Title

An aging tug of war: The accumulation of the spontaneous L-isoaspartate modification in proteins vs. canonical and novel maintenance pathways

Author

Warmack, Rebeccah

Publication Date

2019

Peer reviewed|Thesis/dissertation

UNIVERISTY OF CALIFORNIA

Los Angeles

An aging tug of war: The accumulation of the spontaneous L-isoaspartate modification in proteins *vs.* canonical and novel maintenance pathways

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Biochemistry and Molecular Biology

by

Rebeccah Warmack

2019

ABSTRACT OF THE DISSERTATION

An aging tug of war: The accumulation of the spontaneous L-isoaspartate modification in proteins *vs.* canonical and novel maintenance pathways

by

Rebeccah Warmack

Doctor of Philosophy in Biochemistry and Molecular Biology

University of California, Los Angeles, 2019

Professor Steven G. Clarke, Chair

Alterations in protein stability and propensity for aggregation are hallmarks of many prevalent and costly age-related diseases, and are actively investigated in cases like Alzheimer's and cataracts. However, as expensive clinical trials of drugs based on the native targets fail, we must reconsider what the toxic species are within these protein aggregation diseases. One hypothesis focuses on the alteration of aggregative properties by spontaneous protein modifications, which may generate more toxic polymorphs. One such age-related post-translational modification, the isomerization of aspartate residues (L-isoaspartate, L-isoAsp), is a pervasive modification where preliminary evidence for such aggregation enhancement has been obtained. The primary goal of this thesis has been to characterize how this form of protein damage is linked to aggregation of proteins associated with specific aging diseases, as well as how this modification is maintained in the functioning cell. In the context of the water-soluble extracts of aging mammalian lens, this work has shown that L-isoAsp accumulates to high levels, primarily within aggregated low molecular weight species. *In vitro*

experiments based on our localization of L-isoAsp in the lens revealed that introduction of the L-isoAsp modification in a water-soluble derived peptide was accompanied by decreased aggregative properties, while the opposite result was observed for a water-insoluble derived peptide. In further experiments with segments of the Alzheimer's disease-associated amyloid-β (Aβ) peptide it was demonstrated that the rate of Aβ fibril formation is greatly accelerated with an isoaspartyl residue in agreement with earlier literature. Structural studies revealed the L-isoAsp residue facilitates the formation of a novel protofilament interface, which may represent a distinct polymorph of the full-length Aβ peptide. Here we also begin the characterization of a novel L-isoAsp maintenance pathway distinct from the canonical L-isoAsp repair enzyme – the protein-L-isoaspartyl (D-aspartyl) *O*-methyltransferase or protein carboxyl methyltransferase (PCMT1) - through the protein carboxyl domain-containing proteins (PCMTDs). We have shown that these proteins assemble with cullin-RING ligase proteins and lower detectable levels of L-isoAsp *in vitro*. From this work, it is clear that L-isoAsp residues have significant consequences in protein aggregation diseases, and are attenuated in cells through multiple pathways.

The dissertation of Rebeccah Warmack is approved.

David William Walker

Jorge Torres

Steven Gerard Clarke, Committee Chair

University of California, Los Angeles

2019

DEDICATION

I would like to dedicate this work to my wonderful family. To my parents who have always encouraged and supported me, and to my siblings who led the way and inspired me to always reach higher.

TABLE OF CONTENTS

Abstract of Dissertation……..ii

Committee Page……iv

Dedication……v

Table of Contents……...vi

List of Figures……viii

List of Tables……...xi

Commonly Used Abbreviations……...…xii

Acknowledgements……..xiii

Vita…….....xv

Publications……...…xvi

Chapter 1: Introduction & plan of dissertation……...……1

References…….…9

Chapter 2: Racemized and Isomerized Proteins in Aging Rat Teeth and Eye Lens……13

References……22

Chapter 3: The L-isoaspartate modification within protein fragments in the aging lens can promote protein aggregation……..…24

References……...39

Chapter 4: Structure of amyloid-β (20-34) with Alzheimer's-associated isomerization at Asp23 reveals a novel protofibril interface……77

References……98

Chapter 5: The PCMTDs are a novel L-isoAsp maintenance pathway and interact with cullin-RING ligase proteins……..…124

References…….……140

Chapter 6: Future research and concluding remarks…….……154

References…….……159

Appendix Chapter 1: A glutamate/aspartate switch controls product specificity in a protein arginine methyltransferase ……...…160

References…….167

Appendix Chapter 2: Protein Arginine Methyltransferase Product Specificity Is Mediated by Distinct Active-site Architectures……...…168

References…….178

LIST OF FIGURES

Figure 1.1 Pathway of L-isoasapartate formation……..…….……..2

Figure 2.1 Determination of the D-aspartate/L-aspartate ratio in rat tooth molar hydrolysate…18

Figure 2.2 D-aspartate/L-aspartate ratios in rat molars do not significantly correlate with rat age or molar position…….....18

Figure 2.3 L-isoaspartate accumulates in the rat eye lens after 2 months, but plateaus after 18 months of age…….....19

Figure 2.4 The data from Figure 3 plotted as a function of the age of the individual rats…….....20

Figure 2.5 Age dependence of the specific activity of endogenous protein L-isoaspartate methyltransferase in rat eye lens……..……....21

Figure 3.1 Pathway for L-isoaspartate (L-isoAsp) formation and repair……..…48

Figure 3.2 The L-isoAsp modification accumulates to high levels within the aged human lens polypeptides while endogenous methyltransferase repair activity does not significantly decrease with age……..50

Figure 3.3 L-isoAsp damage accumulates with age throughout a range of polypeptide sites in water-soluble extracts of human lenses, primarily within LMW species…….….52

Figure 3.4 Size exclusion chromatography of WS extracts demonstrates age-dependent increases in HMW protein species and in L-isoAsp damage in crystallins, particularly within the HMW fractions……...54

Figure 3.5 Size exclusion chromatography of WS extracts demonstrates age-dependent increases in HMW protein species and in L-isoAsp damage in crystallins, particularly within the HMW fractions …….…....55

Figure 3.6 Identification of crystallin fragments from the gel filtration WS-HMW fraction reveals the presence of peptides from a variety of crystallins and other lens proteins with no strong trends in residue specificity or localization of cleavage site.……..…...56

Figure 3.7 The isomerized forms of WS-HMW derived synthetic peptides $\alpha A^{52\text{-}65,\ isoAsp58}$ and $\alpha A^{89\text{-}98,\ isoAsp91}$ do not inhibit αB crystallin chaperone activity, but modestly increase protein precipitation *in vitro*.……..…..58

Figure 3.8 The native form of WS-HMW derived synthetic peptide $\alpha A^{52\text{-}65}$ shows enhanced self-aggregation and protein precipitation compared to the isomerized form.……..…..60

Figure 3.9 Isomerization of the anti-chaperone peptide $\alpha A^{66\text{-}80}$ increases its ability to form amyloid and enhances its anti-chaperone activity in vitro.…….……..…....62

Figure S3.1 Exogenous PCMT1 labeling of L-isoAsp is not affected by urea..…….69

Figure S3.2 Exogenous PCMT1 cannot methylate aggregated $\alpha A^{66\text{-}80,\ isoAsp76}$ peptide.…..……69

Figure S3.3 Analysis of L-isoAsp damage in the native size-exclusion fractions of human nuclear lens extracts by SDS-PAGE reveals the HMW gel filtration fraction contains significant amounts of LMW L-isoaspartyl-containing species.…….……...……..71

Figure S3.4 The $\alpha A^{66\text{-}80,\ isoAsp76}$ peptide aggregates purified αB crystallin more than the native $\alpha A^{66\text{-}80}$ peptide.…….……...…….72

Figure 4.1 L-isoAsp in $A\beta^{20\text{-}34}$ accelerates fiber formation and can seed native segment ……..……...…….107

Figure 4.2 Modified fibers have increased resistance to SDS disaggregation.……..…...109

Figure 4.3 $A\beta^{20\text{-}34,\ isoAsp23}$ structure contains an altered protofilament interface……...110

Figure 4.4 A putative model of the heritable Iowa mutation in interface B.…….....…...112

Figure 4.5 Model of $A\beta^{20\text{-}34,\ isoAsp23}$ as the core of an Aβ 1-42 modified polymorph...…..114

Figure 4.6 $A\beta^{20\text{-}34,\ isoAsp23}$ assumes a similar fold to full-length native Aβ structures.…...115

Figure 4.7 $A\beta^{24\text{-}34}$ peptide structure shares similarities with the $A\beta^{20\text{-}34}$ structures……...116

Figure S4.1 PCMT1 is unable to fully methylate aggregated $A\beta^{20\text{-}34,\ isoAsp23}$……....119

Figure S4.2 Seeded $A\beta^{20\text{-}34}$ aggregates display identical fiber diffraction patterns, and the full-length aggregation matches that of the segment ……120

Figure S4.3 $A\beta^{20\text{-}34}$ structures reveal two interfaces and a polar zipper ……...…....121

Figure S4.4 $A\beta^{20\text{-}34,\ isoAsp23}$ matches β-helical structures and other amyloid fibril…….....122

Figure S4.5 Total atom RMSD relationships between Aβ structures.……...123

Figure 5.1 Sequence alignment of PCMT1 and PCMTD proteins.……...145

Figure 5.2 PCMTD1 associates with Elongins B and C, which aid in binding of the AdoMet cofactor.……..…...146

Figure 5.3 The PCMTD1-Elongins do not display methyltransferase activity.……..…147

Figure 5.4 The PCMTD1-Elongins associate with Cul5 and Rbx2 *in vitro*.……....148

Figure 5.5 PCMTD1-Elongins coupled with cullin-RING ligase components lower detectable levels of L-isoaspartate *in vitro*.……149

Figure 5.6 *In vitro* PCMTD1-Elongins can lower PCMT1-detectable L-isoaspartate without active ubiquitination.……...…151

Figure 5.7 *In vitro* ubiquitination assay with the carbonic anhydrase substrate reveals PCMTD1-dependent ubiquitination……153

LIST OF TABLES

Table 1.1 Structures containing L-isoaspartate residues……..5

Table 3.1 Parent proteins of lens fragments ……...…46

Table 3.2 Top five peptides identified from each of 10 abundant lens proteins ……47

Table S3.1 αA crystallin peptides ……63

Table S3.2 αB crystallin peptides ……65

Table S3.3 βA3 crystallin peptides ……...……66

Table S3.4 βA4 crystallin peptides ……...……67

Table S3.5 βB1 crystallin peptides ……...……68

Table S3.6 βB2 crystallin peptides ……...……69

Table S3.7 Filensin peptides ……70

Table S3.8 γS crystallin peptides ……71

Table S3.9 Phakinin peptides ……72

Table 4.1 Data collection and refinement statistics ……...……..106

COMMONLY USED ABBREVIATIONS IN THIS DISSERTATION

L-isoAsp = L-isoaspartate

(isoD) = L-isoaspartate within the sequence

AdoMet = *S*-adenosyl-L-methionine

[^{3}H] AdoMet = *S*-adenosyl-L-[-methyl-^{3}H]methionine

AdoHcy = *S*-adenosyl-L-homocysteine

ATP = adenosine triphosphate

PTM = Post-translational modification

PCMT1 = Protein carboxyl methyltransferase

PCMTD = Protein carboxyl methyltransferase domain-containing protein

Aβ = amyloid-β peptide

Cul5 = Cullin 5

ACKNOWLEDGEMENTS

I would like to wholeheartedly thank my graduate mentor Dr. Steven Clarke for his generosity in allowing me to pursue a diversity of projects. This freedom was crucial for my growth as a scientist. I am also forever indebted to Dr. Jonathan Lowenson for his boundless knowledge and limitless patience for all questions great and small. I would like to thank past and present members of the Clarke laboratory for their constructive feedback on the projects presented in this dissertation. I owe gratitude to my committee for guiding this dissertation, and I would also like to extend a special thanks to the many incredible collaborators I had the opportunity to work with, who I will mention in association with the projects below.

This work was supported in part by the Dissertation Year Fellowship, particularly Chapters 3 and 4, as well as the Cellular and Molecular Biology Pre-Doctoral Training Grant, the Edwin W. Pauley Fellowship, and the Audree Fowler Fellowship in Protein Science.

Chapter 2 of this dissertation is a reprint of the publication “Racemized and Isomerized Proteins in Aging Rat Teeth and Eye Lens” in *Rejuvenation Research* (Volume 19, pages 309-317). I would like to acknowledge the fellow authors Eduardo Mansilla, Rodolfo Goya, and Steven Clarke.

Chapter 3 is a manuscript in preparation titled “The L-isoaspartate modification within protein fragments in the aging lens can promote protein aggregation.” I would like to thank my collaborators and co-authors Harrison Shawa, Kate Liu, Katia Lopez, Joseph A. Loo, Joseph Horwitz, and Steven G. Clarke.

Chapter 4 is a manuscript in preparation titled “Structure of amyloid-β (20-34) with Alzheimer’s-associated isomerization at Asp23 reveals a novel protofibril interface.” I would like to thank my collaborators and co-authors David R. Boyer, Chih-Te Zee, Logan S. Richards, Michael R. Sawaya, Duilio Cascio, Tamir Gonen, David S. Eisenberg, and Steven G. Clarke.

Chapter 5 is an ongoing project titled “The PCMTDs are a novel L-isoAsp maintenance pathway and interact with cullin-RING ligase proteins.” This project has involved a number of

collaborators, I would like to especially thank students Esther Peluso and Dylan Valencia for their work on this project.

Appendix chapters 1 and 2 are reprints of articles titled “A glutamate/aspartate switch controls product specificity in a protein arginine methyltransferase” in *Proc. Natl. Acad. Sci USA* (volume113, pages 2068-2073), and “Protein Arginine Methyltransferase Product Specificity Is Mediated by Distinct Active-site Architectures,” in *J. Biol. Chem* (volume 291, pages18299-18308), respectively. These papers were collaborations with Kanishk Jain, Andrea Hadjikyriacou, You Feng, Erik Debler, Peter Stavropoulos, Günter Blobel, and Steven Clarke.

VITA

2009 – 2013	Bachelor of Science, UC Davis, High Honors Major: Biotechnology with Microbiology/Fermentation emphasis Minor: Technology Management
2013 – 2014	Teaching Assistant, Department of Chemistry and Biochemistry University of California, Los Angeles, Los Angeles, CA
2015	Candidate in Philosophy in Biochemistry and Molecular Biology University of California, Los Angeles, Los Angeles, CA
2014 – 2017	Cellular and Molecular Biology Pre-Doctoral Training Grant, Trainee University of California, Los Angeles, Los Angeles, CA
2017 – 2018	Edwin W. Pauley Fellowship UCLA College of Letters and Sciences, Division of Physical Sciences
2018	John M. Jordan Memorial Award University of California, Los Angeles, Los Angeles, CA
2018 – 2019	Dissertation Year Fellowship University of California, Los Angeles, Los Angeles, CA
2018 – 2019	Audree Fowler Fellowship in Protein Science University of California, Los Angeles, Los Angeles, CA

PUBLICATIONS

Warmack RA, Boyer DR, Zee C-T, Richards LS, Sawaya MR, Cascio D, Gonen T, Eisenberg DS, Clarke SG. Structure of amyloid-β (20-34) with Alzheimer's-associated isomerization at Asp23 reveals a novel protofibril interface. (2019) ***Manuscript submitted for publication.***

Warmack RA, Shawa H, Liu K, Lopez K, Loo JA, Horwitz J, Clarke SG. The L-isoaspartate modification within protein fragments in the aging lens can promote protein aggregation. (2019) ***Manuscript submitted for publication.***

Jain K, **Warmack RA**, Debler EW, Hadjikyriacou A, Stavropoulos P, Clarke SG. Protein arginine methyltransferase product specificity is mediated by distinct active-site architectures. J Biol Chem. 2016 Aug 26. 291(8):18299-308.

Warmack RA, Mansilla E, Goya RG, Clarke SG. Racemized and isomerized proteins in aging rat teeth and eye lens. Rejuvenation Res. 2016 Feb 26. 19(4):309-17.

Debler EW, Jain K, **Warmack RA**, Feng Y, Clarke SG, Blobel G, Stavropoulos P. A glutamate /aspartate switch controls product specificity in a protein arginine methyltransferase. Proc Natl Acad Sci USA. 2016 Feb 23; 113(8):2068-73.

Chapter 1

Introduction & Plan of the Dissertation

Welcome to the Isoaspartyl-ome

Introduction

A majority of post-translational modifications (PTMs) are covalent, enzyme-mediated additions to the polypeptide chain that can alter protein functionality, affect protein-protein interactions, and clear the cell of proteins that are past their prime (1). These include ubiquitination, acetylation, phosphorylation, methylation, and glycosylation among others. Unfortunately, due to the reactive nature of certain amino acid residues and modifying groups, there are a number of PTMs, which occur spontaneously that can interrupt protein function or cause aggregation (2). Tyrosine nitration and protein glycation are examples of the non-enzymatic addition of a modifying moiety (a nitro group, sugar group, respectively (3, 4)). While other spontaneous PTMs do not involve the addition of an external entity, such as pyroglutamylation, racemization, and isomerization. Pyroglutamylation occurs exclusively on N-terminal glutamine/glutamate residues, while racemization and isomerization affect primarily serine, aspartate, and asparagine residues at any site within a polypeptide chain (5, 6). These spontaneous modifications accumulate over time and can strongly contribute to loss of protein function and aggregation.

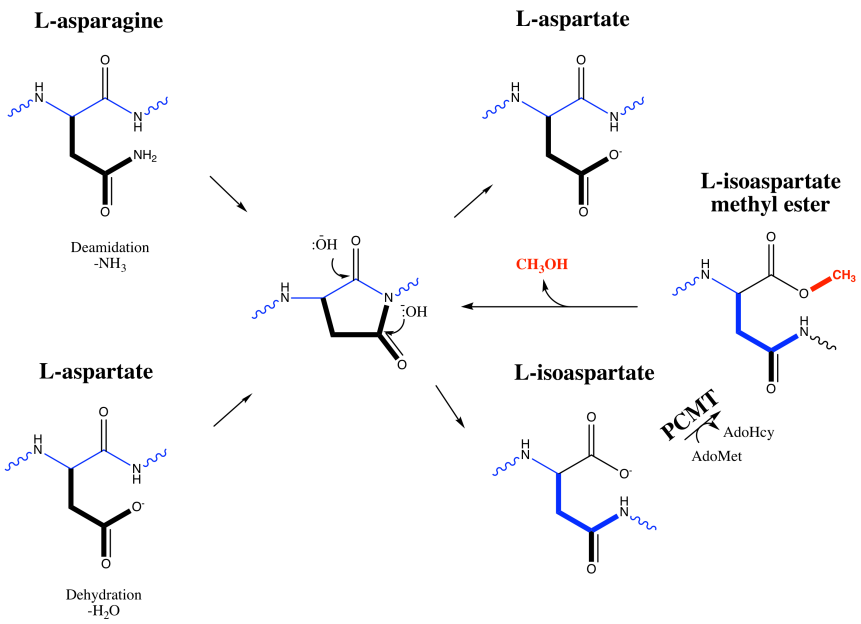


Figure 1: Pathway of L-isoaspartate formation.

The deamidation and isomerization of asparagine and aspartate residues, respectively, both primarily proceed through a five-membered succinimide ring intermediate (**Fig. 1**, adapted from ref. 7). This ring can be hydrolyzed at either carbonyl, but due to the asymmetry of the

succinimide, the isomerized aspartate (L-isoaspartate; L-isoAsp) form is usually 60-80% of the product compared to the normal L-aspartate (L-asp) product (8). The L-succinimide ring can also racemize to yield a D-succinimide intermediate, which can in turn be hydrolyzed to form D-aspartate (D-asp) and D-isoaspartate (D-isoAsp). The isomerized residues reroute the main chain peptide backbone through the side chain methylene, which elongates the backbone with a flexible carbon and shortens the negative side chain by one carbon-carbon bond. The L-isoAsp products can be recognized and repaired by the protein-L-isoaspartyl (D-aspartyl) *O*-methyltransferase (PCMT1). This enzyme specifically recognizes L-isoAsp with high affinity and D-asp residues with 700 – 1000 fold lower affinity (9), and methylates the side chain, creating a labile methyl ester. This methyl ester is quickly hydrolyzed, catalyzing the reformation of the succinimide ring and subsequently the normal L-asp.

The significance of the L-isoaspartate damage and PCMT1 repair pathway is evidenced not only by growth retardation, but also by the average life span of 42 days within *Pcmt1* knockout mice (10). The cause of death at two months is grand mal seizures. Within tissues of equivalent wild-type mice, levels of L-isoAsp are low, reaching maximal levels of 250 pmol L-isoAsp per mg soluble brain extract, 100 pmol L-isoAsp per mg soluble heart extract, and 200 pmol L-isoAsp per soluble testis extract. While the short lifespan of the *Pcmt-1* knockout mice precludes seeing a full plateau in the accumulation of L-isoAsp, in each of the aforementioned tissues the levels are significantly higher, with over 2000 pmol L-isoAsp observed per mg soluble brain extract, 350 pmol L-isoAsp per mg soluble heart extract, and 450 pmol L-isoAsp per soluble testis extract (11). The exact relationship between the knockdown of *Pcmt1* and the onset of the grand mal seizures remains unknown, but the brains within Pcmt1 knockout mice are observed to be ~20-30% larger than wild-type mice brains (12). Intriguingly, treatment with the phosphoinositide-3-kinase (PI3K) inhibitor wortmannin decreases this enlargement of the brain to within 6% the size of the normal brain, and nearly doubles the lifespan of the knockout mice (13). These results suggest that either the enzyme PCMT1 itself or the resultant

accumulation of L-isoAsp modifications plays an important role in brain growth and development, and perhaps specifically through the insulin-signaling pathway.

The importance of the PCMT1/L-isoAsp system is emphasized by the values listed in the above paragraph, which show that the Pcmt1 knockout brain tissue has over 4-fold higher levels of L-isoAsp than heart and testis tissues (11). This was seen by Vigneswara et al., in experiments separating PCMT1-radiolabeled L-isoAsp containing proteins by SDS-PAGE and visualized on film (14). In these experiments, PCMT1 substrates within Pcmt1 knockout brain, heart, spleen, liver, kidney, thymus, and skeletal muscle tissues were resolved by 1D SDS-PAGE and again demonstrated that the number of L-isoAsp containing PCMT1 substrates are many fold higher within brain tissue than the other tissues analyzed. These researchers went on to identify 9 major substrates from the Pcmt1 knockout brain ranging from a-synuclein to the microtubule-associate protein-2. Other studies further identified 19 L-isoAsp containing PCMT1 substrates, which also had a considerable range of cellular functions (15). This wide range of substrates is perhaps not surprising in light of earlier *in vitro* studies which highlighted the high affinity the PCMT1 enzyme has for a large number of L-isoAsp containing peptides with varying sequences (11). The only sequence-based deviations that appear to affect the PCMT1 methylation are prolyl and negatively charged residues on the carboxyl side of the L-isoAsp residue.

