. G.; Liu, Y.; Wang, J. *Anal. Chem.* **2016**, *88* (23), 11521–11526.

[26] Wakankar, A. A.; Borchardt, R. T. *J. Pharm. Sci.* **2006**, *95* (11), 2321–2336.
[27] Xu, A.; Kim, H. S.; Estee, S.; Viajar, S.; Galush, W. J.; Gill, A.; Hötzel, I.; Lazar, G. A.; McDonald, P.; Andersen, N.; Spiess, C. *Mol. Pharm.* **2018**, *15* (10), 4529–4537.
[28] Patel, K.; Borchardt, R. T. *Pharmaceutical Research: An Official Journal of the American Association of Pharmaceutical Scientists*. 1990, pp 787–793.
[29] Yüksel, K. Ü.; Gracy, R. W. *Arch. Biochem. Biophys.* **1986**, *248* (2), 452–459.
[30] Li, X.; Lin, C.; O'connor, P. B. *Anal. Chem.* **2010**, *82* (9), 3606–3615.
[31] Robinson, N. E.; Robinson, Z. W.; Robinson, B. R.; Robinson, A. L.; Robinson, J. A.; Robinson, M. L.; Robinson, A. B. *J. Pept. Res.* **2004**, *63* (5), 426–436.
[32] Hanson, S. R.; Smith, D. L.; Smith, J. B. *Exp. Eye Res.* **1998**, *67* (3), 301–312.
[33] Hains, P. G.; Truscott, R. J. W. *Investig. Ophthalmol. Vis. Sci.* **2010**, *51* (6), 3107–3114.
[34] Serra, A.; Gallart-Palau, X.; Wei, J.; Sze, S. K. *Anal. Chem.* **2016**, *88* (21), 10573–10582.
[35] Hains, P. G.; Truscott, R. J. W. *J. Proteome Res.* **2007**, *6* (10), 3935–3943.
[36] Asomugha, C. O.; Gupta, R.; Srivastava, O. P. *Mol.Vis.* **2010**, *16* (1090–0535 (Electronic)), 476–494.
[37] Boros, S.; Wilmarth, P. A.; Kamps, B.; de Jong, W. W.; Bloemendal, H.; Lampi, K.; Boelens, W. C. *Experimental Eye Research*. **2008**, pp 383–393.
[38] Lorand, L.; Hsu, L. K.; Siefring, G. E.; Rafferty, N. S. *Proc. Natl. Acad. Sci.* **2006**, *78* (3), 1356–1360.
[39] Tao, Y.; Julian, R. R. *Anal. Chem.* **2014**, *86* (19), 9733–9741.
[40] Lyon, Y. A.; Sabbah, G. M.; Julian, R. R. *J. Proteome Res.* **2017**, *16* (4), 1797–1805.
[41] Riggs, D. L.; Gomez, S. V.; Julian, R. R. *ACS Chem. Biol.* **2017**, *12* (11), 2875–2882.
[42] Lyon, Y. A.; Sabbah, G. M.; Julian, R. R. *Experimental Eye Research*. **2018**.
[43] 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.
[44] Ly, T., Julian, R. R. (2008) Residue-Specific Radical-Directed Dissociation of Whole Proteins in the Gas Phase, *J. Am. Chem. Soc. 130*, 351–358.
[45] Tao, W. A., Zhang, D., Nikolaev, E. N., and 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.*
[46] Sun, Q.; Nelson, H.; Ly, T.; Stoltz, B. M.; Julian, R. R. *J. Proteome Res.* **2009**, *8* (2), 958–966.
[47] Tyler-Cross, R.; Schirch, V. *J. Biol. Chem.* **1991**, *266* (33), 22549–22556.
[48] Aswad, D. W.; Paranandi, M. V.; Schurter, B. T. *J. Pharm. Biomed. Anal.* **2000**, *21* (6), 1129–1136.

CHAPTER 5

Concluding Remarks

Aging is an inevitable phenomenon which has daunted mankind since the dawn of history and has inspired stories of a mystical fountain of youth which date back thousands of years. Despite this longstanding interest, the molecular underpinnings of aging remain enigmatic. Many of the recognized aging processes are elusive and analytically challenging to assess. As the instrumental tools for exploring such pathways continue to advance, the development of appropriate methodologies promises to be a fruitful endeavor that will offer rewarding insight into natural biological systems while also providing a framework to guide biomolecular engineering.

Briefly, we have capitalized on the structural sensitivity associated with radical migration through the development and utilization of radical directed dissociation, which takes places following photodissociation inside an ion trap mass spectrometer. This capability, when coupled with liquid chromatography, has allowed us to fully characterize the isomeric products of deamidation and isomerization. In doing so, we established baseline product profiles associated with succinimide ring cleavage. This information allowed us to identify common laboratory practices and conditions that promote racemization, provided mechanistic insight into the deamidation process, and further allowed us to validate approaches toward accelerating in vitro aging studies as outlined in chapter two.

We followed up our initial studies by exploring the even more enigmatic process of glutamine deamidation. We outline the mechanistic underpinnings that lead to diagnostic fragmentation channels, allowing for facile discrimination of Glu and isoGlu residues. Furthermore, we outline several similarities between asparagine and glutamine deamidation, however we also identified several key differences.

Finally, the scope of our radical based technique was expanded in chapter three to evaluate a comprehensive family of isomeric carbohydrates. This work establishes a framework for glycan identification using radical chemistry. Importantly, we demonstrate that a commercially available mass spectrometer is capable of benefitting from radical chemistry and outline a path toward bridging the closely related but often isolated fields of proteomics and glycomics.

Ultimately, we have employed mass spectrometry as an exceptionally informative stand-alone approach, but also benefitted greatly by pairing it with orthogonal and complementary techniques. Our recent aging studies have opened exciting new avenues for isomer proteomics that are poised to continue to make meaningful contributions to our rapidly expanding appreciation of molecular and human aging.