This diversity of L-isoAsp-containing PCMT1 substrates begs that question – how does the L-isoAsp modification affect structure and function in these different proteins? *In vitro* aging studies of proteins reveal that the rate of L-isoAsp formation varies greatly between proteins, within a four week time span it can range from no formation to 1.5 mol L-isoAsp per mol protein (16). In many of the cases where L-isoAsp accumulates to high levels, the result is loss of function, as is the case for calmodulin (17), the HPr phosphocarrier protein (18), and creatine kinase B (19) whose enzyme activities can be partially rescued by repair with the PCMT1 enzyme. While spontaneous alteration of the native form of the protein backbone is likely to

disrupt structure and function, it has been suggested in the case of integrin ligands to cause "gain-of-function" (20). The extracellular matrix protein fibronectin contains Asn-Gly-Arg (NGR) motifs that deamidate to form isoAsp-Gly-Arg (isoDGR), which can mimic the preferred integrin-binding motif Arg-Gly-Asp (RGD) (21). Thus, it is clear that the effects of the L-isoAsp modification on function can vary between proteins.

While it is difficult to characterize the biochemical effects of L-isoAsp on protein function on a case-by-case basis, it is a monumental achievement to characterize the modification structurally. Proteins rarely form the modification at a stoichiometric ratio and creating synthetic constructs of L-isoAsp within large protein sequences is not facile, which makes structural studies that require homogeneous samples very challenging. However, many researchers have persevered and to date there are 41 structures containing L-isoAsp reported in the protein data bank (**Table 1**). A majority of these 41 structures (28 out of 41) are actually an enzyme involved in bacterial cell wall synthesis, murA. In each of these studies, the L-isoAsp was found to be homogeneous throughout the population in a hinge-region of the protein, and it is assumed in this case that it is not an aging incident, but

Table 1: Structures containing L-isoaspartate residues.

PDB ID	Structure Title
1AT6	Hen egg white lysozyme with an isoaspartate residue
1C9P	Complex of bdellastasin with porcine trypsin
1DLG	Crystal structure of the C115S enterobacter cloacae murA in the un-liganded state
1DY5	Deamidated derivative of bovine pancreatic ribonuclease
1EJC	Crystal structure of unliganded mura (type2)
1EJD	Crystal structure of unliganded mura (type1)
1EYN	Structure of mura liganded with the extrinsic fluorescence probe ANS
1JG3	Crystal Structure of L-isoaspartyl (D-aspartyl) O-methyltransferase with adenosine & VYP(ISP)HA substrate
1LSQ	Ribonuclease A with Asn67 replaced by a beta-aspartyl residue
1Q3G	MurA (Asp305Ala) liganded with tetrahedral reaction intermediate
1RTU	USTILAGO SPHAEROGENA RIBONUCLEASE U2
1RYW	C115S MurA liganded with reaction products
1YBG	MurA inhibited by a derivative of 5-sulfonoxy-anthranilic acid
2FI4	Crystal structure of a BPTI variant (Cys14->Ser) in complex with trypsin
2FI5	Crystal structure of a BPTI variant (Cys38->Ser) in complex with trypsin
2FTL	Crystal structure of trypsin complexed with BPTI at 100K
2FTM	Crystal structure of trypsin complexed with the BPTI variant (Tyr35->Gly)
2JV0	SET domain of RIZ1 tumor suppressor (PRDM2)
2Z2C	MURA inhibited by unag-cnicin adduct
3AHS	Crystal Structure of Ustilago sphaerogena Ribonuclease U2B
3KQA	MurA dead-end complex with terreic acid
3KQJ	MurA binary complex with UDP-N-acetylglucosamine
3KR6	MurA dead-end complex with fosfomycin
3LTH	E. cloacae MurA dead-end complex with UNAG and fosfomycin
3SPB	Unliganded E. Cloacae MurA
3SU9	E. Cloacae MURA in complex with UDP-N-acetylmuramic acid and covalent adduct of PEP with Cys115
3SWA	E. Cloacae MurA R120A complex with UNAG and covalent adduct of PEP with CYS115
3SWD	E. coli MurA in complex with UDP-N-acetylmuramic acid and covalent adduct of PEP with Cys115
3SWI	E. Cloacae MurA in complex with Enolpyruvyl-UDP-N-acetylgalactosamine and covalent adduct of PEP with CYS115
3SWQ	E. Cloacae MurA in complex with Enolpyruvyl-UNAG
3UPK	E. cloacae MURA in complex with UNAG
3V4T	E. cloacae C115D MURA liganded with UNAG
3V5V	Unliganded E. cloacae C115D murA
4E7B	E. cloacae MurA in complex with UDP-glucose
4E7C	E. cloacae MurA in complex with UTP
4E7D	E. cloacae MurA in complex with UDP
4E7E	E. cloacae C115D MurA in complex with UDP-glucose
4E7F	E. cloacae C115D MurA in complex with UDP
4E7G	E. cloacae C115D/R120A MurA in the unliganded state
4EII	Unliganded E. cloacae R91K MurA
6H9V	Crystal structure of deaminated P domain from norovirus strain Saga GII-4 in complex with Fuc

rather a necessary modification for the protein, though this has not been probed experimentally (22). Another interesting structure is the L-isoAsp residue 115 in porcine trypsin (1C9P). The authors discovered that the density of their solution did not match that of the expected Asn115, but fit an L-isoAsp (23). Furthermore, this isomerized residue makes contacts to a Thr through two hydrogen-bonded waters. The authors state that these crystal contacts might be possible for a D-asp residue (seen in previous structures), but are enhanced by the L-isoAsp at this position. Thus they suggest that this site might also favor isomerization in order to improve contacts within the structure.

In a very recent study, researchers obtained L-asp and L-isoAsp structures of human Norovirus capsid protein and showed that the introduction of the isomerized residue disrupts the side chain orientation of an adjacent residue, which in turn affects this residues' interactions with the glycan, resulting in a lowered affinity for the ligand (24). Similarly, in studies of the bovine pancreatic ribonuclease and hen-egg white lysozyme, the introduction of the L-isoAsp residue causes large shifts of the local peptide backbone (~5 and ~7.5 Å, respectively) away from ligand binding regions, and in both cases causes decreases in substrate affinity (25, 26).

The structures discussed above support both functional and disruptive roles for the L-isoAsp modification within protein structures, which seem to generally correspond with the location of the isomerized residue relative to the binding/active site. In the cases of murA, porcine trypsin, and other examples such as the *Ustilago sphaerogena* ribonuclease U_2, where the identified L-isoAsp residues are not adjacent to the ligand binding site, the investigators note either neutral or beneficial effects (27). Yet, in the Norovirus capsid protein, bovine pancreatic ribonuclease, and hen-egg white lysozyme where the modified residue sits near the binding site, the result is lowered affinity. There may be examples in which a protein active site-adjacent L-isoAsp improves substrate recognition, but have yet to be fully characterized.

Plan of the Dissertation

The sum of the *Pcmt1* mice knockout studies, and *in vitro* biochemical and structural studies of the L-isoAsp residue underscores the significance of the PCMT1-L-isoAsp balance *in vivo*, as well as the variation in the biochemical effects upon L-isoAsp formation in proteins. To avoid an exhaustive cataloging of L-isoAsp in each protein of the body, in order to address the relevance of this modification in the context of human health, this thesis focuses on the systems in which we already know the L-isoAsp modification to be relevant: the brain and in long-lived proteins. This work also includes the beginning characterization of a novel L-isoAsp maintenance enzyme, which may complement the activity of the repair PCMT1 enzyme.

Chapters 2 and 3 revolve around the characterization of the nature and effects of L-isoAsp in the long-lived proteins of the lens in rodent and human lenses, respectively. In the research of Chapter 2, collaborators at the National University of La Plata provided our lab with teeth and eye samples from their colony of Sprague-Dawley rats. The average life span of this colony was 36 months, but our collaborators also provided us with samples from a “super-old” specimen of 44 months age. Proteins embedded in the mineralized matrices of the teeth were extracted using acid hydrolysis and analyzed for D- *vs.* L-asp content, while eye lenses were homogenized and evaluated for levels of L-isoAsp. My results showed no significant changes with age in D- *vs.* L-asp levels in proteins extracted from the teeth, but extremely high levels of L-isoAsp in the lenses of the aged rats, in spite of the endogenous repair PCMT1 activity being maintained with age.

In Chapter 3, I take these investigations of the eye lens further into the human lens. As in the rodent lens I found high levels of the L-isoAsp modification. In these samples I localized a majority of the damage in water-soluble extracts to high-molecular weight aggregates, specifically within small protein fragments present in the aggregates. In the cases of two water-soluble derived peptides the L-isoAsp modification appeared to decrease aggregative properties compared to native peptides. However, in the case of a water-insoluble derived peptide, the modification increased the tendency to aggregate and chaperone inhibition properties. Thus in a

similar manner to L-isoAsp investigated previously in other proteins, the effects of the modifications within lens protein fragments appear to differ from peptide to peptide.

Chapter 4 emphasizes this point in the characterization of the L-isoAsp modification in a peptide derived from the Alzheimer's disease-associated Amyloid-β. In this section I focused on an L-isoAsp site that has been previously identified in the literature, L-isoAsp23. The replacement of the normal residue with L-isoAsp in synthetic peptides of residues 20-34 greatly enhances aggregation of the peptide *in vitro*. Structural studies of the native and isomerized forms of these peptides revealed that both of the peptides form well ordered protofilament cores, in what we designate a "β-helix like turn," that interacts with a protofilament on either site through one of two distinct steric zipper interfaces. The L-isoAsp residue plays an important role in strengthening one of these two interfaces.

Within the last chapter I turn to the characterization of novel L-isoAsp recognition and maintenance pathways. I focus on a pair of human proteins called the protein carboxyl methyltransferase domain-containing proteins (PCMTDs), so called because their N-terminal sequences have 26% sequence identity with the PCMT1 enzyme. Most importantly, they conserve the cofactor *S*-adenosylmethionine (AdoMet) binding motifs and the L-isoAsp recognition motifs. However, both of these proteins also contain ~130 amino acids beyond the sequences that are homologous to PCMT1, which contain what is known as the "SOCS box" motif, which is known to recruit a complex of proteins which form cullin-RING ligases. These protein complexes form to bring protein substrates into close contact with ubiquitin-activated E2 proteins, which tag the substrate for degradation or other signaling purposes (28). The combinations of motifs in these protein sequences led to my hypothesis that the PCMTD proteins are L-isoAsp-recognizing E3 ligase adaptor proteins, which recognize L-isoAsp containing substrates and recruit E3 ligase machinery for ubiquitination of the substrate. We have been able to show that these proteins effect L-isoAsp levels in lysates, bind AdoMet, and associate with E3 ligase components. These experiments are summarized in Chapter 5.

Finally, in Chapter 6 I will summarize the results presented in this thesis and their significance within the context of the field, as well as lay out possible future directions.

References:

1. Walsh, Christopher T. (2006). Posttranslational modification of proteins: expanding nature's inventory. Englewood: Roberts and Co. Publ. ISBN 9780974707730.
2. Clarke, S. Aging as war between chemical and biochemical processes: protein methylation and the recognition of age-damaged proteins for repair. (2003) *Ageing Res. Rev.* **2**, 263-285.
3. Radi, R. Protein tyrosine nitration: biochemical mechanisms and structural basis of functional effects. (2013) *Acc. Chem. Res.* **64**, 550-559.
4. Vistoli, G., De Maddis, D., Cipak, A., Zarkovic, N. Carini, M. Aldini, G. Advanced glycoxidation and lipoxidation end products (AGEs and ALEs): an overview of their mechanisms of formation. (2013) *Free Radic. Res*. **47**, 3-27.
5. Kumar, A., Bachhawat, A.K. Pyroglutamic acid: throwing light on a lightly studied metabolite. (2012) *Current Sci*. **102**, 288-297.
6. Geiger, T., Clarke, S. Deamidation, Isomerization, and Racemization at Asparaginyl and Aspartyl Residues in Peptides. (1987) *J. Biol. Chem*. *282,* 785-794.
7. Patananan, A.N., Capri, J. Whitelegge, J.P., Clarke, S.G. Non-repair Pathways for Minimizing Protein Isoaspartyl Damage in the Yeast *Saccharomyces cerevisiae*. (2014) *J. Biol. Chem*. **289**, 16936-16953.
8. McFadden, P.N., Clarke, S. Conversion of isoaspartyl peptides to normal peptides: Implications for the cellular repair of damaged proteins. (1987) *Proc. Natl. Acad. Sci. USA*. **84**, 2595-2599.
9. Lowenson, J.D., Clarke, S. (1992) Recognition of D-aspartyl residues in polypeptides by the erythrocyte L-isoaspartyl/D-aspartyl protein methyltransferase. Implications for the repair hypothesis. *J. Biol. Chem.* **267**, 5985-5995.

10. Kim, E., Lowenson, J.D., MacLaren, D.C., Clarke, S., Young, S.G. Deficiency of a protein-repair enzyme results in the accumulation of altered proteins, retardation of growth, and fatal seizures in mice. (1997) *Proc. Natl. Acad. Sci. USA*. **94**, 6132-6137.
11. Lowenson, J.D., Kim, E., Young, S.G., Clarke, S. Limited accumulation of damaged proteins in l-isoaspartyl (D-aspartyl) O-methyltransferase-deficient mice. (2001) *J. Biol. Chem*. **276**, 20695-20702.
12. Farrar, C.E., Huang, C.S., Clarke, S.G., Houser, C.R. Increased cell proliferation and granule cell number in the dentate gyrus of protein repair-deficient mice. (2005) *J. Comp. Neurol.* **493**, 524-537.
13. MacKay, K.B., Lowenson, J.D., Clarke, S.G. Wortmannin Reduces Insulin Signaling and Death in Seizure-Prone Pcmt1-/1 Mice. (2012) *PLoS One*. **7**, e46719.
14. Vigneswara, V., Lowenson, J.D., Powell, C.D., Thakur, M., Bailey, K., Clarke, S., Ray, D.E., Carter, W.G. Proteomic identification of novel substrates of a protein isoaspartyl methyltransferase repair enzyme. (2006) *J. Biol. Chem.* **281**, 32619-32629.
15. Zhu, J.X., Doyle, H.A., Mamula, M.J., Aswad, D.W. Protein repair in the brain, proteomic analysis of endogenous substrates for protein L-isoaspartyl methyltransferase in mouse brain. (2006) *J. Biol. Chem***. 281**, 33802-33813.
16. Reissner, K.J., Aswad, D.W. Deamidation and isoaspartate formation in proteins: unwanted alterations or surreptitious signals? (2003) *Cell. Mol. Life. Sci*. **60**, 1281-1295.
17. Johnson, B.A., Langmack, E.L., Aswad, D.W. Partial repair of deamidation-damaged calmodulin by protein carboxyl methyltransferase. (1987) *J. Biol. Chem*. **262,** 12283-12287.
18. Brennan, T.V., Anderson, J.W., Jia, Z., Waygood, E.B., Clarke, S. Repair of spontaneously deamidated HPr phosphocarrier protein catalyzed by the L-isoaspartate-(D-aspartate) O-methyltransferase. (1994) *J. Biol. Chem*. **269**, 24586-24595.

19. Dimitrijevic, A., Qin, Z., Aswad, D.W. Isoaspartyl Formation in Creatine Kinase B Is Associated with Loss of Enzymatic Activity; Implications for the Linkage of Isoaspartate Accumulation and Neurological Dysfunction in the PIMT Knockout Mouse. (2014) *PLos One*. **9**, e100622.
20. Corti, A., Curnis, F. Isoaspartate-dependent molecular switches for integrin-ligand recognition. (2011) *J. Cell. Sci*. **124**, 515-522.
21. Curnis, F., Longhi, R., Crippa, L., Cattaneo, A., Dondossola, E., Bachi, A., Corti, A. Spontaneous formation of L-isoaspartate and gain of function in fibronectin. (2006) *J. Biol. Chem*. **281**, 36466-36476.
22. Eschenburg, S., Schönbrunn, E. Comparative X-ray analysis of the un-liganded fosfomycin-target murA. (2000) *Proteins*. **40**, 290-298.
23. Rester, U., Bode, W., Moser, M., Parry, M.A., Huber, R., Auerswald, E. Structure of the complex of the antistasin-type inhibitor bdellastasin with trypsin and modeling of the bdellastasin-microplasmin system. (1999) *J. Mol. Biol*. **293**, 93-106.
24. Mallagaray, A., Creutznacher, R., Dülfer, J., Mayer, P.H.O., Grimm, L.L., Orduña, J.M., Trabjerg, E., Stehle, T., Rand, K.D., Blaum, B.S., Uetrecht, C., Peters, T. A post-translational modification of human Norovirus capsid protein attenuates glycan binding. (2019) *Nat. comm*. **10**, 1320.
25. Caspasso, S., Di Donato, A., Esposito, L., Sica, F., Sorrentino, G., Vitagliano, L., Zagari, A., Massarella, L. Deamidation in proteins: the crystal structure of bovine pancreatic ribonuclease with an isoaspartyl residue at position 67. (1996) *J. Mol. Biol*. **257**, 492-496.
26. Noguchi, S., Miyawaki, K., Satow, Y. Succinimide and isoaspartate residues in the crystal structures of hen egg-white lysozyme complexed with tri-N-acetylchitotriose. (1998) *J. Mol. Biol*. **278**, 231-238.

27. Noguchi, S., Satow, Y., Uchida, T., Sasaki, C., Matsuzaki, T. Crystal structure of Ustilago sphaerogena ribonuclease U2 at 1.8 A resolution. (1995) *Biochemistry*. **34**, 15583-15591.

28. Kile, B.T., Schulman, B.A., Alexander, W.S., Nicola, N.A., Martin, H.M., Hilton, D.J. The SOCS box: a tale of destruction and degradation. (2002) *Trends Biochem. Sci*. **27**, 235-241.

CHAPTER 2

Racemized and Isomerized Proteins in Aging Rat Teeth and Eye Lens

The work described in this chapter has been reproduced from:

Rebeccah A. Warmack, Eduardo Mansilla E, Rodolfo G. Goya, Steven G. Clarke. “Racemized and isomerized proteins in aging rat teeth and eye lens.” *Rejuvenation Res*. 2016 Feb 26. **19,** 309-317.

REJUVENATION RESEARCH
Volume 19, Number 4, 2016

DOI: 10.1089/rej.2015.1778

Racemized and Isomerized Proteins in Aging Rat Teeth and Eye Lens

Rebeccah A. Warmack,[1] Eduardo Mansilla,[2] Rodolfo G. Goya,[3] and Steven G. Clarke[1]

Abstract

The quantification of aspartic acid racemization in the proteins of nonmetabolically active tissues can be used as a measure of chronological aging in humans and other long-lived organisms. However, very few studies have been conducted in shorter-lived animals such as rodents, which are increasingly used as genetic and metabolic models of aging. An initial study had reported significant changes in the ratio of D- to L-aspartate in rat molars with age. Using a sensitive HPLC method for the determination of D- and L-aspartate from protein hydrolysates, we found no accumulation of D-aspartate in the molars of 17 rats that ranged in age from 2 to 44 months, and the amount of D-aspartate per molar did not correspond with molar eruption date as had been previously reported. However, developing an alternate approach, we found significant accumulation of isomerized aspartyl residues in eye lens proteins that are also formed by spontaneous degradation processes. In this study, we used the human protein L-isoaspartate/D-aspartate *O*-methyltransferase (PCMT1) as an analytical reagent in a sensitive and convenient procedure that could be used to rapidly examine multiple samples simultaneously. We found levels of isomerized aspartyl residues to be about 35 times higher in the lens extracts of 18-month-old rats versus 2-month-old rats, suggesting that isomerization may be an effective marker for biological aging in this range of ages. Importantly, we found that the accumulation appeared to plateau in rats of 18 months and older, indicating that potentially novel mechanisms for removing altered proteins may develop with age.

Introduction

THE USE OF RODENT MODELS in aging research can provide information on how drugs, diet, or environment affect longevity and health.[1–4] This research often involves transgenic or knockout rodent lines for the analysis of a specific gene, protein, or phenotype. For these studies, it is useful to correlate age, genetic alterations, or therapeutic interventions with changes in molecular markers of aging.

L-Aspartate and L-asparagine residues in aging proteins are subject to a variety of spontaneous chemical reactions, including racemization, isomerization, and deamidation.[5,6] The accumulation of proteins containing these altered residues has been correlated with the decline in physiological function in the aging process.[7] These amino acid residues are particularly prone to degradation due to the electrophilicity of the γ-carbonyl carbon in their side chains. In less metabolically active biological tissues, such as bone, teeth, or eye lens, structural proteins may have lifetimes nearly as long as the organism itself and can accumulate these spontaneous modifications.[8] In organisms with long life spans such as humans, aspartic acid racemization has been a commonly used measure for age determination in forensic science.[9,10] The accuracy of age validation by D-aspartate accumulation for human tooth dentin has been estimated to be ± 3 years.[11] Similar studies have investigated the aspartic acid racemization in human bones and blood stains.[12,13]

However, very little work has been done to optimize these racemization dating methods for smaller shorter-lived animals, such as rats. Only two studies have been done within rodents investigating the correlation of aspartic acid racemization in teeth with the age of the animal. In 1995, Ohtani et al. reported significant differences between D-aspartate in rat molar enamel from rats ranging in age from 55 to 250 days.[14] They also reported that D/L-aspartate ratios corresponded with the eruption dates of the individual

[1]Department of Chemistry and Biochemistry, The Molecular Biology Institute, University of California, Los Angeles, Los Angeles, California.
[2]Tissue Engineering, Regenerative Medicine and Cell Therapies Laboratory, CUCAIBA, Buenos Aires Province Ministry of Public Health, Buenos Aires, Argentina.
[3]Institute for Biochemical Research (INIBIOLP)-Cathedra of Histology B, Cathedra of Pathology B, School of Medicine, National University of La Plata, La Plata, Argentina.

molars. Interestingly, the recorded D/L ratios showed aspartate racemization rates some ten times faster than those observed in humans, possibly reflecting a species-specific rate of racemization.[12,15] The second study was performed in senescence-accelerated mice using 2-, 6-, 12-, and 16-month-old animals. The D/L ratios reported from these whole molar samples also were shown to correlate with age.[16]

Human eye lens has also been analyzed for the accumulation of racemized aspartate derivatives. Like the mineralized collagens and associated proteins in teeth, the long-lived proteins of the eye lens, including α-, β- and γ-crystallin, can accumulate damaged residues. Masters et al. (1977) reported significant racemization rates within the human eye lens, with a linear correlation between age and the D/L aspartate ratio.[17] In the case of crystallins, these modifications are associated with insolubilization and aggregation.[18] The isomerization of aspartate and asparagine residues, resulting in the formation of L-isoaspartate and D-isoaspartate residues and linked with racemization through a common succinimide intermediate,[19,20] has been linked to cataract formation.[21] Although quantification of aspartate stereoisomers has been studied as another potential age marker in various other tissues,[22–24] this has only been indirectly studied in the eye lens.[21,25,26]

The studies described above demonstrate the time-dependent accumulation of aspartic acid derivatives and its correlation with age, particularly in long-lived species. With the prevalent use of mice and rats in aging and disease research, the development and optimization of methods used to investigate aging markers in these organisms are important. The goal of the work reported here was to explore the use of two age-related markers of aging in rats, racemization in molars and isomerization in the lens. Our results show that significant increases in racemization of aspartate/asparagine residues with age cannot be detected in molars, but that increases in the isomerization of these residues with age can be readily quantified in the lens.

Experimental Procedures

Collection of eye lens and teeth from aged rats

Sprague-Dawley (SD) rats were raised in the INIBIOLP rat colony at the National University of La Plata, Argentina. Animals were housed in a temperature-controlled room (22°C ± 2°C) on a 12-hour light/12-hour dark cycle. Food and water are available *ad libitum*. In this rat colony, the average 50% survival time for females, studied in groups of 50–60 animals, is approximately 32 months, provided mammary tumors are systematically removed when detected by palpation. The 2-month-old, 18-month-old, and some of the 29-month-old animals received no treatment. However, some of the 29-month-old animals received, 3 months before sacrifice, an introcerebroventricular injection of a placebo adenovector expressing the reporter gene for the DsRed2 red fluorescent protein and were fixed as described below. Assessment of spatial memory by the Barnes Maze test[27] as well as immunohistochemical analysis of glial fibrillary acidic protein (for astrocytes) and doublecortin immunoreactive cells in the hippocampus (for neuroblasts) revealed no significant differences between these placebo control adenovector-treated and untreated animals. Furthermore, the incidence of cataracts in both groups at the time of sacrifice was comparable (unpublished data). Thus, it is unlikely that the placebo adenovector treatment had a significant effect on the rate of eye lens or tooth protein racemization and isomerization. One rat (44 months old, designated SO) was intravenously injected with human bone marrow-derived mesenchymal stem cells as described.[28] Although the mesenchymal stem cell treatment seems to have extended the life span of the 44-month-old rat, it did not prevent the development of strong bilateral cataracts,[28] which again makes it improbable that the treatment had a significant impact on eye lens and tooth protein metabolism.

All animals used here were virgin females. Animals were sacrificed by rapid decapitation, except for the group of 29-month-old animals that were fixed. These animals were anesthetized with ketamine hydrochloride (70 mg/kg, i.p.) and xylazine (15 mg/kg, i.m.) and perfused with phosphate-buffered 4% paraformaldehyde (pH 7.4). After removal of the brain, the heads were frozen at −80°C and shipped to UCLA packed in dry ice. All experiments with animals were performed in accordance with the Animal Welfare Guidelines of NIH (INIBIOLP's Animal Welfare Assurance No. A5647-01); IACUC approved on April 3, 2007, extended December 19, 2011.

At UCLA, heads were thawed at 4°C for 3 days in 300 mL of deionized water. The jaws were separated and molars were removed using a burlisher. Excess soft tissue was removed and the molars were then cleaned in an ultrasonic water bath for 20 min in deionized water, followed by 20 min in ethanol, and vacuum-dried. Eyes were excised using surgical scissors and lenses were extracted under a Zeiss Stemi 2000 light microscope. Lenses were washed gently in 500 μL of deionized water in 1.5-mL microcentrifuge tubes on a rotator for 20 minutes at 4°C. Rat heads for lens extraction were not subjected to any tissue fixation.

Determination of D/L-aspartate ratios in molar hydrolysates

After drying, intact individual molars were acid hydrolyzed in 500 μL of 6 N HCl at 110°C for 6 hours under vacuum in a Waters PicoTag apparatus. Samples were vacuum-dried, resuspended in 50 μL of deionized water, and any remaining particulate material was removed by centrifugation. Analysis of *o*-phthalaldehyde/*N*-acetyl-L-cysteine (OPA-NAC) derivatives of aspartic acid was performed after the method of Aswad.[29] After a 20-fold dilution in water, 5 μL of the hydrolysate was added to 20 μL of the OPA-NAC reagent containing 4 mg *o*-phthalaldehyde (MP Biomedicals, LLC #102648) dissolved in 300 μL methanol, 250 μL 0.4 M potassium borate pH 10.3, 390 μL deionized water, and 60 μL 1 M *N*-acetyl-L-cysteine and was incubated for 3 minutes at room temperature. The reaction was stopped with the addition of 40 μL of 0.2 M sodium citrate, pH 5.5, and the solution analyzed on an HP Series II 1090 HPLC, using a 5 μm Agilent Eclipse C18 reverse-phase column (150 mm in length, 4.6 mm inner diameter) at 25°C. Solvent A was 50 mM Na acetate, pH 5.4. Solvent B was 80% (v/v) methanol and 20% (v/v) solvent A. Isocratic elution with 90% solvent A: 10% solvent B was carried out for 5 minutes. Then, solvent B was increased linearly to 100% over 5 minutes and maintained at 100% solvent B for 15 minutes. The flow rate was 1.0 mL/min throughout. Fluorescent derivatives were monitored using a Gilson Model 121 fluorimeter with a 50 watt halogen EPZ

projector lamp, an excitation filter of 305–395 nm (Gilson catalog number 095312), and a 450 wide-band emission filter (Gilson catalog number 095442). The RFU sensitivity was set at 0.05. The D-aspartate/L-aspartate ratio was calculated by integrated area under the appropriate peaks after correction for the background.

Determination of L-isoaspartate levels in eye lens proteins

Lenses from unfixed rat heads were thawed as described above, weighed, and then 20 μL of T-PER Total Protein Extraction Reagent (Life Technologies 78510) was added per mg wet weight of lens at room temperature. Sodium EDTA, pH 7.5, and phenylmethanesulfonyl fluoride dissolved in isopropanol were added to final concentrations of 1 mM as protease inhibitors. Lenses were homogenized at room temperature in a 1.5-mL microcentrifuge tube with 200 strokes of a form-fitting plastic pestle (Kimble-Chase Kontes pellet pestle 7495150000; Fisher Scientific). Samples were centrifuged at 10,000 *g* for 5 minutes at room temperature to pellet unbroken cells and membrane debris. The supernatant was collected and stored at −20°C before further analyses.

The human recombinant protein L-isoaspartate/D-aspartate *O*-methyltransferase (PCMT1)—an enzyme that recognizes L-isoaspartate residues with high affinity[30]—was used as an analytical reagent to quantify L-isoaspartate levels in the lens extract proteins. In a final volume of 100 μL, 2.5 to 5 μg of lens extract protein (as determined by a Lowry assay after precipitation with trichloroacetic acid) was incubated for 2 hours at 37°C with 5 μg PCMT1 (purified as a His-tagged enzyme from *Escherichia coli* containing the expression plasmid #34852 available from Addgene.com as described by Patananan et al.,[31] with a specific activity at 37°C of 3,361 pmol of methyl esters formed on ovalbumin/min/mg of enzyme) and final concentrations of 135 mM Bis-Tris-HCl, pH 6.4, and 10 μM *S*-adenosyl-L-[*methyl*-^{3}H]methionine ([^{3}H]AdoMet) (prepared by a 1600-fold isotopic dilution of a stock of 72 Ci/mmol [^{3}H]AdoMet (PerkinElmer Life Sciences, NET155H00) with nonisotopically labeled AdoMet (*p*-toluenesulfonate salt; Sigma-Aldrich A2408)). The reaction was stopped by adding 10 μL of 2 M sodium hydroxide, and 100 μL of the 110 μL mixture was transferred to a 9×2.5 cm piece of folded thick filter paper (Bio-Rad; catalog number 1650962), wedged in the neck of a 20-mL scintillation vial above 5 mL scintillation reagent (Safety Solve, Research Products International, catalog number 121000), tightly capped, and incubated at room temperature. After 2 hours, the folded filter papers were removed, the caps replaced, and the vials were counted thrice for 3 minutes each in a Beckman LS6500 scintillation counter. Background radioactivity in a no substrate blank was determined by incubating the recombinant human PCMT1, 135 mM Bis-Tris-HCl buffer, and 10 μM [^{3}H]AdoMet as described above and was subtracted from the value obtained in experimental samples. Samples were analyzed in triplicate.

Rat eye lens endogenous protein L-isoaspartyl methyltransferase activity assays

In a final volume of 100 μL, 10 μL of lens extract protein was incubated for 2 hours at 37°C with final concentrations of 100 μM KASA(isoD)LAKY peptide, 125 mM Bis-Tris-HCl, pH 6.4, and 10 μM [^{3}H]AdoMet, as prepared above. The reaction was stopped by adding 10 μL of 2 M sodium hydroxide, and 100 μL of the 110 μL mixture assayed for volatile radioactivity, as described in the section above. Background radioactivity was determined in a control lacking the lens extract and was subtracted from the value obtained in samples containing the lens extracts. Assays were performed in duplicate.

Results and Discussion

D-Aspartate accumulation in teeth cannot be used to determine rat age

Tooth enamel is 96% mineralized and contains only 4% protein.[32] Most of the protein that is present is trapped in the solid matrix of hydroxyapatite and lasts for years with little to no metabolic turnover. Dentin is similarly solid and is approximately 70% mineralized. Given that the entrapped proteins can accumulate spontaneous chemical modifications, enamel and dentin can be useful for age determination through the quantification of protein damage.

D-Aspartate accumulation has been well correlated with age in human tooth enamel and dentin.[33,34] Human teeth are relatively large, and the separation of enamel and dentin is a well-documented procedure.[32,35] Ohtani et al.[14] reported the successful application of these techniques to rat molar tooth enamel, a potentially exciting finding as rats live much shorter lives than humans, and rat teeth (4–10 mg) are much smaller than those in humans (500 to 2500 mg). In this article, the enamel was apparently separated from the rest of the molar with a rotary saw, and after hydrolysis, samples were analyzed for D- and L-aspartate by gas chromatography.

To optimize rodent age determination using reverse-phase liquid chromatography, attempts were made to replicate the removal of the dental crowns from the molars as described in Ohtani et al.[14] However, the very small size of the rat teeth and the distribution of the enamel prevented separation in our hands. Personal communication with Ohtani revealed that in fact whole molars were used for their analyses as they also found that the teeth were too small for the separation of enamel and dentin and that the description of the separation was incorrect in their article. Thus, we also analyzed whole molars in this study.

After extraction and acid hydrolysis of the whole molar, OPA-NAC derivatives of amino acids were detected and quantified through RP-HPLC. Representative chromatographs can be seen in Figure 1, with D- and L-aspartate eluting at approximately 7 and 8 minutes, respectively, with baseline resolution (Fig. 1D). The D-aspartate peak was confirmed by running a hydrolysate alone (Fig. 1A), followed by a run containing half the amount of hydrolysate with added D-aspartate standard. In this study, the D-aspartate peak increased in area by sixfold, while all other peak areas were halved (Fig. 1B). Unhydrolyzed D- and L-Asp standards were also run to verify the elution times of the peaks of interest (Fig. 1C). These results demonstrate the reproducible separation and identification of D- and L-aspartate levels within the rat teeth with RP-HPLC analysis of OPA-NAC amino acid derivatives.

Respective levels of D- and L-aspartate were then quantified to calculate the ratios shown in Figure 2A, representing the amounts of aspartate racemization found in the individual

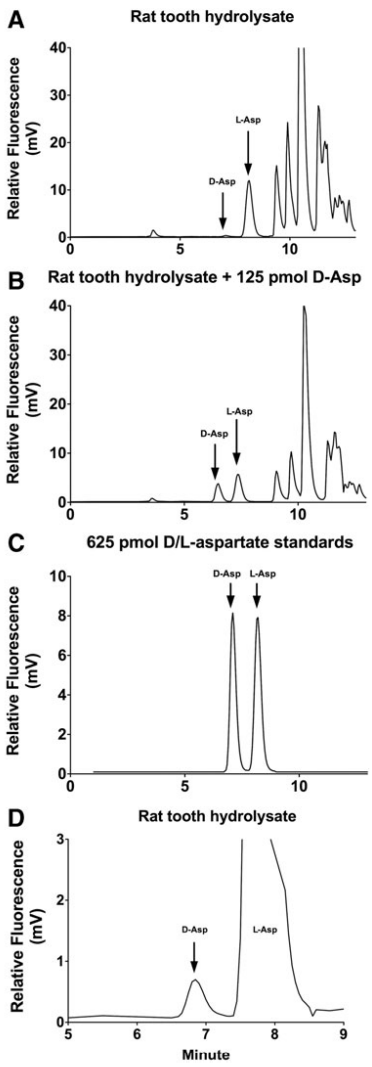


FIG. 1. Determination of the D-aspartate/L-aspartate ratio in rat tooth molar hydrolysates. **(A)** A representative chromatograph from reverse-phase HPLC analysis of OPA-NAC amino acid derivatives with fluorescence detection from a tooth hydrolysate of the first molar of a 29-month-old rat that was processed as described in the Experimental Procedures section. The position of D-Asp and L-Asp was ascertained by comparison of the elution positions of standards, chromatographed separately (not shown). **(B)** HPLC analysis of a mixture of half of the amount of tooth hydrolysate chromatographed in **(A)** with 125 pmol of a D-Asp standard. **(C)** An unhydrolyzed mixture of 625 pmol of D-Asp and 625 pmol of L-Asp was analyzed as described above. **(D)** An enlarged view of a portion of **(A)** showing the separation of the derivatives of D- and L-aspartate.

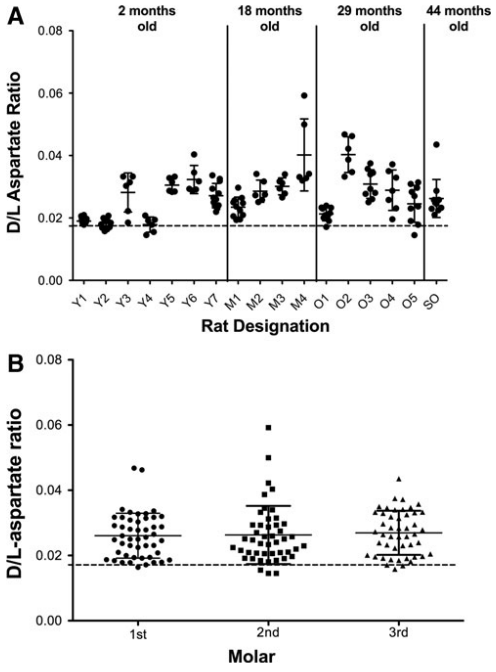


FIG. 2. D-Aspartate/L-aspartate ratios in rat molars do not significantly correlate with rat age or with the molar position. **(A)** Determination of D/L aspartate ratios of rat molar hydrolysates was performed as described in the Experimental Procedures section and Figure 1. Sets of first, second, and third molars were analyzed from seven young rats (Y1 to Y7; 2 months old), four middle-aged rats (M1 to M4; 18 months old), five old rats (O1 to O5; 29 months old), and one exceptionally old rat (S01; 44 months old). For some rats, independent analyses were made of each first, second, and third molar from the upper and lower jaw, as well as the left and right side, giving 12 samples per rat. In other cases, only molars from the right or left side were analyzed, giving 6 samples per rat. In a few cases, teeth were unavailable due to shattering during extraction from the jaw. Error bars indicate the standard deviation of the D/L aspartate ratio from all of the molars of an individual rat. A two-tailed Student's *t*-test with unpaired samples and unequal variance between the D/L-aspartate ratios of the groups of 2-, 18-, 29-, and 44-month-old samples revealed no significant differences in the D/L aspartate ratio (p-value >0.05). The dashed horizontal line shows the background racemization observed (D/L-aspartate ratio = 0.0174) when a sample of L-aspartic acid was acid hydrolyzed and analyzed under our conditions. **(B)** D/L-aspartate ratios plotted for each of the first, second, and third molars for the data shown in **(A)**. Error bars indicate the standard deviation.

rat molars. Statistical analysis between the different age groups did not reveal significant differences in the D/L-aspartate ratios in contrast to what had been reported previously.[14] In these experiments, four of the rat heads from the group at 29 months (designated O2, O3, O4, and O5 in Fig. 2A) had been fixed with 4% paraformaldehyde, while the

other rat head (designated O1) was simply frozen before analysis. We saw no differences in the degree of racemization of this group and the other age groups, suggesting that fixation did not affect this process.

Rat molars are known to develop in a predictable manner, with the front molar erupting at approximately 19 days after birth, the second set on the 21st day, and the last set 35–40 days after birth.[36] Thus, in addition to exploring a relationship to age, we investigated the link between racemization and eruption date. Data for individual teeth were plotted according to the particular molar, shown in Figure 2B, and demonstrated that aspartate racemization does not significantly correspond with molar age relative to other molars in the mouth, contrary to the results reported by Ohtani et al.[14]

The discrepancies seen between our results and those in Ohtani et al.[14] may have several origins. Our sampling included multiple rats for each age group excepting the 44-month-old rat, including seven 2-month-old rats, four 18-month-old rats, and five 29-month-old rats. The Ohtani group analyzed five individual rats of different ages ranging from about 2 months to 8 months. The increase in sample size may accurately represent the variability between rats of the same age group. Our age groups also spanned a much longer length of time.

It is also possible that there was some confusion in the Ohtani et al. publication[14] since the techniques reported for the separation of molar enamel and dentin are not possible under the methods given. However, we demonstrate that the variability between rats of the same age prevents this method from reliably identifying rat age or from using D-aspartate levels as a marker of aging. Finally, our results suggest that the rates of racemization of rat molar proteins are probably not in fact markedly higher than in human teeth, as was indicated by the data of Ohtani et al.[14]

These results indicate that the rate of racemization of rat tooth proteins is slow enough that we cannot distinguish the

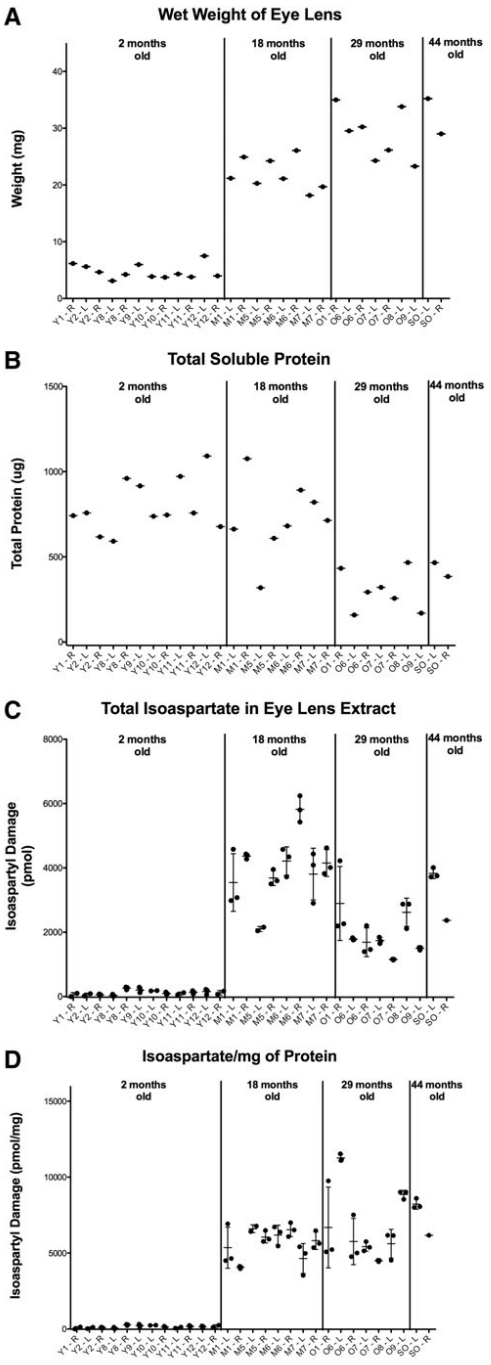


FIG. 3. L-isoaspartate accumulates in the rat eye lens after 2 months, but plateaus after 18 months of age. Rat eye lenses were extracted from seven young rats (2 months), four middle-aged rats (18 months), five old rats (29 months), and one exceptionally old rat (44 months), including several of the animals whose teeth were analyzed in Figure 2. Rats were identified as in Figure 2; R and L represent the left and right lenses. Cytosolic extracts were made and L-isoaspartate content measured as described in the Experimental Procedures section. **(A)** Wet weight of eye lens. **(B)** Data are shown for the total average amount of soluble protein from homogenized lens extract determined by triplicate Lowry protein assays after precipitation with 10% trichloroacetic acid. Student's *t*-test performed as in Figure 2 shows a *p*-value of 0.0007 for the averages of the technical replicates of the 18-month-old versus 29-month-old samples. **(C)** Data are shown for triplicate replicates of the total L-isoaspartate content in each extract. Error bars indicate standard deviation. Student's *t*-test shows a *p*-value of 0.0005 for the average values of the 18-month-old versus 29-month-old samples. **(D)** L-isoaspartate per mg protein. Error bars indicate standard deviation. No significant difference was found in the average values for the 18-month-old and 29-month-old samples ($p > 0.05$).

age-dependent accumulation above the background of spontaneous racemization resulting from the acid hydrolysis procedure. The acid hydrolysis conditions have been optimized to ensure near complete hydrolysis of peptide bonds linked to aspartic acid and asparagine residues while minimizing the background racemization.[23]

L-isoaspartate levels in rat eye lens correspond with increased age

The degradation of L-aspartate and L-asparagine residues proceeds mainly through an L-succinimide intermediate.[5,6,19] The rate of L-isoaspartyl formation from this intermediate is much greater than the rate of D-succinimide formation, resulting in a much faster accumulation of isomerized residues than racemized residues.[19] Thus, damage to proteins at these residues may be better approached by analyzing L-isoaspartate.

We took advantage of the availability of an enzyme that specifically methylates both L-isoaspartate and D-aspartate residues, the human protein carboxyl methyltransferase 1 (PCMT1).[20] This enzyme recognizes L-isoaspartate residues with great preference over D-aspartate residues, resulting in its ability to serve as an analytical reagent for L-isoaspartyl residues.[30,31,37 39] Although the methyltransferase has not be used in mineralized tissues such as tooth, it has been utilized to detect abnormal aspartyl residues that accumulate with age in mouse brain, heart, and testis proteins.[40] Given the slow protein turnover rates of the eye lens tissue and the higher abundance of proteins than in rat teeth, we hypothesized that L-aspartate isomerization to L-isoaspartate would be a better marker of age than protein racemization to D-aspartate.

Previous reports have indicated that the rat eye lens grows throughout life, but at a declining rate over the lifetime of the organism,[41,42] a situation distinct from that seen in the human eye lens where most of the development occurs before one year of age and only increases very slowly thereafter.[43] In fact, lens wet weight has been used as a marker of rat age.[42,44 46] From the results seen in Figures 3A and 4A, we also demonstrated an age-dependent increase in eye lens wet weight. However, it seems that as the lens grows, the amount of soluble protein decreases, suggesting the formation of damaged and aggregated proteins, dropping down from about 1000 μg in a 2-month-old rat to just below 500 μg in the 44-month-old rat (Figs. 3B and 4B).

Figures 3C, 3D, 4C, and 4D show the quantification of the levels of L-isoaspartate in proteins of the soluble extracts of lens from the vapor diffusion assay. From 2 months of age to 44 months, the level of isoaspartate per mg of protein (Figs. 3D and 4D) increases from an average of about 140 to 7200 pmol isoaspartate/mg eye lens protein, corresponding to about 0.00018 to 0.0090 residues of isoaspartate per residues of aspartate and asparagine. The largest differences occur between 2 months and 18 months of age. While the total levels of isoaspartate appear to be maximal at 18 months of age and decreasing at older ages following the decreases in soluble protein (Figs. 3B, C and 4B, C), the amount of isoaspartyl residues/mg protein appear to plateau after 18 months of age (Fig. 3D and 4D). Thus, as the eye lens ages, there is a significant increase in the isoaspartate level from 2 to 18 months, followed by decreases in both soluble protein and total isoaspartate in soluble proteins in

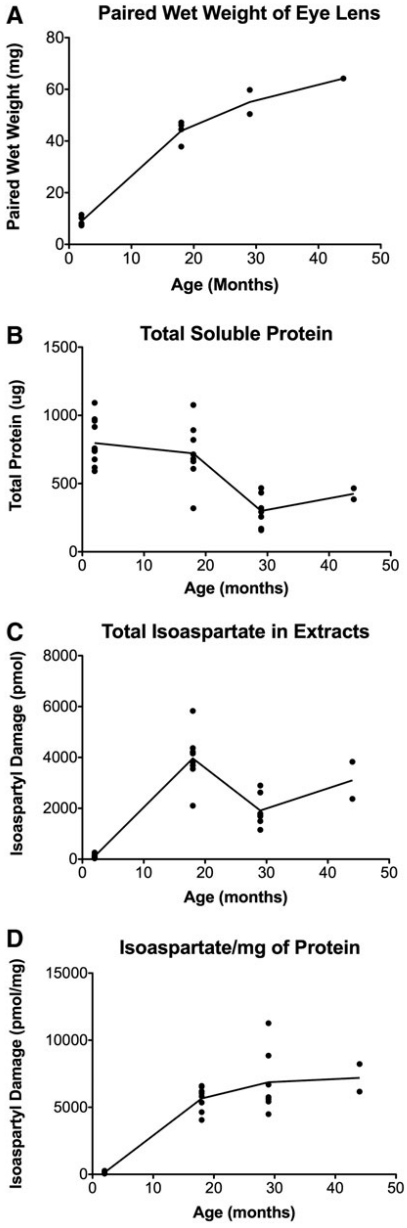


FIG. 4. The data from Figure 3 are plotted as a function of the age of the individual rats. In **(A)**, the wet weights of the left and right lenses were added for each rat where both lenses were available. In **(B–D)**, only the average values are shown.

the 29-month-old and 44-month-old animals. These results lead to a plateau in the amount of isoaspartate/mg protein by 18 months of age. The link between the amount of soluble protein and isoaspartate is not clear, but it is possible that isomerization itself may lead to aggregation. Interestingly, a similar situation may occur in human eye lens, where the endogenous PCMT1 enzyme is much less saturated with endogenous substrates before 30 years of age than after 50 years of age.[26]

The plateau/decrease in total L-isoaspartate levels in soluble proteins and in the ratio of L-isoaspartate to total protein seen in the data of Figures 3C, 3D, 4C, and 4D by 18 months of age could have several origins. In the first place, lens tissues are known to contain endogenous PCMT1.[26] This enyzme catalyzes the first step of a pathway by which L-isoaspartate residues are converted to L-aspartate residues.[20,30] It is possible that the activity of this enzyme is significantly increased in the older lenses, thus maintaining L-isoaspartate at a constant level.

To investigate this possibility, the specific activity of the endogenous PCMT1 enzyme was assayed in the eye lens extracts using saturating amounts of an L-isoaspartyl-containing peptide, KASA(isoD)LAKY (Fig. 5). Although we found a slight increase in the specific activity between the 2-month-old and 18-month-old samples, there was no further increase after 18 months and perhaps a decrease. Thus, it is unlikely that the plateau in the level of isoaspartate within the eye lens is due to increasing repair activity.

Since we observed that the levels of total soluble protein decreased over the lifetime of the rats, a second possibility is that the formation of aggregates may prevent the accumulation of L-isoaspartate residues. Finally, it is possible that other repair or maintenance pathways compensate in older animals by removing proteins with L-isoaspartate residues. The yeast *Saccharomyces cerevisiae*, for example, has no known homolog of the PCMT1 repair enzyme and yet is able to maintain low levels of L-isoaspartate, suggesting the presence of other repair or degradation mechanisms, including a possible role of metalloproteases.[31] It is conceivable then that proteases in the eye lens may recognize the high levels of aggregates or isomerized proteins and degrade these species to peptides and or amino acids.

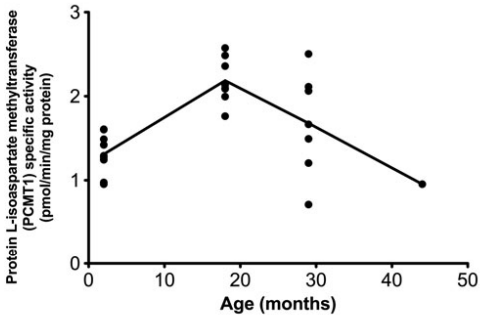


FIG. 5. Age dependence of the specific activity of endogenous protein L-isoaspartate methyltransferase in rat eye lens. Enzyme activity with an L-isoaspartyl-containing peptide was performed as described in the Experimental Procedures section. Each point represents the average of duplicate measurements. The connecting line shows the mean value of each age group. Student's *t*-test performed as in Figure 2 indicated that the only significant difference in the age groups was between the 2-month-old and 18-month-old groups ($p=4\times10^{-6}$).

In mice, it was similarly observed that in heart and testis extracts the amount of isoaspartate initially accumulates rapidly, and then levels off, suggesting again possible compensation mechanisms.[40] It is possible that the number of proteins with labile aspartate/asparagine residues is limited and that the plateau is due to all possible sites being racemized or isomerized. However, within an mg of protein, there are approximately 800 nmols of asparagine and aspartate residues, which may racemize or isomerize. The maximum level of damage that was observed in our study was about 12 nmol isoasparate/mg protein (Fig. 3D), suggesting that we are only observing approximately 1.5% of sites forming aberrant derivatives.

As stated previously, whole tooth analysis for D-aspartate can validate a human individual's age to within 3 years. This variability exists, in part, due to the relatively slow rate of racemization and the background of spontaneous racemization that occurs during the acid hydrolysis step. In humans, the long life span can overcome most of these issues, simply because the age groups span across a greater length of time, which allows for the development of discernable levels of D-aspartate in teeth. Within rodents, however, where the life span is typically 2.5–3.5 years, a range of ±3 years would not allow for accurate age determination. However, by taking advantage of L-isoaspartate's much more rapid rate of accumulation and measuring the accumulated levels of L-isoaspartate within eye lens extracts, we have shown clear differences between young and old rats. The relative ease of the preparation for these samples, as well as the more distinct differences in the measurements made, makes this technique an attractive alternative to previously reported methods.

Acknowledgments

The authors thank Dr. Aubrey D. N. J. de Grey of the SENS Research Foundation (Mountain View, CA) for initiating this work and for providing guidance throughout. The authors thank Dr. Shane White of the UCLA School of Dentistry for his help in preparing tooth samples and Drs. Rajendra Gangalum and Suraj Bhat of the UCLA Jules Stein Eye Institute for assistance and advice on the preparation of lens samples. They thank Dr. Jonathan Lowenson for his guidance in this work and for his help in preparing the manuscript. This work was supported by a grant from the Life Extension Foundation, Inc., and funds from the Elizabeth and Thomas Plott Chair in Gerontology of the UCLA Longevity Center. R.A.W. was supported by the Ruth L. Kirschstein National Research Service Award GM007185.

R.G.G. and E.M. provided the animal material; E.M. took responsibility for monitoring the health of the 44-month-old rat; R.G.G., R.A.W., and S.G.C. planned the study; R.A.W. performed the analytical experiments; R.A.W. and S.G.C. wrote the article and analyzed the data; and R.A.W., E.M., R.G.G., and S.G.C. reviewed and revised the article.

Author Disclosure Statement

No competing financial interests exist.

References

1. Lebel M, Picard F, Ferland G, Gaudreau P. Drugs, nutrients, and phytoactive principles improving the health span of rodent models of human age-related diseases. J Gerontol A Biol Sci Med Sci 2012;67:140–151.
2. Spindler SR, Mote PL, Flegal JM, Teter B. Influence on longevity of blueberry, cinnamon, green and black tea, pomegranate, sesame, curcumin, morin, pycnogenol, quercetin, and taxifolin fed iso-calorically to long-lived, F1 hybrid mice. Rejuvenation Res 2013;16:143–151.
3. Foster TC, DeFazio RA, Bizon JL. Characterizing cognitive aging of spatial and contextual memory in animal models. Front Aging Neurosci 2012;4:12.
4. Stefanova NA, Kozhevnikova OS, Vitovtov AO, Maksimova KY, Logvinov SV, Rudnitskaya EA, Korbolina EE, Muraleva NA, Kolosova NG. Senescence-accelerated OXYS rats: A model of age-related cognitive decline with relevance to abnormalities in Alzheimer disease. Cell Cycle 2014;13:898–909.
5. Radkiewicz JL, Zipse H, Clarke S, Houk KN. 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;123:3499–3506.
6. Wakankar AA, Borchardt RT. Formulation considerations for proteins susceptible to asparagine deamidation and aspartate isomerization. J Pharm Sci 2006;95:2321–2336.
7. Nowotny K, Jung T, Grune T, Höhn A. Accumulation of modified proteins and aggregate formation in aging. Exp Gerontol 2014;57:122–131.
8. Sakuma A, Ohtani S, Saitoh H, Iwase H. Comparative analysis of aspartic acid racemization methods using whole-tooth and dentin samples. Forensic Sci Int 2012;223:198–201.
9. Ritz-Timme S, Rochholz G, Schütz HW, Collins MJ, Waite ER, Cattaneo C, Kaatsch HJ. Quality assurance in age estimation based on aspartic acid racemization. Int J Legal Med 2000;114:83–86.
10. Arany S, Ohtani S, Yoshioka N, Gonmori K. Age estimation from aspartic acid racemization of root dentin by internal standard method. Forensic Sci Int 2004;141:127–130.
11. Waite ER, Collins MJ, Ritz-Timme S, Schutz HW, Cattaneo C, Borrman HI. A review of the methodological aspects of aspartic acid racemization analysis for use in forensic science. Forensic Sci Int 1999;103:113–124.
12. Collins MJ, Waite ER, van Duin AC. Predicting protein decomposition: The case of aspartic-acid racemization kinetics. Philos Trans R Soc Lond B Biol Sci 1999;354: 51–56.
13. Arany S, Ohtani S. Age estimation of bloodstains: A preliminary report based on aspartic acid racemization rate. Forensic Sci Int 2011;212:e36–e39.
14. Ohtani S, Matsushima Y, Ohhira H, Watanabe A. Age-related changes in D-aspartic acid of rat teeth. Growth Dev Aging 1995;59:55–61.
15. Ohtani S, Yamamoto T. Age estimation by amino acid racemization in human teeth. J Forensic Sci 2010;55: 1630–1633.
16. Ohtani S, Yamada Y, Yamamoto I, Marumo T, Sugeno H, Sugimoto H, Ogasawara A, Yamagishi M. Estimation of age from teeth using the racemization of aspartic acid (racemization method). Nihon Hoigaku Zasshi 2000;54:207–218.
17. Masters PM, Bada JL, Zigler JS Jr. Aspartic acid racemisation in the human lens during ageing and in cataract formation. Nature 1977;268:71–73.
18. Vasavada AR, Praveen MR, Jani UD, Shah SK. Preoperative prediction of posterior capsule plaque in eyes with posterior subcapsular cataract. Indian J Ophthalmol 2006;54:169–172.
19. Geiger T, Clarke S. Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. J Biol Chem 1987;262:785–794.
20. Johnson BA, Murray ED Jr., Clarke S, Glass DB, Aswad DW. Protein carboxyl methyltransferase facilitates conversion of atypical L-isoaspartyl peptides to normal L-aspartyl peptides. J Biol Chem 1987;262:5622–5629.
21. 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;287:39992–40002.
22. Ghandour H, Lin B, Choi S, Mason JB, Selhub J. Folate status and age affect the accumulation of L-isoaspartyl residues in rat liver proteins. J Nutr 2002;132:1357–1360.
23. Brunauer LS, Clarke S. Age-dependent accumulation of protein residues which can be hydrolyzed to D-aspartic acid in human erythrocytes. J Biol Chem 1986;261: 12538–12543.
24. Fonseca MI, Head E, Velazquez P, Cotman CW, Tenner AJ. The presence of isoaspartic acid in β-amyloid plaques indicates plaque age. Exp Neurol 1999;157:277–288.
25. Fujii N, Takata T, Fujii N. Quantitative analysis of isomeric (l-α-, l-β-, d-α-, d-β-) aspartyl residues in proteins from elderly donors. J Pharm Biomed Anal 2015; 116: 25–33.
26. McFadden PN, Clarke S. Protein carboxyl methyltransferase and methyl acceptor proteins in aging and cataractous tissue of the human eye lens. Mech Ageing Dev 1986;34: 91–105.
27. Barnes CA. Memory deficits associated with senescence: A neurophysiological and behavioral study in the rat. J Comp Physiol Psychol 1979;93:74–104.
28. Mansilla E, Roque G, Sosa YE, Tarditti A, Goya RG. A Rat treated with mesenchymal stem cells lives to 44 months of age. Rejuvenation Res 2016;19:318–321.
29. Aswad DW. Determination of D- and L-aspartate in amino acid mixtures by high-performance liquid chromatography after derivatization with a chiral adduct of o-phthaldialdehyde. Anal Biochem 1984;137:405–409.
30. Lowenson JD, Clarke S. Structural elements affecting the recognition of L-isoaspartyl residues by the L-isoaspartyl/D-aspartyl protein methyltransferase. Implications for the repair hypothesis. J Biol Chem 1991; 266:19396–19406.
31. Patananan AN, Capri J, Whitelegge JP, Clarke SG. Non-repair pathways for minimizing protein isoaspartyl damage in the yeast *Saccharomyces cerevisiae*. J Biol Chem 2014; 289:16936–16953.
32. Nanci A. *Ten Cate's Oral Histology*. St. Louis: Elsevier Mosby, 2013.
33. Helfman PM, Bada JL. Aspartic acid racemization in tooth enamel from living humans. Proc Nat Acad Sci USA 1975;72:2891–2894.
34. Helfman PM, Bada JL. Aspartic acid racemisation in dentine as a measure of ageing. Nature 1976;262:279–281.
35. Griffin RC, Chamberlain AT, Hotz G, Penkman KE, Collins MJ. Age estimation of archaeological remains using amino acid racemization in dental enamel: A comparison of morphological, biochemical, and known ages-at-death. Am J Phys Anthropol 2009;140:244–252.

36. Schour I, Massler M. The teeth. In: Farris EJ, Griffith, JQ Jr. (eds): *The Rat in Laboratory Investigation*. Philadelphia: Lippincott Co., 1949.
37. Johnson BA, Aswad DW. Optimal conditions for the use of protein L-isoaspartyl methyltransferase in assessing the isoaspartate content of peptides and proteins. Anal Biochem 1991;192:384–391.
38. Lowenson JD, Clarke S. Recognition of D-Aspartyl Residues in Polypeptides by the Erythrocyte L-Isoaspartyl/D-Aspartyl Protein Methyltransferase–Implications for the Repair Hypothesis. J Biol Chem 1992;267:5985–5995.
39. Qin Z, Dimitrijevic A, Aswad DW. Accelerated protein damage in brains of PIMT+/- mice; a possible model for the variability of cognitive decline in human aging. Neurobiol Aging 2015;36:1029–1036.
40. Lowenson JD, Kim E, Young SG, Clarke S. Limited accumulation of damaged proteins in l-isoaspartyl (D-aspartyl) O-methyltransferase-deficient mice. J Biol Chem 2001;276:20695–20702.
41. Bours J, Hockwin O, Fink H. Biochemistry of the ageing rat lens: Lens wet weight and lens dry weight with respect to sex differences. Ophthalmic Res 1983;15:198–203.
42. Augusteyn RC. Growth of the eye lens: I. Weight accumulation in multiple species. Mol Vis 2014;20:410–426.
43. Augusteyn RC. Growth of the human eye lens. Mol Vis 2007;13:252–257.
44. Hardy AR, Quy RJ, Huson LW. Estimation of age in the Norway rat (*Rattus norvegicus* Berkenhout) from the weight of the eyelens. J Appl Ecol 1983;20: 97–102.
45. Vejaratpimol R, Liat LB. Age estimation of *Rattus argentiventer* from eye lens weight. J Sci Soc Thailand 1983;9: 107–118.
46. Tanikawa T. An eye-lens weight curve for determining age in black rats, *Rattus rattus*. J Mamm Soc Jpn 1993;18: 49–51.

Address correspondence to:
Steven G. Clarke
Department of Chemistry and Biochemistry
The Molecular Biology Institute
University of California
607 Charles E. Young Drive East
Los Angeles, CA 90095-1569

E-mail: clarke@mbi.ucla.edu

Received: September 22, 2015
Accepted: December 2, 2015

CHAPTER 3

The L-isoaspartate modification within protein fragments in the aging lens can promote protein aggregation

The L-isoaspartate modification within protein fragments in the aging lens can promote protein aggregation

Rebeccah A. Warmack, Harrison Shawa, Kate Liu, Katia Lopez, Joseph A. Loo, Joseph Horwitz, Steven G. Clarke

ABSTRACT

Transparency in the lens is accomplished by the dense packing and short-range order interactions of the crystallin proteins in fiber cells lacking organelles. These features are accompanied by a lack of protein turnover, leaving lens proteins susceptible to a number of damaging modifications and aggregation. The loss of lens transparency is attributed in part to such aggregation during aging. Among the damaging post-translational modifications that accumulate in long-lived proteins, isomerization at aspartate residues has been shown to be extensive throughout the crystallins. In this study, we localize the accumulation of L-isoaspartate within water-soluble protein extracts primarily to crystallin peptides in high molecular weight aggregates, and show with mass spectrometry that these peptides are from a variety of crystallins. To investigate the consequences of aspartate isomerization, we investigated two water soluble-derived αA crystallin peptides, 52-LFRTVLDSGISEVR-65 and 89-VQDDFVEIH-98, with the L-isoaspartate modification introduced at Asp58 and Asp91, respectively. Importantly, while both peptides modestly increase protein precipitation, the native 52-LFRTVLDSGISEVR-65 peptide shows higher aggregation propensity. In contrast, the introduction of L-isoaspartate within a previously identified anti-chaperone peptide from water-insoluble aggregates, αA crystallin 66-SDRDKFVIFL(isoAsp)VKHF-80, results in enhanced amyloid formation in vitro. The modification of this peptide also increases aggregation of the lens chaperone αB crystallin. These findings may represent multiple pathways within the lens wherein the isomerization of aspartate residues in crystallin peptides differentially results in peptides associating with water-soluble or water-insoluble aggregates. Here the eye lens serves as a model for the cleavage and modification of long-lived proteins within other aging tissues.

INTRODUCTION

The synthesis of the main structural proteins of the mammalian lens, the crystallins, begins during embryonic lens development within the primary fiber cells, which eventually comprise the lens core, termed the lens nucleus (1). Crystallin synthesis is followed by the loss of the cellular nucleus and other organelles via autophagy, mitophagy, and nucleophagy in order to minimize light scattering (2). The resulting fiber cells are largely devoid of protein turnover machinery yet contain protein concentrations upwards of 450 mg/ml in the human lens. These high protein concentrations provide lens transparency via short-range order interactions that minimize errors in refraction by destructive interference (3, 4). These proteins, many of which have been synthesized by the time of birth, are not protected by the same turnover mechanisms present in normal somatic cells, and are thus susceptible to spontaneous, age-related covalent modifications and aggregation.

Post-translational modifications that accumulate over time have been identified within all the major crystallin families – α, β, and γ. These alterations are largely age-dependent spontaneous reactions leading to deamidation, isomerization, racemization, oxidation, and glycation (5, 6). Among these, isomerization and deamidation has been extensively characterized in the aged lens (7-11). The primary mechanism of asparagine deamidation and aspartate isomerization involves the formation of an intermediate L-succinimide ring, which can be hydrolyzed at either of its two carbonyls resulting in either L-aspartate (L-Asp) or L-isoaspartate (L-isoAsp; Fig. 1). Additionally, the L-succinimide intermediate can racemize to D-succinimide and yield D-aspartate (D-Asp) or D-isoaspartate (D-isoAsp). The major products of these reactions are L-isoaspartyl residues (12). Deamidation can affect the structural integrity of proteins through the introduction of a negative charge, but L-isoAsp is particularly harmful at both asparagine and aspartate sites within proteins due to the addition of a backbone methylene group that effectively “kinks” the polypeptide chain (13). Aspartate isomerization products can be difficult to distinguish from normal aspartyl residues with canonical mass spectrometry

techniques and their identification remains challenging, often requiring specific fragmentation methods such as electron transfer dissociation, electron capture dissociation, or complex labeling strategies (14, 15). Because of these difficulties, our understanding of L-isoAsp modifications within the context of the human lens is still incomplete.

There is a repair pathway within most cells, including the lens fiber cells, that results in the conversion of L-isoaspartyl to L-aspartyl residues by the L-isoaspartate (D-aspartate) O-methyltransferase (PCMT1) (16, 17). The major PCMT1 activity initiates the recognition of L-isoaspartyl residues forming methyl esters that can then spontaneously result in their conversion to normal L-aspartate residues. This enzyme can also recognize D-aspartyl residues with a much lower affinity (at least 700 times) in reactions that can lead to D-isoaspartyl formation (18). However, it has been shown that aspartate isomers are still present within lens proteins. Various groups have shown the accumulation of the four Asp isomers at specific sites, particularly within the α crystallins of aged and cataractous lenses (19-24). The accumulation of these residues presumably reflects the rate of their formation and the rate of their removal by either repair or degradation reactions. However, the effects of these modifications on lens function remain unclear. Takata et al., 2016 linked aspartate isomerization to dissociation of crystallins from the native oligomeric form (25). Other studies have shown deamidation within crystallins in vitro to be associated with aggregation (6).

Here we investigated the localization, extent, and possible consequences of L-isoAsp accumulation within the human lens. We took advantage of the specificity of PCMT1 and the high sensitivity of radiolabeling techniques to demonstrate that L-isoAsp residues accumulate within the aged lens to high levels, with the greatest accumulation in the urea-solubilized water-insoluble (WI) nuclear extract proteins. Size exclusion separation of water-soluble (WS) lens extracts followed by radiolabeling and SDS-PAGE separation of polypeptides containing L-isoaspartyl reveals that the highest levels of labeled residues are localized to aggregated low molecular weight (LMW) protein fragments that migrate primarily below 14 kDa on SDS-PAGE.

Mass spectrometry of these LMW species shows these fragments come from a number of different lens proteins. Finally, to investigate the potential consequences of this modification within lens protein fragments, we probed the effects of L-isoAsp residues in αA crystallin-derived peptides. Our assays within three peptides revealed that the introduction of the L-isoAsp residue can increase or decrease the aggregation tendencies of the given peptide, which may dictate whether peptides in the lens eventually become primarily WI- or WS-associated.

RESULTS

L-isoaspartyl residues accumulate to high levels within water-soluble (WS) and urea solubilized water-insoluble (WI) extracts of aged lenses despite endogenous PCMT1 activity

Recombinantly purified L-isoaspartyl/D-aspartyl protein methyltransferase PCMT1 was used as an analytical probe with [3H]AdoMet to specifically label L-isoAsp residues and to quantitate the extent of isomerization by detecting [3H]methanol in a volatility assay after base hydrolysis of the [3H]methyl esters formed in the incubation. L-isoAsp content in the WS and urea solubilized WI extracts of whole human lenses from ages spanning 26-76 years old was observed to increase with age (Fig. 2A). Levels for the WI extracts were ~five-fold higher than the WS extracts in the younger samples and some two-fold higher in the older samples. Control experiments using the peptide substrate KASA-isoD-LAKY showed that urea content of the WI samples did not affect the PCMT1 activity (Fig. S1).

The nuclear regions of the lens develop prior to birth, while the cortical regions continue to grow throughout an organism's life span. Thus the interior nuclear sections of the spherical lens are more aged than the exterior cortical regions. To see how L-isoAsp levels compared between these regions, separate lenses from 47-76 year olds were further dissected into nuclear and cortical regions. The 75-year-old WS dissected lens sample averages 6960 pmols L-isoAsp/mg protein extract, while in the 76-year-old WS whole lens sample has a mean value of 6590 pmols L-isoAsp/mg protein extract, showing that the averages of the aging WS cortex

and nucleus correspond well with the values observed in the WS whole lens extract. The urea solubilized WI extracts again demonstrated the greatest level of the L-isoAsp modification, and in both the WS and WI extracts the lens nucleus contained higher amounts than the cortex (Fig. 2B). However, the averages of the WI nucleus and cortex protein extracts were significantly higher than the WI whole lens protein extracts, with the 75-year-old WI nucleus and cortex averaging 17,200 pmols L-isoAsp/mg protein extract, and the 76-year-old WI whole lens averaging 12,300 pmols L-isoAsp/mg protein extract. These variations may be the result of differences in efficiency of urea-solubilization of the whole lenses versus the dissected lens regions.

The high levels of the L-isoAsp modification - up to 17,200 pmol L-isoaspartate/mg protein in the urea-solubilized water-insoluble nuclear lens extract - corresponds to an average of approximately 0.4 isoaspartyl residues per polypeptide chain, indicating that almost half of all proteins in this fraction may be modified. We did not quantify total L-isoAsp within water-insoluble urea-insoluble (UI) fractions due to the inability to fully solubilize this protein fraction in a buffer compatible with the PCMT1 enzyme. This UI fraction has been shown to increase to up to 30% of the fiber cell mass after 50 years of age (26, 27). Thus, within our assays we may only be quantifying a small portion of the total L-isoAsp within the lens. However, our calculation of 0.4 isoaspartyl residues per polypeptide chain can be taken as a minimal estimate of the total levels of aspartate isomerization within lens proteins.

The L-isoAsp repair enzyme PCMT1 is present in the aging lens and should mediate the accumulation of L-isoaspartyl residues (16). In Fig. 2C, endogenous PCMT1 activity levels between the human 26-year-old and 44-year-old whole lens extracts were found to decrease from approximately 18.5 pmol/min/mg to 12.5 pmol/min/mg, then remained relatively constant throughout the 61-, 66-, and 76- year-old tissues, reaching a minimum activity of 10.4 pmol/min/mg in the 76-year-old extract. In the dissected nuclear and cortical lens extracts, enzyme activity levels were an average of 7 pmol/min/mg higher within the cortical extracts than

the nuclear extracts of the same lens, suggesting that endogenous PCMT1 activity within the nucleus may be affected by the age of the enzyme. However, the levels of PCMT1 activity observed within the 76-year-old whole lens sample and the 51 and 75 year-old nuclear samples are still some 50% of the highest activities in younger lenses. These data show that while the activity of the L-isoaspartyl repair enzyme may decrease slightly with age in the human lens, it is still significant within the oldest samples. The lack of PCMT repair activity is unlikely to be due to limiting amounts of the AdoMet cofactor, as studies have reported that in the aged nucleus the minimum level of AdoMet observed is 10 μM (28). Thus in vivo, the repair enzyme may be inhibited by another small molecule or occluded from repairing the L-isoAsp sites, resulting in the accumulation of L-isomerized aspartyl residues seen in Figures 2A and 2B.

To investigate the possibility that endogenous PCMT1 may be obstructed from repairing L-isoAsp sites, exogenous PCMT1 activity was tested against soluble and aggregated αA66-80, isoAsp76 peptide (SDRDKFVIFL(isoAsp)VKHF). The purified enzyme was able to methylate over 100% of the analyzed soluble peptide, which may indicate that the other aspartate residues within the peptide are partially isomerized (Fig. S2). In contrast, only 2% of analyzed peptide aggregates were observed to be methylated. These results support the hypothesis that the repair enzyme cannot access L-isoAsp sites within heavily aggregated species for repair.

Age-related L-isoaspartyl sites are largely localized to low molecular weight polypeptides

PCMT1 was again used as an analytical probe with [3H]AdoMet to specifically label and visualize L-isoAsp sites within lens polypeptides by radiolabeling and subsequent SDS-PAGE separation. PCMT1-radiolabeled whole lens extracts were separated via SDS-PAGE as shown in Figures 3A and B. The gels were exposed to a film to localize radioactive signals corresponding to polypeptides containing L-isoaspartyl residues. In the fluorographs of whole lens extracts from the 26-year-old sample, radioactivity is found predominately in the positions corresponding to the molecular weights of the α- (~19-20 kDa), β- (~22-28 kDa), and γ- (~20-21 kDa) crystallins (lower panels, Figures 3A & B). However, in the older lens samples, there is

reduced radioactivity in the crystallin polypeptides but now significant amounts of radioactivity are present in the extreme high (>200 kDa) and low molecular weight (<14 kDa) regions of the gel. These results suggest that isomerization of crystallins may result in their aggregation to species resistant to SDS denaturation or in their proteolysis.

The localization of L-isoAsp-containing polypeptides to low and high molecular weight species seen in the older whole lens extracts in Fig. 3A was also observed in lens samples dissected into nuclear and cortical regions (Fig. 3B). Here, we only observed distinct bands corresponding to intact crystallin molecular weights in the cortical extracts. These results further support the idea that soluble crystallins accumulating isomerized L-aspartate residues may be become insoluble or degraded over time in the nucleus of the aged lens.

L-isoaspartyl-containing polypeptides of high and low molecular weight regions from SDS-PAGE gels are localized to the void volume fraction of size exclusion chromatography

WS whole and nuclear lens extracts were fractionated by size on a native gel filtration column that has a void volume corresponding to proteins of greater than 670 kDa. Four distinct UV absorbing peaks were detected which match those identified by Harding (29): a high molecular weight (HMW) peak in the void volume, an α crystallin peak, βH/βL crystallin peaks, and a γ crystallin peak (Fig. 4A & 4B). We found that the UV absorbance of the HMW peak in both the whole lens and the nuclear extracts increases with age in respect to the soluble crystallin fractions, likely correlating to an increase in aggregated protein. When fractions of the whole lens extracts were analyzed for L-isoaspartyl levels by PCMT1-radiolabeling, little signal was found in the 26-year-old and 44-year-old samples, while a large peak of radioactivity was found in the HMW fraction of the 76-year-old extract. In the nuclear extracts, radioactivity is found in all age samples in the HMW and crystallin fractions, but does increase with age (Fig. 4B). Interestingly, no peak was seen to correspond to very low molecular weights below the γ crystallin peak, nor was damage observed in that region, as might be expected from the results

seen in Fig. 3, where large amounts of L-isoaspartate signal was seen in the LMW region of the SDS-PAGE fluorograph. This result suggests that the low molecular weight isoaspartyl-containing polypeptides seen on SDS-PAGE are aggregated in the absence of SDS.

To investigate the possibility suggested above, the gel filtration fractions were concentrated by precipitation with trichloroacetic acid, and resuspended pellets were radiolabeled by PCMT1, separated by SDS-PAGE, and the L-isoAsp signal detected by fluorography in whole lens samples (Fig. 5) and in nuclear extracts (Fig. S3). Trichloroacetic acid precipitation would be expected for peptides of 2 kDa or higher (30). Strikingly, the LMW species observed in Fig. 3 reappear via fluorography within the HMW gel filtration fraction in both whole lens extracts and nuclear extracts and again show large amounts of the L-isoAsp signal. These LMW species are below the intact weights of the human crystallins suggesting they are cleaved fragments that must be aggregated or otherwise interacting with larger protein species.

The aggregated LMW species from the HMW gel filtration fraction represent a variety of lens protein fragments, which accumulate with age and contain isomerization-prone sites

The HMW fractions from the 44-year-old and 76-year-old WS whole lens samples were resuspended in urea, and LMW species were isolated by passage through a molecular weight cut off filter as described in the “Experimental Procedures” section. These LMW eluents were analyzed by mass spectrometry (without any protease treatment) to identify specific endogenous protein fragments against a database of 10 highly abundant lens proteins: αA crystallin, αB crystallin, βA3 crystallin, βA4 crystallin, βB1 crystallin, βB2 crystallin, γC crystallin, γS crystallin, filensin, and phakinin. Peptide fragments were identified from all of these proteins, although only one peptide was identified from γC crystallin. All peptides with an ion score of 26 or higher are shown in Tables S1-S9, while the top five peptides as ranked by ion score of each protein are represented in Table 2. Of the proteins, αA crystallin, αB crystallin, filensin, and phakinin had the highest numbers of identified peptide sequences perhaps reflecting high levels

of these lens protein fragments within the WS-HMW aggregates, the sequence coverage of each of the 10 proteins is shown in Table 1. The frequency of specific residues at the N- and C-termini of all the protein fragments combined is represented in Fig. 6A. Although there is not a strong occurrence of any one residue at the termini, the most frequent residues found at the N-termini were Ser, Arg, and Gly residues, while at the C-terminus Leu, Gln, Lys, and Phe residues were most abundant (Fig. 6A). Notably, neither Asn nor Asp was prevalent at the termini, which suggests that non-enzymatic cleavages via the succinimide ring intermediate are not common within the aged WS extract (31).

The relative abundances of the peptide fragments were quantified to see which peptides increased between the 44- and 76-year-old samples. A majority of these peptides increase in abundance in the 76-year-old sample. In Fig. 6B, the sites of cleavage for αA and αB crystallin were mapped onto a diagram of the domains of these crystallins. Although there are a higher absolute number of cleavages within the core α crystallin domain, there are also a significant number within the terminal domains. Together these results suggest that there is no strong sequence specificity of the crystallin fragmentation of αA and αB crystallin in the WS aggregates, but that the α-crystallin domain of the proteins may be more accessible to cleavage due to the involvement of the terminal domains in oligomerization of the crystallins (32). There are , however, strong biases within the βA4 crystallin, βB1 crystallin, βB2 crystallin, and Filensin proteins. The majority of fragments observed for these proteins are from the C-terminus, perhaps indicating that these are derived from truncations of the proteins (Tables S4-S7).

Isomerized forms of MS-identified peptides αA52-65 and αA89-98 modestly increase protein precipitation

Of the 38 peptides with an ion score of 26 or higher identified from αA crystallin, 34 were found to contain at least one of the following aspartate residues: Asp58, 67, 76, 84, 91, 92, and 136. All of these sites, excepting 67 and 136, have been shown in previous literature to be highly isomerized, particularly Asp58 and Asp 91/92 (19, 23), which lie in the N-terminal region

and the α-crystallin domain, respectively. To investigate the possible effects of L-isoaspartate accumulation in short peptides within the water-soluble extract of the aging lens, we used synthetic peptides of two αA crystallin peptides that were identified from the WS-HMW peak by mass spectrometry (Table S1), 52-LFRTVLDSGISEVR-65 (αA52-65) and 89-VQDDFVEIH-98 (αA89-98), with the L-isoaspartate modification introduced at Asp58 (αA52-65, isoAsp58) and Asp91 (αA89-98, isoAsp91), respectively.

The α-crystallins are related to small heat shock proteins and act as molecular chaperones within the lens, suppressing aggregation of misfolded crystallins (33). In order to see the effect of these peptides on the chaperone activities of αB crystallin, the αA crystallin segments were first incubated with purified αB crystallin and alcohol dehydrogenase (ADH) under conditions that cause ADH to denature and aggregate. The ability of αB crystallin to prevent ADH aggregation was then monitored over time by light scattering at A360. In Fig. 7A, the aggregation of ADH is largely prevented by the addition of chaperone αB crystallin, as shown by the significant reduction in light scattering. Neither the addition of the αA52-65, isoAsp58 nor the αA89-98, isoAsp91 peptide appears to alter the levels of aggregation when added to ADH with αB crystallin. Thus inhibition of the activity of the αB crystallin chaperone by the isomerized peptides was not observed under these conditions.

In order to probe the peptides effects on crystallin solubility, two different amounts of αA52-65, isoAsp58 and αA89-98, isoAsp91 peptides were then incubated with the 47-year-old WS nuclear extract. After centrifugation, the pellets were solubilized and separated by SDS-PAGE (Fig. 7B). Increases in the amount of protein in the pellet were seen when both 25 and 50 μg of the peptides were present as quantified by densitometry (Fig. 7C). Thus, the addition of the isoaspartyl-containing peptides, especially αA89-98, isoAsp91, may modestly promote insolubilization of the lens proteins.

The isomerized αA52-65, isoAsp58 peptide shows decreased self-aggregation and lens protein precipitation compared to the native αA52-65 peptide

The αA52-65, αA52-65, isoAsp58, and αA89-98, isoAsp91 peptides were all soluble at physiological pH, while the native αA89-98 could only be solubilized in 3% ammonium hydroxide. To facilitate comparison between native and isomerized peptides, the αA52-65 peptides were solubilized in Tris buffer and the αA89-98 peptides were solubilized in ammonium hydroxide as described in the "Experimental Procedures" section. While none of the peptides were observed to bind thioflavin T (ThT), self-aggregation of the native and isomerized forms was monitored by light scattering at 340 nm (Fig. 8A). The only peptide to show significant increases in light scattering as a result of precipitation was the native αA52-65 peptide.

We observed in the previous experiment that the isomerized peptides were able to induce moderate increases in precipitated lens proteins. To investigate the effects of the isomerized peptides relative to the native forms, the αA52-65 and αA52-65, isoAsp58 peptides were incubated with 47-year-old WS nuclear extract. The αA89-98 peptides were not used due to the high alkaline conditions required for solubilization of the native peptide which alone may affect lens protein stability. In Fig. 8B, the increases in protein levels within the precipitated pellet with the native αA52-65 peptide can be readily visualized. Densitometry analyses of these gel lanes again show limited increases in protein with the αA52-65, isoAsp58 peptide, but more significant shifts are observed with the native αA52-65 peptide (Fig. 8C).

The previously identified WI-derived αA crystallin peptide 66-SDRDKFVIFLDVKHF-80 displays enhanced amyloid formation and precipitation of αB crystallin in vitro upon the introduction of L-isoAsp at Asp76

The αA crystallin 66-SDRDKFVIFLDVKHF-80 (αA66-80) anti-chaperone peptide fragment was previously shown to accumulate within WI lens extracts (34). Our mass spectrometry of WS-HMW peptides showed overlapping fragments in this region of αA, but the full αA66-80 was not found (Table S1). Isomerization at the Asp76 within αA crystallin has been previously demonstrated (19, 23). We were thus interested in the effects of isomerization of the αA66-80 peptide on its aggregation and anti-chaperone activity. We show in Fig. 9A that a

synthetic αA66-80 peptide containing L-isoAsp at Asp76 (αA66-80, isoAsp76) demonstrates increased rates of aggregate formation compared to the unmodified peptide as measured by ThT fluorescence. The peptides were incubated in buffer without ThT, and resultant fibers are shown in Fig. 9B. Thus, in vitro the peptides are forming ordered, fibrillar aggregates.

The αA crystallin peptides were then incubated with αB crystallin and alcohol dehydrogenase (ADH) under conditions that cause ADH to denature, which was monitored by light scattering at A360 nm. In Fig. 9C, the light scattering of ADH alone plateaus at approximately 0.5 absorbance units. When the chaperone αB crystallin is introduced, the level of light scattering is significantly reduced. However, when the αA anti-chaperone peptide is introduced, the levels are similar to ADH alone. When the L-isoAsp-modified αA segment is used, the levels rise over those of ADH alone. This is unlikely to be due to extra scattering by the aggregation of the peptide with itself, as this yields relatively low light scattering (Fig. 9D). However, when the peptide is incubated with αB crystallin alone, there is significant light scattering (Fig. 9D). We tested the peptides' abilities to aggregate the chaperone at lower concentrations, and found that as little as 5 μg of L-isoAsp modified peptide to 40 μg of αB crystallin was sufficient to cause light scattering (Fig. S4). Thus the inhibition of αB crystallin's chaperone abilities may not be entirely direct, but rather an increase in the aggregation of the chaperone itself, exacerbated by the L-isoAsp modification in the α-crystallin fragment.

DISCUSSION

In this study, we aimed to establish the extent and potential repercussions of the isomerization of asparaginyl and aspartyl residues across the human lens proteome using highly sensitive and specific radiolabeling of modified residues by PCMT1 with [3H]AdoMet. This method revealed that the L-isoAsp modification in the aged lens localizes to aggregates that elute above 670 kDa on native gel filtration but is found there in low molecular weight peptide species after SDS-PAGE. Roy and Spector (35) previously demonstrated that WS-HMW species were identical to the insoluble protein fraction in human lenses, with the only

distinction being the size of the particles still in suspension. Remarkably they also found that a major component of these HMW protein fractions appeared to be degraded polypeptides near 11 kDa (35). It has since been determined that crystallin fragments ranging in size between 2.8 and 18 kDa can comprise up to 14-27% total protein content in WS-HMW aggregates of aged and cataract lens extracts (36-38). Our mass spectrometry results reveal the presence of endogenous short peptides in the WS-HMW native gel filtration void volume fraction from ten different lens proteins (Table 2, Tables S1-S9). The mechanisms by which these fragments are produced are not fully understood. Due to the presence of fragments cleaved at specific sites and truncated crystallins during aging it is believed that some protease activity remains in the lens during aging and slowly degrades the crystallin proteins (39-42).

Previous work shows αA and αB crystallins to be extensively truncated, and some hot spots have been identified such as the between Asp129-Prol130 bond in αB (43, 44). Two αB peptides identified within the work presented here have N- or C-terminal cleavage at or near the Asp129-Prol130 bond, including RIPADVD and RIPADVDPL (Table S2). More recently, MALDI-tissue imaging was performed on lenses of different ages, identifying intact crystallin fragments in both the nuclear and cortical regions of the tissue (45). The smallest peptide identified in this study was αA crystallin 1-34 (4,265 Da), while the largest αA peptide identified in the work presented here corresponds to residues 130-147 (1,971 Da). Thus, there were no exact matches in identified peptides. However, several residues found at the C-terminus of truncated products overlapped between this study and the MALDI-imaging results, including Gln50, Asp58, Arg65, and Phe80. Arg65 appears repeatedly within the peptides presented here and in all cases is more abundant in the 76-year-old sample than the 44-year-old. The consistent identification of truncation sites between studies may highlight their susceptibility to truncation.

Interestingly, Harrington et al., found that the WS-HMW fraction of normal aged lenses contained primarily fragments of αA and αB crystallin, while cataractous lenses also had βA3 and βB1 crystallin fragments (44). In the work presented here, tables S3 and S5 list five and 16

peptides of βA3 and βB1 crystallin, respectively. All of the βB1 peptides are higher in abundance within the 76-year-old sample, including eight which occur only within this aged sample, while four of the five βA3 increase within the 76-year-old sample. This information parallels literature that has shown that over half of deamidations within the insoluble protein fraction of aged lens were found in the β-crystallins, particularly βA3 and βB1 (7). From this study we know that residues Asn108 of βB1 and Asn133 of βA3 have mass shifts of +1 Da each. While the mass shift does not reveal whether the product of the deamidation is Asp or isoAsp, both residues sites were found contained within fragments here and these peptides should be examined for the effects of isomerization. Therefore, the βA3 and βB1 peptides identified here which increase with age may be significant formations for loss of transparency in both aged and cataractous lenses.

Previous studies have investigated the possible combinatorial effects of truncations and proteolysis on the properties of crystallins, including βB1 and γS. In the case of βB1, the combination of deamidation and truncation had the most severe effects on protein stability, while truncation alone had little effect, however, deamidation has not been examined in short βB1 peptides (46). Similarly, isolated N- and C-terminal domains of human γS crystallin showed high heat and pH stabilities (47). Thus truncated crystallin monomers may be relatively stable. However, it is known that truncation of the β crystallins affects assembly formation, which may in turn affect lens transparency (48). A number of other studies have investigated the effects of shorter fragments produced by truncations on protein aggregation in vitro, and found that peptides derived from both αA- and βL- crystallins can act as anti-chaperones, interfering with the action of intact αA and αB crystallins against protein misfolding, or as mini-chaperones promoting stability of other proteins (34, 49, 50). A majority of these studies have focused on the unmodified sequences of these crystallin fragments.

Most of the peptides identified within our WS-HMW native gel filtration void volume fraction contained Asp residues that are known to isomerize in the aging lens (23, 51; Tables 1

and 2; Tables S1-S9). Our studies of two peptides found in this fraction showed that while the L-isoAsp peptides could modestly promote protein precipitation, the native peptides had more severe solubility and aggregation properties. Thus, it appears the L-isoAsp residue within these two WS HMW-derived peptides actually decreases their propensities for self-aggregation and protein precipitation. In contrast with these results, we found that for a WI-derived anti-chaperone peptide, the L-isoAsp form showed increased levels of self-aggregation over the L-Asp peptide, as well as increased precipitation of the lens chaperone αB crystallin. It is important to note that while we demonstrated the aggregates of the peptide alone were amyloid in nature, amyloid structures have not definitively been demonstrated within the human lens, but these may be representative of other forms of aggregation in vivo (52). Thus, the effects of the L-isoAsp modification within cleaved peptides can be variable.

Our current work does not fully resolve the fundamental question of how this modification affects cataract formation. Truscott and Friedrich emphasized the difficulties of establishing causative relationships between specific PTMs and the opacification of the lens (11). It is conceivable that modification of an intact enzyme by isoaspartyl formation may affect its activity as reviewed by Reissner and Asward (53). On the other hand, it could be that for many enzymes or other structural lens proteins, where the modification is not near the active site, isomerization may be benign. However, it could well be that the isomerization of the peptides may lead to aggregation, or such peptides may serve as inhibitors of a particular enzyme or a seed for aggregation of other proteins.

In the molecular heterogeneity of the lens, these peptides will be interacting and aggregating with other species in complex manners. In these cases it is too early to predict the effects of the isomerization based solely on the sequence of the damaged peptide. However, in the experiments conducted here, we see a correlation between propensity for self-aggregation and the precipitation of other proteins (Figures 8 & 9). Thus, if a peptide is predicted to have

higher self-aggregative properties, then it may be more likely to aid in the precipitation of other proteins.

Here we have only presented three cases of isomerization within lens peptides, therefore, it is difficult to definitively extend these results to other peptide sequences. However, some theories concerning the propensity to self-aggregate can be put forth combining the peptide work presented here with previous literature. Assuming amyloid as a model of self-aggregation, the interdigitation of amino acid side chains is crucial for the formation of the steric zipper, and is highly sequence-dependent (54). Thus, the isomerization of an aspartate or asparagine in the core of a steric zipper is likely to disrupt interdigitation due to the introduction of the extra carbon-carbon bond in the main chain. Additionally, if the aspartate residue is important for forming salt bridges or similar interactions, the isomerization will likely distance the aspartyl side chain and break these bonds. The disruption of a salt bridge by Asp109 isomerization in □B crystallin has recently been probed by molecular dynamics and suggests that the loss of this interaction destabilizes the protein and leads to insolubility (55). These points may help explain decreased aggregation of certain peptides in vitro, which we have shown to form amyloid. If the aspartate is removed from steric zipper forming regions, it may have no effect, or provide needed flexibility to form a second steric zipper interface, enhancing aggregation.

In light of these results and in the context of the aging lens, we suggest that L-isoAsp containing proteins and peptides take a number of different pathways towards becoming associated with WS and WI aggregates. We believe it is unlikely that the L-isoAsp residues are specifically targeted by proteases, as in our own data set N- or C-terminal aspartate residues were rare, nor have aspartate residues been identified as a significant site of truncation by other literature studying the LMW lens fragments (44). Instead, to explain the high concentration of the L-isoAsp modification found in the WS aggregated lens protein fragments, we note that the native peptides themselves were highly prone to aggregation and insolubility, while the

isomerized αA52-65 and αA89-98 forms aggregated less. The diminished capacity for these L-isoAsp peptides to cause aggregation may represent a tendency for certain isomerized peptides to become associated with WS protein aggregates, while their native counterparts associate with WI protein aggregates. Conversely, our experiments with a WI-derived anti-chaperone peptide demonstrated more severe protein precipitation and anti-chaperone activity in the isomerized form, and this peptide could not be found in our mass spectrometric analyses of the WS-HMW aggregates (34). Thus it may be that the peptides in which the L-isoAsp modification produces more aggregation-prone properties are more likely to be found in the WI aggregates, while the peptides in which the L-isoAsp modification lessens these characteristics are more likely to be associated with WS aggregates.

Our results do not rule out that intact crystallins could become associated with HMW aggregates and isomerize while aggregated. These modified, aggregated crystallins could then become cleaved while in the aggregate form possibly in attempts by lens proteases to clear the light-scattering aggregates, and the modified peptides would then only be detected as distinct fragments by denaturing SDS-PAGE. However, the formation of the L-isoAsp residue may be less favorable in an aggregate due to the fact that L-isoAsp residues are less readily formed within rigid structures, such as that of a large aggregate. In these structures, the peptide bond nitrogen is not often posed to attack the side chain carbonyl group of the aspartate or asparaginyl residue (Fig. 1, Ref. 56). The flexibility of short peptides naturally allows the backbone to sample a number of conformations in which the backbone nitrogen approaches the L-Asp side chain close enough for nucleophilic attack.

Irrespective of the path that gives rise to L-isoAsp-containing peptides in HMW aggregates, our results show that L-isoAsp may either increase or decrease the aggregation propensity of peptides depending on the sequence of the fragment, and both are likely to contribute to WI and WS aggregation. Thus, the initial production of short crystallin fragments may represent a significant risk for the stability of other intact crystallins, and regulation of

protease activity within the aging lens may help attenuate the onset of cataract. The lens tissue also acts as an ideal model of aging due to the lack of cellular and protein turnover, and the results presented here may be representative of a number of age-related diseases, such as the polypeptide aggregates of amyloid-β and tau in Alzheimer's disease, in which protein modifications at different sites have variable effects on protein aggregation.

EXPERIMENTAL PROCEDURES

Water-soluble (WS) and water-insoluble (WI) lens extract preparation

We studied 10 pairs of unfixed eye lenses, collected within three days post-mortem from three eye banks (San Diego Eye Bank, San Diego, CA; Sightlife, Seattle, WA; and OneLegacy, Los Angeles, CA). None of the lenses had known ocular diseases. Eyes were a kind gift from Dr. Joseph Demer and were obtained in conformity with legal requirements. Lenses were thawed on ice. Whole lenses were lysed in 300-500 µL ice cold 50 mM Tris-HCl, pH 7.9, 150 mM NaCl in a glass tube with a Teflon pestle rotating at 300 rpm for two periods of 30 sec each. In some cases lens samples first had the nucleus removed using a 6 mm trephine. The remaining peripheral material was pooled for the preparation of the cortical extract. Both nuclear and cortical regions were then lysed with the same procedure as whole lens extract. Lysates were spun at 20,000g for 20 min at 4 °C. The supernatant was removed as the water-soluble extract and stored at -80 °C.

Pellets were resuspended in 300 µL of 6 M urea in 50 mM Tris-HCl, pH 7.9, 150 mM NaCl and homogenized by hand in an Eppendorf 1.6 mL microcentrifuge tube with 80 strokes of a form fitting plastic pestle and rotated for 30 min at 4 °C (Kimble-Chase Kontes pellet pestle 7495150000). After centrifugation at 15,000g for 15 min at 4 °C, the supernatant fraction was set aside and the pellet was re-homogenized with 50 strokes, then spun at 15,000g for 15 min at 4 °C. The pellet was then re-homogenized and re-centrifuged one additional time. The three supernatants were then combined as the WI extract and stored at -80 °C. Protein

concentrations were quantified by a Lowry assay after protein precipitation with 10% trichloroacetic acid (57).

Determination of L-isoaspartate levels by the PCMT1 methanol vapor diffusion assay

PCMT1 was used as an analytical reagent to quantify L-isoAsp levels in the lens extract proteins. In a final volume of 100 µL, 12.5 to 25 µg of lens extract protein was incubated for 2 h at 37 °C with 5 µg PCMT1 (purified as a His-tagged enzyme from Escherichia coli (E. coli) containing the expression plasmid #34852 available from Addgene.com as described by Patananan et al. (58) with a specific activity at 37 °C of 5,300 pmol of methyl esters formed on KASA(isoD)LAKY/min/mg of enzyme). Final concentrations in the reactions included 135 mM Bis-Tris-HCl, pH 6.4, and 10 µM S-adenosyl-l-[methyl3H]methionine ([3H]AdoMet) (prepared by a 1600-fold isotopic dilution of a stock of 72 Ci/mmol [3H]AdoMet (PerkinElmer Life Sciences, NET155H00) with nonisotopically labeled AdoMet (p-toluenesulfonate salt; Sigma-Aldrich A2408)). The reaction was stopped by adding 10 µL of 2 M sodium hydroxide, and 100 µL of the 110 µL mixture was transferred to a 9 by 2.5 cm piece of folded thick filter paper (Bio-Rad; catalog number 1650962) wedged in the neck of a 20-mL scintillation vial above 5 mL scintillation reagent (Safety Solve, Research Products International, catalog number 121000), tightly capped, and incubated at room temperature. After 2 h, the folded filter papers were removed, the caps replaced, and the vials were counted thrice for 5 minutes each in a Beckman LS6500 scintillation counter. Background radioactivity in a reaction containing no substrate was determined by incubating the recombinant human PCMT1, 135 mM Bis-Tris-HCl buffer, and 10 µM [3H] AdoMet as described above and was subtracted from the value obtained in experimental samples. Samples were analyzed in duplicate or triplicate as indicated in the figure legends.

Buffered-urea control reactions were carried out by resuspending the KASA(isoD)LAKY peptide in 50 mM Tris-HCl, pH 7.9, 150 mM NaCl with or without 6 M urea. Buffer alone, 10

pmol KASA(isoD)LAKY, or 100 pmol KASA(isoD)LAKY in volumes of 25 µL were PCMT1 and [3H]AdoMet in a final volume of 100 µL as described above.

L-isoAsp quantification in soluble versus aggregated 66-SDRDKFVIFLDVKHF-80 peptide was performed in the same reaction conditions as described above. Aggregated peptide was prepared by dissolving 66-SDRDKFVIFLDVKHF-80 at 1 mg/mL in 50 mM Tris-HCl, pH 7.6, 150 mM NaCl, (TBS) and shaking continuously for 5 days. Aggregate solutions were stored at room temperature until analysis. Soluble peptide solutions were prepared by dissolving 66-SDRDKFVIFLDVKHF-80 peptide in TBS with 10 % DMSO and filtering through Corning® Costar® Spin-X® centrifuge tube filters immediately prior to analysis (Millipore Sigma, catalog # CLS8161).

Determination of endogenous lens protein L-isoaspartyl methyltransferase (PCMT1) activity by the methanol vapor diffusion assay

In a final volume of 100 µL, 10 µg of lens extract protein was incubated for 2 h at 37 °C with final concentrations of 100 µM KASA(isoD)LAKY peptide, 125 mM Bis-Tris-HCl, pH 6.4, and 10 µM [3H] AdoMet, as prepared above. The reaction was stopped by adding 10 µL of 2 M sodium hydroxide, and 100 µL of the 110 µL mixture assayed for volatile radioactivity, as described in the section above. Background radioactivity was determined in a control lacking the lens extract and was subtracted from the value obtained in samples containing the lens extracts. Samples were analyzed in duplicate or triplicate as indicated in the figure legends.

L-isoaspartate analysis in lens polypeptides by SDS-PAGE fluorography

Analyses were performed using the approach described by Patananan et al. (56). Briefly, 25 µg extracts were analyzed in a 30 µL reaction volume with final concentrations of 74 mM Bis-Tris-HCl, pH 6.4, 6 µg recombinant human PCMT1, 0.3 µM S-adenosyl-L-[methyl-3H] methionine (PerkinElmer Life Sciences; 75-85 Ci/mml, 0.55 mCi/ml in 10 mM H2SO4:ethanol (9:1, v/v)), and incubated for 2 h at 37 °C. The reaction was stopped by adding 5 µL SDS-PAGE loading buffer (250 mM Tris-HCl, pH 6.8, 10% (w/v) SDS, 50% (v/v) glycerol, 5% (v/v) β-

mercaptoethanol, and 0.05% (w/v) bromophenol blue). Samples were heated at 100 °C for 3 min and separated on a 12% SDS-polyacrylamide gel prepared in Bis-tris-HCl, pH 6.4 and run at 140 V for 1 h. Gels were stained with Coomassie (0.1% (w/v) Brilliant Blue R-250, 10% (v/v) glacial acetic acid, and 50% (v/v) methanol) for 1 h and destained with 10% (v/v) acetic acid and 15% (v/v) methanol. For fluorography, gels were subsequently incubated with EN3HANCE (PerkinElmer Life Sciences, catalog number 6NE9701) for 1 h, incubated in water for 30 min, and dried before the gels were exposed to film (Denville Scientific, 8 × 10-inch Hyblot CI) for 2-3 days at −80 °C.

Size exclusion chromatography and fluorography of lens extract protein fractions

Size exclusion chromatography was performed on an ÄKTA prime system. A Superose 6, 10/300 GL gel filtration column (GE Healthcare, 17-5172-01, column length 30 cm, column inner diameter 10 mm, 13 µm average particle size) was equilibrated with 50 mM Tris-HCl, pH 7.9, 150 mM NaCl. Approximately 1 mg of lens extract protein was loaded for each run. One mL fractions were collected at a flow rate of 0.4 mL/min at room temperature.

Twenty-five µL were removed from fractions for subsequent L-isoaspartate quantification by the PCMT1 methanol vapor diffusion assay as described above. The remaining portion of the fraction was precipitated with 10% trichloroacetic acid overnight at 4 °C, pelleted at 20,800g for 10 min at 4 °C, and the supernatant was discarded. Pellets were resuspended in 100 µL of 100 mM Tris-HCl, pH 7.9, 150 mM NaCl, 6 M urea, 0.1% sodium dodecyl sulfate. Thirty µL of the resuspended pellet was radiolabeled in a reaction volume of 60 µL with a final concentration of 80 mM Bis-Tris-HCl, pH 6.4, 6 µg recombinant human PCMT1, 0.3 µM S-adenosyl-L-[methyl-3H] methionine, and incubated for 2 h at 37 °C as described in the section above. The reaction was stopped with 15 µL of SDS-PAGE loading buffer. Samples were heated at 100 °C for 3 min and separated on a 4-20%, 10 well ExpressPlus PAGE gel (Genscript, catalog # M42010) at

140 V for 1 h. Staining, enhancing, and fluorography proceeded as described in the section above.

Mass spectrometric identification of low molecular weight peptides in WS HMW gel filtration fraction

Lens samples were separated by size exclusion chromatography. The HMW fraction was concentrated in a Savant SpeedVac concentrator, and resuspended in 200 uL of 50 mM Tris-HCl, pH 7.9, 150 mM NaCl, 6 M urea. Diothiothrietol (DTT) was added to a final concentration of 5 mM and the solution was incubated at 56 °C for 30 min. Iodoacetamide was then added to a final concentration of 14 mM, and incubated at room temperature for 30 min. An additional amount of DTT was added at a final concentration of 5 mM and the sample was incubated for 15 min at room temperature in the dark. Sample was then applied to a Microcon YM-30 30,000 MW cut off filter (Millipore, Burlington, MA) and filtrate was collected for analysis. Equal amount of the two samples were injected into the mass spectrometer. LC-MS/MS analysis was performed on an EASY-nLC 1000 coupled to a Q Exactive mass spectrometer with nanoESI source (Thermo Fisher Scientific). The peptides were separated on a 75 um diameter × 25 cm C18 reversed phase column (Thermo Fisher Scientific) with a gradient from 5% solvent B (0.1% formic acid in acetonitrile), 95% solvent A (0.1% formic acid in water) to 40% solvent B for 30 min at a constant flow of 300 nl/min, followed by an increase to 80% B during 30-50 min. The mass spectrometer was operated in data-dependent acquisition mode with a top 10 MS/MS method. Orbitrap resolving power was set to 70,000 at m/z 200 for MS1 and 17,500 at m/z 200 for MS2. Peptides in raw data were searched against a database of 10 highly abundant human lens proteins using Proteome Discoverer, version 2.2 (Thermo Fisher Scientific) for identification and label-free precursor ion quantification. Identification was performed with Mascot search engine with no enzyme specificity. The mass tolerances were set to 10 ppm and 0.02 Da for precursor and fragment ions. Methionine oxidation and cysteine carbamidomethylation were considered as dynamic modifications. Fixed value PSM validator

was used with maximum delta Cn of 0.05. For label free quantitation, feature detection and retention time alignment were performed. Quantification values were based on intensities at apex of chromatographic peak. Peptide abundance was calculated as a sum of abundances of individual peptide-spectral matches that pass a quality threshold, and protein abundance was calculated as sum of peptide abundances. Because there were major differences observed in the total abundances of searched peptides between the two samples, no normalization was performed. Instead, we compared total ion counts of the two runs and they indicated close to equal peptide loading amounts. The raw abundances were scaled such that the average between the two samples was 100 for better visualization purposes such as comparing the two samples and noticing peptides that exist only in 1 sample.

Purification of recombinant αB crystallin

The pET20-αB crystallin plasmid in the E. coli BL21 strain was a kind gift from Dr. Wayne Hubbell at the UCLA Stein Eye Institute. Cells were grown at 37 °C to an OD600 of 0.6. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mM and cell growth continued for 3 h at 37 °C. Cells were harvested at 5,000 g for 15 min at 4 °C. Cell pellets were resuspended in 20 mM Tris-HCl, pH 8.5, 10% glycerol, 1 M NaCl with 1 mM phenylmethylsulfonyl fluoride (PMSF), 5 mM β-mercaptoethanol, 25 U/mL benzonase, and a Pierce protease inhibitor tablet, EDTA-free (ThermoFisher Scientific, A32965). Cells were lysed on an Emulsiflex with 3 passages at 15,000 psi. The lysate was spun at 9,700 g for 50 min at 4 °C. Nucleic acids were precipitated from the supernatant by the addition of 0.1% final polyethyleneimine and incubation at room temperature for 15 min. The mixture was spin at 13,000 rpm for 50 min. The supernatant was removed and subsequently loaded onto a 5 mL GE Healthcare HisTrap HP (catalog # 17-5248-01) equilibrated with wash buffer (20 mM Tris-HCl, pH 7.5; 100 mM NaCl; 5% glycerol; 20 mM imidazole). Proteins were eluted with an isocratic gradient from 0 – 100% elution buffer (20 mM Tris-HCl, pH 8.5; 100 mM NaCl; 5% glycerol; 500 mM imidazole) over 60 min at 1 mL/min. All fractions containing αB crystallin were pooled and

applied to a 5 mL GE Healthcare HiTrap Q HP anion exchange column (GE29-0513-25) equilibrated with wash buffer (20 mM Tris-HCl, pH 7.5; 5% glycerol). Proteins were eluted with an isocratic gradient from 0 – 100% elution buffer (20 mM Tris-HCl, pH 8.5; 5% glycerol; 1 M NaCl) over 70 mL at 1 mL/min. All fractions containing the polypeptide corresponding to αB crystallin after SDS-PAGE were pooled and loaded onto a HiPrep 16/60 Sephacryl S-200 HR gel filtration column (GE Healthcare 17116601). Proteins were separated at 0.4 mL/min over 200 mL wash buffer (20 mM Tris-HCl, pH 7.5; 100 mM NaCl). Fractions containing aB crystallin were pooled, glycerol was added to a final concentration of 5%, and the protein was concentrated in an Amicon Ultra-15 centrifugal filter unit (Millpore Sigma, UFC901008).

Thioflavin T binding and light scattering conditions for peptide aggregation assays

Synthetic peptides of 66-SDRDKFVIFLDVKHF-80 with either an L-Asp or L-isoAsp residue at the 76 position were obtained from Genscript and dissolved in water at a concentration of 100 μM. Thioflavin T (Sigma, T3516; ThT) was dissolved at a concentration of 100 μM in 100 mM Tris-HCl, pH 7.6, 300 mM NaCl (TBS, Sigma, 94158). For a final volume of 200 μL in a 96 well plate (Fisherbrand, flatbottom, clear, nonsterile, 12565501) 100 μL of peptide stock was mixed with 100 μL of the 100 μM ThT stock. Plates were continuously shaken at 60 rpm and fluorescence readings were taken every 15 min (excitation of 450 nm and 482 nm emission) in a Varioskan plate reader.

Synthetic αA crystallin peptides, 52-LFRTVLDSGISEVR-68 and 89-VQDDFVEIH-98, were obtained from Genscript. For assessment of self-aggregation by light scattering, both native and isomerized 52-LFRTVLDSGISEVR-68 peptides were dissolved at a concentration of 3 mg/ml in 50 mM Tris-HCl, pH 7.5, 1% DMSO. Because of the highly insoluble nature of the native 89-VQDDFVEIH-98 peptide, both of the native and isomerized 89-VQDDFVEIH-98 peptides were first dissolved at 2.5 mg/ml in 3% ammonium hydroxide in water (final concentration of 1.58 M from a 28% stock of Fisher certified ACS Plus reagent) and diluted to 2 mg/ml in a final concentration of 50 mM Tris-HCl, pH 7.5. The final pH of this solution as tested

by pH strip (BDH VWR analytical; catalog #BDH35309.606, pH range 0-14), was found to be between pH 9 and 10. Two wells of 100 µL volume were assayed for each peptide solution in a Varioskan plate reader at 340 nm. Plates were continuously shaken at 600 rpm and readings were taken every 15 min.

Chaperone inhibition assays

Saccharomyces cerevisiae alcohol dehydrogenase (ADH; Sigma A7011), 125-150 µg; αB crystallin, 40-50 µg; and peptides, 40-50 µg, were mixed in a final volume of 300 µL of reaction buffer (50 mM sodium phosphate, pH 7.0, 100 mM NaCl, and 10 mM phenanthroline). Denaturation was monitored by absorbance at 360 nm every 15 min with continuous shaking at 60 rpm in a Varioskan plate reader.

Transmission electron microscopy

Formvar/Carbon grids (Ted Pella, catalog 01754-F) were prepared with 3 µL αA66-80, isoAsp76 fiber stock solution for 3 min, washed with water, then 2 µL Ted Pella uranyl acetate alternative (Ted Pella, catalog 19485) was applied for 2 min, rinsed with water and air dried. Images were collected on an FEI T12 instrument.

Protein precipitation gel assays of peptide and lens extract mixtures

The αA52-65, isoAsp58 and αA89-98, isoAsp91 peptides were incubated with 500 µg of 47 year old lens WS protein extract for 14 h at 37 °C in a final volume of 120 µl 50 mM Tris-HCl, pH 7.9. Reactions were then spun down for 10 min at 960 g. The supernatant was removed and the pellet was resuspended in 40 µl of 1x SDS-PAGE loading dye and separated by SDS-PAGE on a 4-20%, 10 well ExpressPlus PAGE gel (Genscript, catalog # M42010) at 140 V for 1 h. Densitometry of lanes was performed in ImageJ (59).

ACKNOWLEDGMENTS

The authors would like to thank Linlin Ding for providing materials and advice, and Dr. Joseph Demer for the gift of lens material obtained with support from grant EY008313. This work was supported by grants to S.G.C. from the National Science Foundation (MCB-1714569), the UCLA

Academic Senate Faculty Research Program, the Life Extension Foundation, Inc., and the Elizabeth and Thomas Plott Chair in Gerontology of the UCLA Longevity Center. J.H. is supported by NIH grant R01EY023588. J.A.L. is supported by NIH grant GM103479. Funding for R.A.W. was provided by the UCLA Dissertation Year Fellowship and by the Ruth L. Kirschstein National Research Service Award GM007185. K. Liu is supported by a Canadian NSERC Postgraduate Scholarship. The authors would like to thank the Electron Imaging Center for NanoMachines (EICN) of California NanoSystems Institute (CNSI) at UCLA for the use of the electron microscopes.

REFERENCES

1. Forrester, John V., Dick, Andrew D., McMenamin, Paul G., Roberts, Fiona. (2016) The Eye: Basic Sciences in Practice, Third Edition. New York: Elsevier, p. 102. Print.

2. Cvekl, A., Ashery-Padan, R. (2014) The cellular and molecular mechanisms of vertebrate lens development. Development. 141, 443-4447.

3. Bassnett, S., Shi, Y., Vrensen, G.F. (2011) Biological glass: structural determinants of eye lens transparency. Philos. Trans. R. Soc. Lond. B. Bio. Sci. 366, 1250-1264.

4. Delaye, M., Tardieu, A. (1983) Short-range order of crystallin proteins accounts for eye lens transparency. Nature. 302, 415-417.

5. MacCoss, M.J., McDonald, W.H., Saraf, A., Sadygov, R., Clark, J.M., Tasto, J.J., Gould, K.L., Wolters, D., Washburn, M., Weiss, A., Clark, J.I., Yates III, J.R. (2002) Shotgun identification of protein modifications from protein complexes and lens tissue. Proc. Natl. Acad. Sci. USA. 99, 7900-7905.

6. Lampi, K.J., Wilmarth, P.A., Murray, M.R., David, L.L. (2014) Lens β-crystallins: The role of deamidation and related modifications in aging and cataract. Prog. Biophys. Mol. Biol. 115, 21-31.

7. Wilmarth, P.A., Tanner, S., Dasari, S., Nagalla, S.R., Riviere, M.A., Bafna, V., Pevsner, P.A., David, L.L. (2006) Age-related changes in human crystallins determined from comparative analysis of post-translational modifications in young and aged lens: does deamidation contribute to crystallin insolubility? J. Proteome Res. 5, 2554-2566.

8. Hains, P.G., Truscott, R.J.W. (2007) Post-translational modifications in the nuclear region of young, aged, and cataract human lenses. J. Proteome Res. 6, 3935-3943.

9. Hains, P.G., Truscott, R.J.W. (2010) Age-Dependent Deamidation of Lifelong Proteins in the Human Lens. Invest. Ophthalmol. Vis. Sci. 51, 3107-3114.

10. Ray, N.J. (2015) Biophysical chemistry of the ageing eye lens. Biophys. Rev. 7, 353-368.

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. Mcfadden, P.N., Clarke, S. (1987) Conversion of isoaspartyl peptides to normal peptides: Implications for the cellular repair of damaged proteins. Proc. Natl. Acad. Sci. 84, 2595-2599.

13. 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.

14. Ni, W., Dai, S., Karger, B.L., Zhou, Z.S. (2010) Analysis of Isoaspartic Acid by Selective Proteolysis with Asp-N and Electron Transfer Dissociation Mass Spectrometry. Anal. Chem. 82, 7485-7491.

15. Liu, M., Cheetham, J., Cauchon, N. Ostovic, J., Ni, W., Ren, D., Zhou, Z.S. (2012) Protein isoaspartate methyltransferase-mediated 18O-labeling of isoaspartic acid for mass spectrometry analysis. Anal. Chem. 84, 1056-1062.

16. McFadden, P.N., Clarke, S. (1986) Protein carboxyl methyltransferase and methyl acceptor proteins in aging and cataractous tissue of the human eye lens. Mech. Ageing Dev. 34, 91-105.

17. Chondrogianni, N., Petropoulos, I., Grimm, S., Georgila, K., Catalgol, B., Friguet, B., Grune, T., Gonos, E.S. (2014) Protein damage, repair, and proteolysis. Mol. Aspects Med. 35, 1-71.

18. Lowenson, J.D., Clarke, S. (1992) Recognition of D-aspartyl residues in polypeptides by the erythrocyte L-isoaspartyl/D-aspartyl protein methyltransferase. Implications for the repair hypothesis. J. Biol. Chem. 267, 5985-5995.

19. Fujii, N., Hiroaki, S., 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.

20. Hooi, M.Y.S., Raftery, M.J., Truscott, R.J.W. (2012) Racemization of Two Proteins over Our Lifespan: Deamidation of Asparagine 76 in γS Crystallin is Greater in Cataract than in Normal Lenses across the Age Range. Invest. Ophthalmol. Vis. Sci. 53, 3554-3561.

21. Hooi, M.Y., Raftery, M.J., Truscott, R.J. (2013) Accelerated aging of Asp 58 in αA crystallin and human cataract formation. Exp. Eye. Res. 106, 34-39.

22. Zhu, X., Zhang, K., He, W., Du, Y., Hooi, M., Lu, Y. (2018) Racemization at the Asp 58 residue in αA-crystallin from the lens of high myopic cataract patients. J. Cell. Mol. Med. 22, 1118-1126.

23. Lyon, Y.A., Sabbah, G.M., 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. 171, 131-141.

24. Takata, T. Shimo-Oka, T., Kojima, M., Miki, K., Fujii, N. (2006) Differential analysis of D-beta-Asp-containing proteins found in normal and infrared irradiated rabbit lens. Biochem. Biophys. Res. Commun. 344, 263-271.

25. Takata, T., Fujii, N. (2016) Isomerization of Asp residues plays an important role in αA-crystallin dissociation. FEBS J. 283, 850-859.

26. Li, L.K., Roy, D., Spector, A. (1986) Changes in lens protein in concentric fractions from individual normal human lenses. Curr. Eye. Res. 5, 127-135.

27. Harrington, V., Srivastava, O.P., Kirk, M. (2007) Proteomic analysis of water insoluble proteins from normal and cataractous human lenses. Mol. Vis. 13, 1680-1694.

28. Truscott, R.J.W., Mizdrak, J., Friedrich, M.G., Hooi, M.Y., Lyons, B., Jamie, J.F., Davies, M.J., Wilmarth, P.A., David, L.L. (2012) Is S-methylation in the human lens a result of non-enzymatic methylation by Sadenosylmethionine? Exp. Eye. Res. 99, 48-54.

29. Harding, J.J., Crabbe, M.J.C. (1984) The lens: development, proteins, metabolism and cataract. In: Davson H (ed) The eye, vol 1B. New York, Academic Press p. 207-492.

30. Yvon, M., Chabanet, C., Pélissier, J.P. (1989) Solubility of peptides in trichloroacetic acid (TCA) solutions. Hypothesis on the precipitation mechanism. Int. J. Pept. Protein Res. 34, 166-176.

31. Voorter, C.E., de Haard-Hoekman, W.A., van den Oetelaar, P.J., Bloemendal, H., de Jong, W.W. (1988) Spontaneous peptide bond cleavage in aging alpha-crystallin through a succinimide intermediate. J. Biol. Chem. 263, 19020-19023.

32. Laganowsky, A., Benesch, J.L.P., Landau, M., Ding, L., Sawaya, M.R., Cascio, D., Huang, Q., Robinson, C.V., Horwitz, J., Eisenberg, D. (2010) Crystal structures of truncated alpha and alphaB crystallins reveal structural mechanisms of polydispersity important for eye lens function. Protein Sci. 19, 1031-1043.

33. Horwitz, J. (1992) Alpha-crystallin can function as a molecular chaperone. Proc. Natl. Acad. Sci. USA. 89, 10449-10453.

34. Santhoshkumar, P., Raju, M., Sharma, K.K. (2011) αA-crystallin peptide SDRDKFVIFLDVKHF accumulating in aging lens impairs the function of α-crystallin and induces lens protein aggregation. PLoS One. 6, e19291.

35. Roy, D., Spector, A. (1976) High molecular weight protein from human lenses. Exp. Eye Res. 22, 273-279.

36. Srivastava, O.P. (1988) Age-related increase in concentration and aggregation of degraded polypeptides in human lenses. Exp. Eye Res. 47, 525-543.

37. Santhoshkumar, P., Udupa, P., Murugesan, R., Sharma, K.K. (2008) Significance of interactions of low molecular weight crystallin fragments in lens aging and cataract formation. J. Biol. Chem. 283, 8477-8485.

38. Srivastav, O.P., Srivastava, K., Silney, C. (1996) Levels of crystallin fragments and identification of their origin in water soluble high molecular weight (HMW) proteins of human lenses. Curr. Eye. Res. 15, 511-520.

39. Groenen, P.J., Merck, K.B., de Jong, W.W., Bloemendal, H. (1994) Structure and modifications of the junior chaperone alpha-crystallin. From lens transparency to molecular pathology. Eur. J. Biochem. 225, 1-19.

40. Van Fleef, F.S.M., Willems-Thijssen, W., Hoenders, H.J. (1976) Intracellular Degradation and Deamidation of a-Crystallin Subunits. Eur. J. Biochem. 66, 477-483.

41. Sharma, K.K., Santhoshkumar, P. (2009) Lens Aging: Effects of the Crystallins. Biochim. Biophys. Acta. 1790, 1095-1108.

42. Takemoto, L.J. (1995) Identification of the in vivo truncation sites at the C-terminal region of alpha-A crystallin from aged bovine and human lens. Curr. Eye Res. 14, 837-841.

43. Srivastava, O.P., Srivastava, K. (2003) Existence of deamidated aB-crystallin fragments in normal and cataractous human lenses. Mol. Vis. 9, 110-118.

44. Harrington, V., McCall, S., Huynh, S., Srivastava, K., Srivastava, O.P. (2004) Crystallins in water soluble-high molecular weight protein fractions and water insoluble protein fractions in aging and cataractous human lenses. Mol. Vis. 10, 476-489.

45. Grey, A.C., Schey, K.L. (2009) Age-Related Changes in the Spatial Distribution of Human Lens α-Crystallin Products by MALDI Imaging Mass Spectrometry. Invest. Ophthalmol. Vis. Sci. 50, 4319-4329.

46. Lampi, K.J., Kim, Y.H., Bächinger, H.P., Boswell, B.A., Lindner, R.A., Carver, J.A., Shearer, T.R., David, L.L.m Kapfer, D.M. (2002) Decreased heat stability and increased chaperone requirement of modified human betaB1-crystallins. Mol. Vis. 8, 359-366.

47. Wenk, M., Herbst, R., Hoeger, D., Kretschmar, M., Lubsen, N.H., Jaenicke, R. (2000) Gamma S-crystallin of bovine and human eye lens: solution structure, stability and folding of the intact two-domain protein and its separate domains. Biophys. Chem. 86, 95-108.

48. Ajaz, M.S., Ma, Z., Smith, D.L., Smith, J.B. (1997) Size of human lens beta-crystallin aggregates are distinguished by N-terminal truncation of betaB1. J. Biol. Chem. 272, 11250-11255.

49. Senthilkumar, R., Chaerkady, R., Sharma, K.K. (2002) Identification and Properties of Anti-Chaperone-like Peptides Derived from Oxidized Bovine Lens βL-Crystallins. J. Biol. Chem. 277, 39136-39143.

50. Raju, M., Santhoshkumar, P., Sharma, K.K. (2012) αA-Crystallin-Derived Mini-Chaperone Modulates Stability and Function of Cataract Causing αAG98R-Crystallin. PLoS One. 7, e44077.

51. Fujii, N., Takata, T., Fujii, N., Aki, K. (2016) Isomerization of aspartyl residues in crystallins and its influence upon cataract. Biochim. Biophys. Acta. 1860, 183-191.

52. Truscott, R.J.W. (2007) Eye Lens Proteins and Cataracts. In: Uversky, V.N., Fink, A.L. (eds) Protein Misfolding, Aggregation, and Conformational Diseases. Protein Reviews, vol. 6. Springer, Boston, MA.

53. Thompson, M.J., Sievers, S.A., Karanicolas, J., Ivanova, M.I., Baker, D., Eisenberg, D. (2006) The 3D profile method for identifying fibril-forming segments of proteins. Proc. Natl. Acad. Sci USA. 103, 4074-4078.

54. Reissner, K.J., Aswad, D.W. (2003) Deamidation and isoaspartate formation in proteins: unwanted alterations or surreptitious signals? Cell. Mol. Life Sci. 60, 1281-1295.

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

56. Clarke, S. (1987) Propensity for spontaneous succinimide formation from aspartyl and asparaginyl residues in cellular proteins. Int. J. Peptide Protein Res. 30, 808-821.

57. Thorne, C.J.R. (1978) Techniques in Protein and Enzyme Biochemistry. 2-18.

58. Patananan, A.N., Capri, J., Whitelegge, J.P., Clarke, S.G. (2014) Non-repair Pathways for Minimizing Protein Isoaspartyl Damage in the Yeast Saccharomyces cerevisiae. J. Biol. Chem. 289, 16936-16953.

59. Schneider, C. A.; Rasband, W. S. & Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis. Nature methods. 9, 671-675.

Accession	Description	Coverage [%]	# Unique Peptides
P02489	Alpha-crystallin A	75	74
P02511	Alpha-crystallin B	78	54
P05813	Beta-crystallin A3	55	15
P53673	Beta-crystallin A4	57	24
P53674	Beta-crystallin B1	52	31
P43320	Beta-crystallin B2	67	25
Q12934	Filensin	30	73
P07315	Gamma-crystallin C	5	1
P22914	Gamma-crystallin S	47	15
Q13515	Phakinin	46	68

TABLE 1. Parent proteins of lens fragments. Lens peptides were isolated from the gel filtration HMW fraction and identified as described in “Experimental Procedures” against the 10 proteins in the table below. The sequence coverage and number of unique peptides are reported.

Parent protein	Sequence	Theo. MH+ [Da]	Relative abundances 44 year old	Relative abundances 76 year old
αA crystallin	52 – LFRTVLDSGISEVR – 65	1591.88023	2.2	197.8
	56 – VLDSGISEVR – 65	1074.57896	12.9	187.1
	134 – SADGMLTFCGPKIQ – 147	1524.71851	6	194
	54 – RTVLDSGISEVR – 65	1331.72775	68.3	131.7
	51 – SLFRTVLDSGISEVR – 65	1678.91226		200
αB crystallin	93 – VLGDVIEVH – 101	980.54112	57.2	142.8
	29 – GEHLLESDLFPT – 40	1357.66342	9.8	190.2
	123 – RIPADVDPLTIT – 134	1310.73144	2.2	197.8
	123 – RIPADVDPLTITS – 135	1397.76347	2.7	197.3
	95 – GDVIEVH – 101	768.38864		200
βA3 crystallin	29 – GPWKITIYD – 37	1092.57242	200	
	152 – GWFNNEVGSMKIQ – 164	1509.71547		200
	127 – TIFEKENFIGRQ – 138	1481.7747	16	184
	124 – SKMTIF – 129	726.38547	1.4	198.6
	139 – WEISDDYPSLQAM – 151	1554.67809	0.4	199.6
βA4 crystallin	108 – TIFEQENFLGKK – 119	1453.76856		200
	109 – IFEQENFLGKK – 119	1352.72088	1.4	198.6
	120 – GELSDDYPSLQAM – 132	1425.62024	1.8	198.2
	105 – SRLTIFEQENFLGKK – 119	1809.98576		200
	104 – DSRLTIFEQENFLGKK – 119	1925.0127		200
βB1 crystallin	238 – HLEGSFPVLA – 247	1069.56767		200
	238 – HLEGSFPVLATEPPK – 252	1621.85843	21	179
	240 – EGSFPVLATEPPK – 252	1371.71546	2.3	197.7
	242 – SFPVLATEPPK – 252	1185.6514	54.3	145.7
	38 – TLAPTTVPITSAK – 50	1299.75184	22.2	177.8
βB2 crystallin	121 – KMEIIDDDVPSFHAH – 135	1753.8214		200
	121 – KMEIIDDDVPSFHA – 134	1616.76248		200
	121 – KMEIIDDDVPSFHAHG – 136	1810.84286		200
	122 – MEIIDDDVPSFHAH – 135	1625.72643		200
	143 – SVRVQSGTWVGYQYPGYRGL – 162	2273.14618		200
Filensin	79 – GELAGPEDALARQVE – 93	1554.77583	66.3	133.7
	324 – FIETPIPLFTQSH – 336	1529.79986		200
	239 – RVELQAQTTTLEQAIK – 254	1829.0127	5.8	194.2
	324 – FIETPIPLFTQ – 334	1305.70892	1.1	198.9
	191 – QQIIHTTPPASIVTS – 205	1592.86425	2.6	197.4
γS crystallin	60 – YILPQGEYPEYQRWM – 74	1972.9262		200
	98 – IFEKGDFSGQ – 107	1127.53677		200
	65 – GEYPEYQRWM – 74	1358.5834	13.2	186.8
	50 – YERPNFAGYM – 59	1247.55137		200
	98 – IFEKGDFSGQMYETTED – 114	1996.84806	10	190
Phakinin	162 – QQVGEAVLENARL – 174	1426.76487		200
	157 – WASSCQQVGEAVLENARL – 174	2017.976		200
	211 – KVIDEANLTKM – 221	1261.68205	6.3	193.7
	266 – TGLDDILETIRIQ – 278	1486.81115		200
	162 – QQVGEAVLENARLM – 175	1557.80535	4	196
γC crystallin	166 – SLRRVVDLY – 174	1120.64732		200

TABLE 2. Top five peptides identified from each of 10 abundant lens proteins. Lens peptides were isolated from the gel filtration HMW fraction and identified as described in "Experimental Procedures." The five peptides with the highest ion scores from each of the proteins listed in Table 1 are displayed with their corresponding theoretical mass and relative abundances in a 44-year-old and 76-year-old whole lens sample on a scale of 0-200. Full peptide tables for each protein can be found in Supporting Tables 1-9.

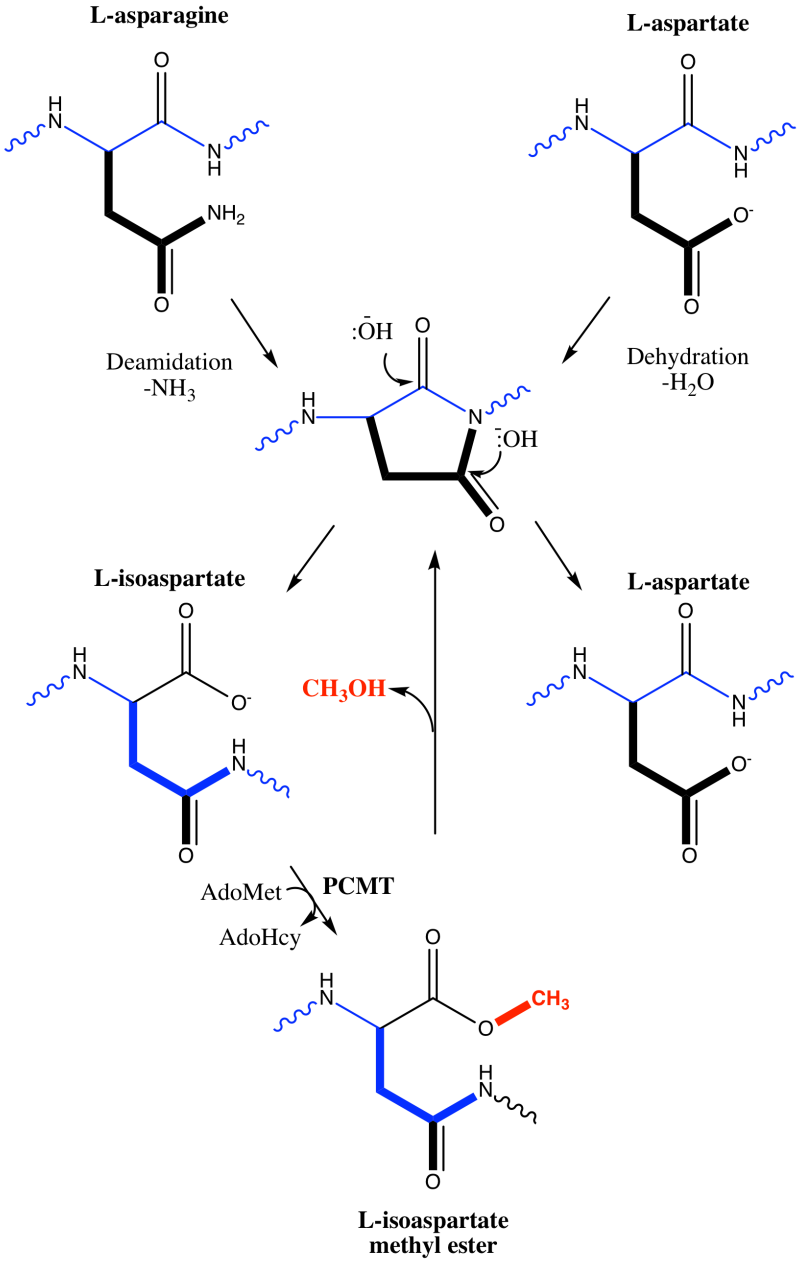


FIG. 1. Pathway for L-isoaspartate (L-isoAsp) formation and repair. Normal L-asparagine and L-aspartate residues (top), can undergo deamidation or dehydration, respectively, to yield a five membered L-succinimide ring (center) which is readily hydrolysed under cellular conditions to yield either L-aspartate (~15-40% of product), or, more frequently, L-isoaspartate (~60-85%; lower left (15)). This abnormal residue can be repaired by reactions initiated by the L-isoaspartyl/D-

aspartyl *O*-methyltransferase PCMT1, which uses *S*-adenosylmethionine (AdoMet) as a methyl donor to create a methyl ester that can be quickly hydrolysed back under cellular conditions to the L-succinimide intermediate, allowing the reformation of L-aspartate residues. The L-succinmide ring can racemize during this pathway and yield D-aspartate and D-isoaspartate isomers which are not shown here. D-aspartate is an additional substrate for the PCMT1 repair enzyme (albeit with k_{cat}/K_m values reduced 1000-fold or more (17), while D-isoaspartate is not.

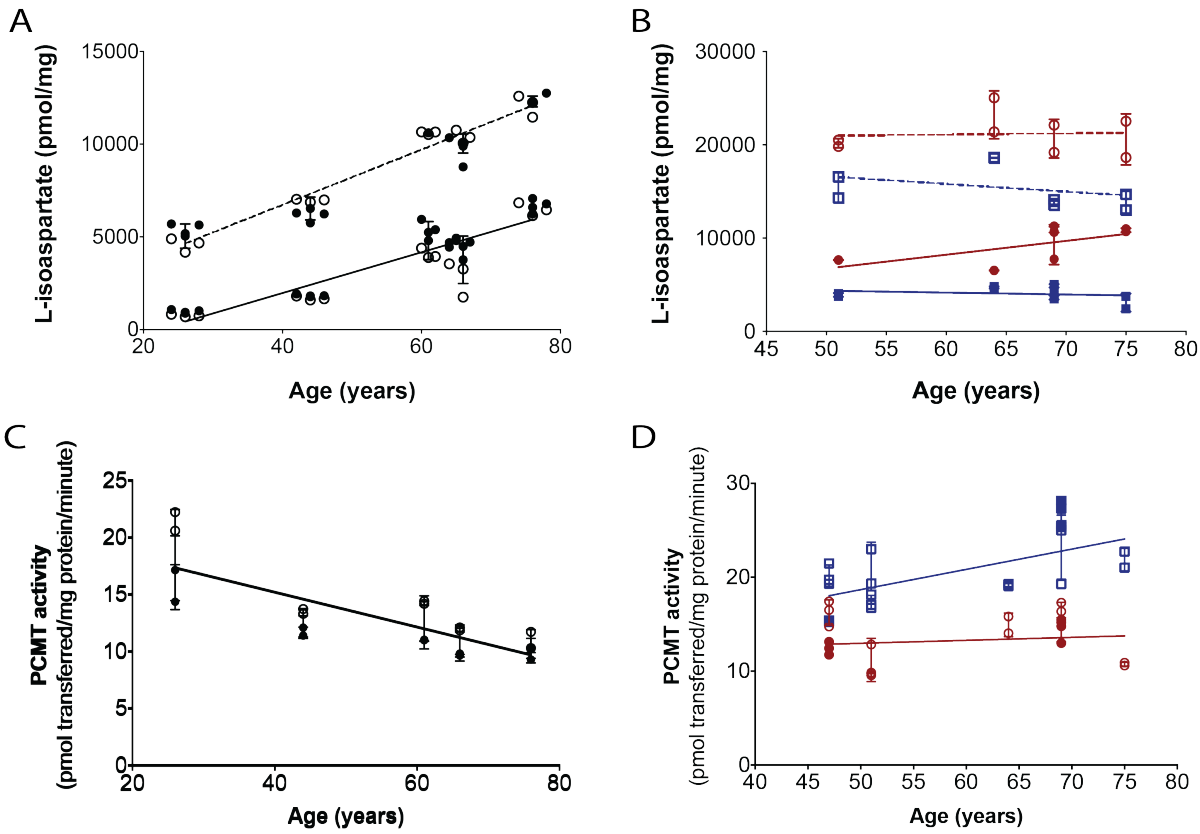


FIG. 2: The L-isoAsp modification accumulates to high levels within the aged human lens polypeptides while endogenous methyltransferase repair activity remains high with age. Lens extracts (25 µg of protein) were radiolabeled by 6 µg PCMT1 and 10 µM [3H] AdoMet for 2 h at 37 °C as described in the “Experimental Procedures” section . A) L-isoAsp quantified in whole lens extracts. Closed circles represent a single eye lens; open circles represent the other eye from the same individual. The solid line includes the soluble (WS) extracts; the dashed line includes the urea-solubilized water-insoluble (WI) extracts. Each symbol represents one technical replicate. Error bars represent standard deviation of the biological replicates. Lines were linear regression fits in Graphpad, and slope was significantly non-zero for both WS and WI extracts at 0.0048 and 0.0045 p-values, respectively. B) L-isoAsp levels quantified in dissected lens nuclear (red circles) and cortical (blue squares) extracts. The solid line and closed symbols represent the WS extracts; the dashed line and open symbols represent the WI extracts. Each symbol represents one technical replicate. Error bars represent the standard

deviation of the technical replicates. Lines were linear regression fits performed in Graphpad and none of the lines had significantly non-zero slopes with p-values greater than 0.05. C) PCMT1 activity in whole lens extracts was quantified by detecting the amount of [3H]AdoMet radioactivity transferred to 100 μM peptide substrate KASA(isoD)LAKY by 15 μg of lens extract protein in 2 h at 37 °C. Closed circles represent one lens from an individual; open circles represent the other lens from that individual. Each symbol represents one technical replicate. Error bars represent the standard deviation of the technical replicates. Line was a linear regression fit performed in Graphpad, and slope was significantly non-zero with a p-value of 0.0258. D) PCMT1 activity was quantified in dissected lens nuclear (red circles) and cortical extracts (blue squares). Open and closed symbols represent each lens from one individual at 47- and 69-years old. Each symbol represents a technical replicate. Error bars represent the standard deviation of the technical replicates. Lines were linear regression fits performed in Graphpad and neither of the lines had significantly non-zero slopes with p-values greater than 0.05.

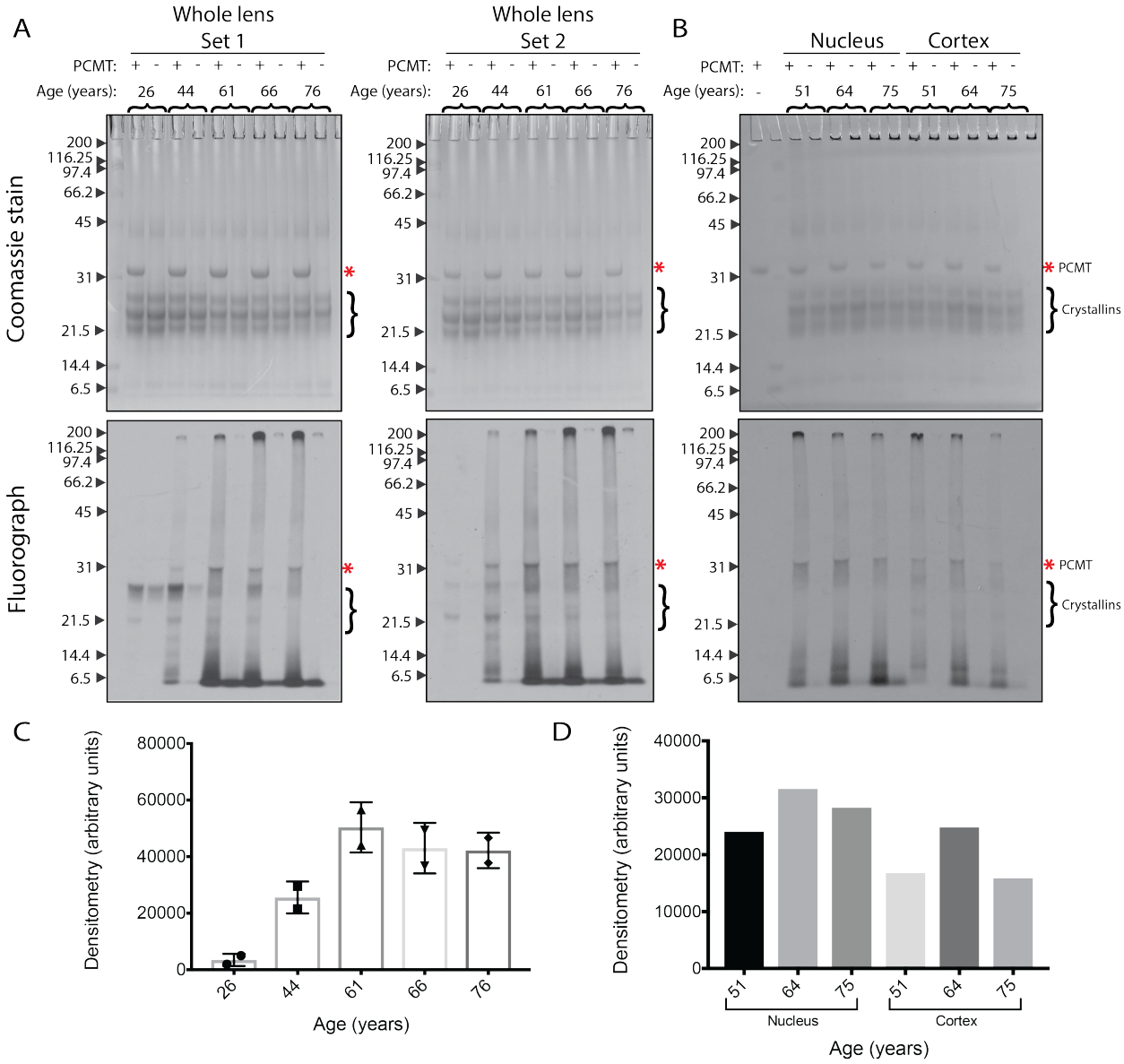


FIG. 3: L-isoAsp damage accumulates with age throughout a range of polypeptide sizes in water-soluble extracts of human lenses, primarily within LMW species. WS extracts (25 µg protein) were incubated with 0.3 µM [3H] AdoMet and with (+) or without (-) 5 µg PCMT1 for 2 h at 37 °C and polypeptides were analyzed by SDS-PAGE using a 12% acrylamide matrix. Coomassie-stained gels are shown at the top; fluorographs (2 day exposures) are shown at the bottom. Molecular weight markers (kDa) are indicated with arrows to the left of each gel and include myosin, β-galactosidase, phosphorylase b, bovine serum albumin, ovalbumin, carbonic anhydrase, soybean trypsin inhibitor, lysozyme, and aprotinin (Bio-Rad SDS-PAGE Molecular

Weight Standards, Broad Range, catalog #161-0317). The band corresponding to PCMT1 is indicated by asterisks. The migration positions of intact crystallins are indicated by brackets on the right hand side of the gels. A) Whole lens extracts from two lenses of each individual. B) Nuclear and cortical separated extracts from a single lens (1 day exposure). C) Densitometry of lanes from whole lens sets 1 and 2. Error bars represent the standard deviation of the two lanes. D) Densitometry of lanes from nuclear and cortical sets.

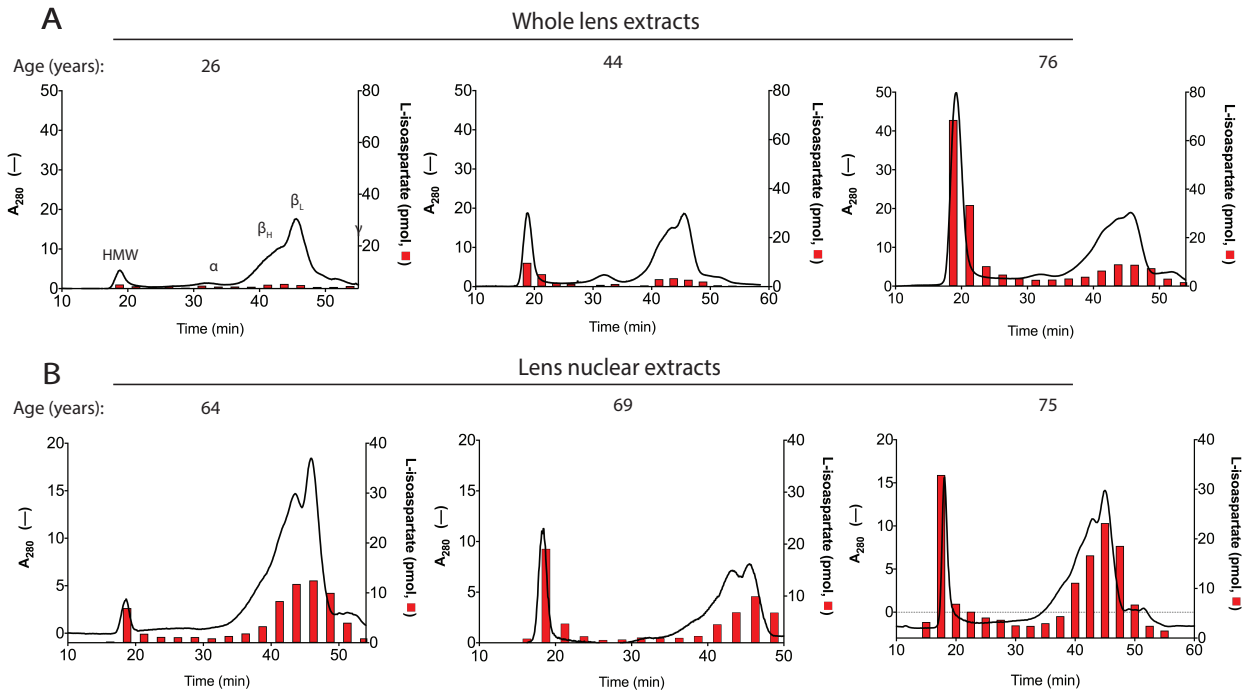


FIG. 4. Size exclusion chromatography of WS extracts demonstrates age-dependent increases in HMW protein species and in L-isoAsp damage in crystallins, particularly within the HMW fractions. One mg of lens extract was loaded onto a 24 mL Superose 6, 10/300 GL gel filtration column and proteins were eluted at 0.4 mL/min in 50 mM Tris-HCl, pH 7.9, 150 mM NaCl. One mL fractions were collected, and 25 µL of each fraction was removed for L-isoAsp quantification (red columns) as described in the "Experimental Procedures" section. The black lines represent UV absorbance at 280 nm. Peaks were defined throughout the text as labeled in the "26-year-old" S.E.C. panel. These designations are based off of the known molecular weights of α- and β-crystallin oligomers, coupled with the void volume of the column. A) Whole lens extracts, B) Nuclear extracts. The age of lens in years is designated on top of each corresponding graph.

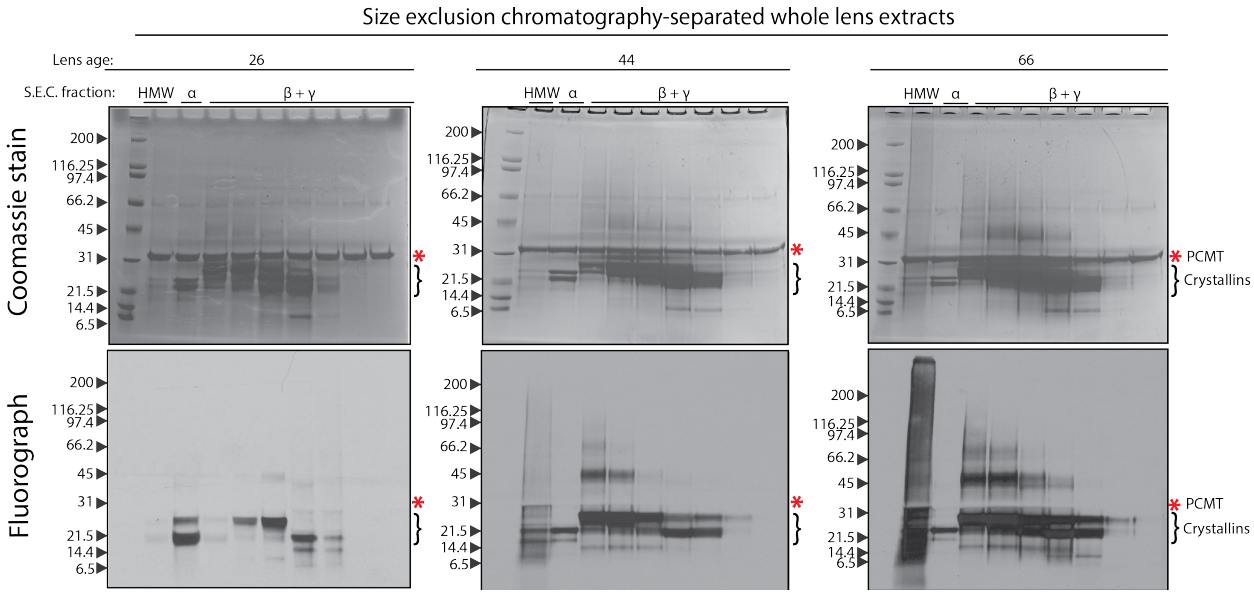


FIG. 5: Analysis of L-isoAsp damage in the native size-exclusion fractions of human whole lens extracts by SDS-PAGE reveals the HMW gel filtration fraction contains significant amounts of LMW L-isoaspartyl-containing species. Fractions from size exclusion chromatography as in Fig. 4 were trichloroacetic acid-precipitated, resuspended in 100 mM Tris-HCl, pH 7.9, 150 mM NaCl, 0.1% SDS, 6 M Urea and L-isoAsp sites within the fractions were radiolabeled by PCMT1 and [^{3}H] AdoMet as described in the Fig. 3 legend. Labeled fractions were then separated by SDS-PAGE on a 4-12% gradient gel. The Coomassie-stained gel is shown in the upper panels; the fluorograph (one week exposure) is shown in the lower panels. The bands corresponding to PCMT1 in the Coomassie-stained gel are indicated by asterisks. The migration positions of intact crystallins are indicated by brackets on the right hand side of the gels. Labels above lanes represent the corresponding peaks from the size exclusion run (see Fig. 4). The age of lens sample in years is designated above the corresponding gel.

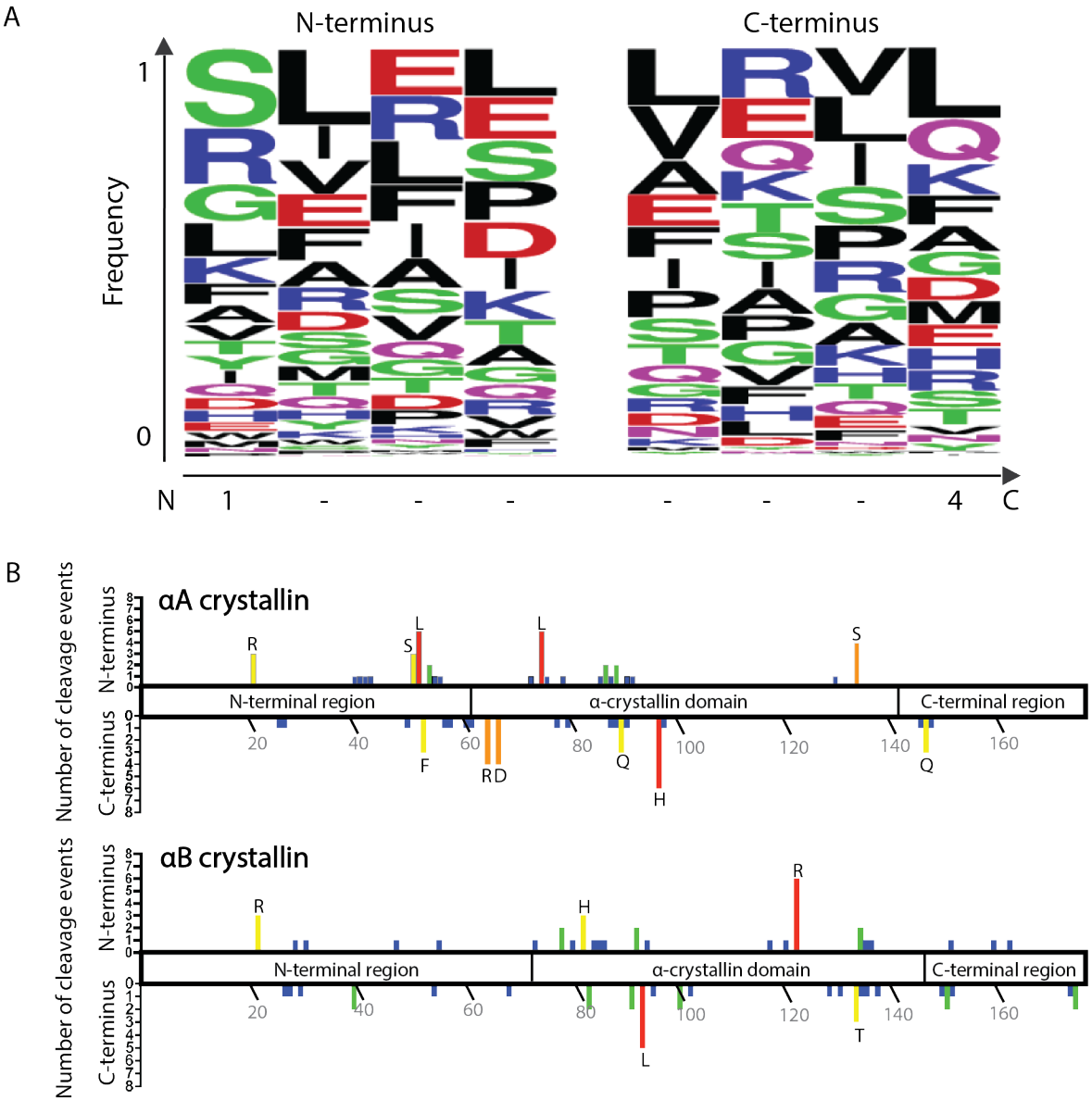


FIG. 6: Identification of crystallin fragments from the gel filtration WS-HMW fraction reveals the presence of peptides from a variety of crystallins and other lens proteins with no strong trends in residue specificity or localization of cleavage sites. A, The four amino acid residues present at the N- and C-terminal sequences of the peptides identified by mass spectrometry in the WS-HMW fraction were inputted into WebLogo. The y-axis represents the fraction of a particular residue that occurs at that position. On the x-axis, position 1 of the N-terminus represents the first residue of the cleaved peptide, position 4 of the C-terminus represents the last residue of the cleaved peptide. B, Graphical representation of the cleavage sites observed in

WS-HMW peptides along the primary sequence of αA-crystallin (upper panel) and αB-crystallin (lower panel). Residues with 3 or more cleavages are designated by their single letter amino acid code. Similar patterns were observed with other lens proteins. Grey numbers designate every 20^{th} residue in the sequence.

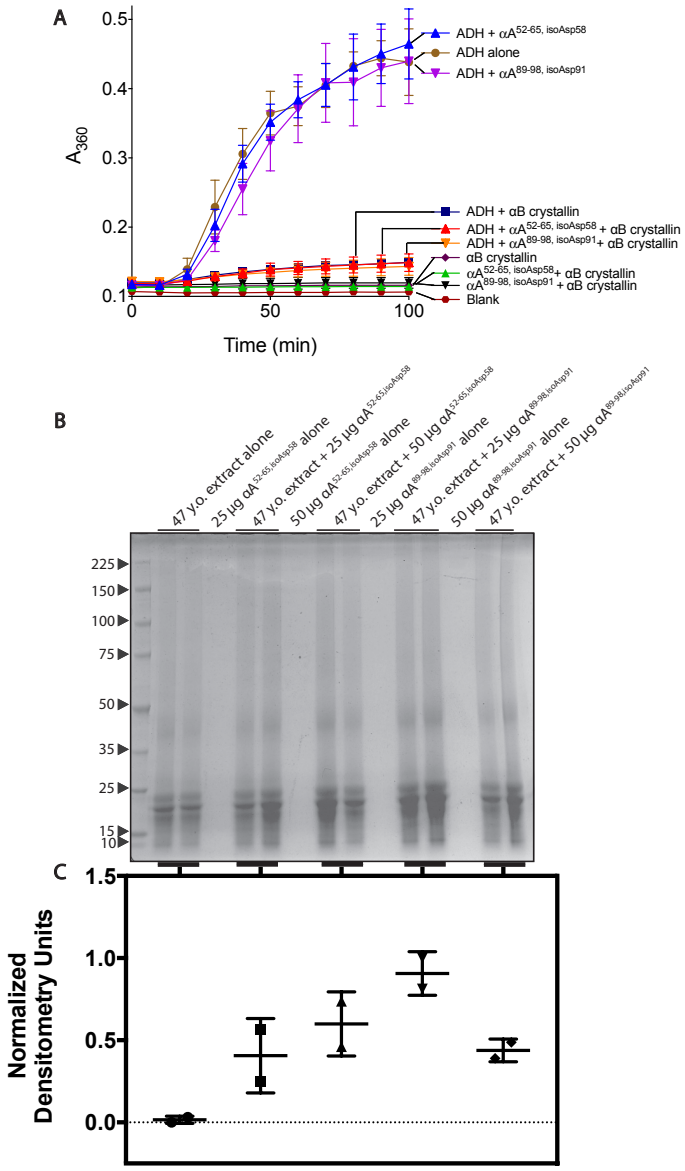


FIG. 7. The isomerized forms of WS-HMW derived synthetic peptides αA$^{52\text{-}65,\ isoAsp58}$ and αA$^{89\text{-}98,\ isoAsp91}$ do not inhibit αB crystallin chaperone activity, but modestly increase protein precipitation *in vitro*. A, Inhibition assays were performed as described in the "Experimental Procedures" section. Briefly, alcohol dehydrogenase (ADH, 125 µg) was incubated under denaturing conditions (50 mM sodium phosphate, pH 7.0, 100 mM NaCl, and 10 mM phenanthroline) with and without **αB** crystallin (50 µg) and peptides (50 µg). Aggregation was monitored by light scattering at $A_{360\ nm}$. Symbols represent the mean and errors bars represent the standard deviation of three technical replicates. B, 47 year old lens extract (500 µg) was

incubated with both 25 and 50 µg of the $\alpha A^{52\text{-}65,\ isoAsp58}$ or $\alpha A^{89\text{-}98,\ isoAsp91}$ peptides (from stocks made in water) for 14 h at 37 °C. Reactions were spun for 10 min at 3,000 rpm. The resulting pellet was separated by SDS-PAGE. Molecular weight markers (kDa) are indicated with arrows to the left of the gel (Perfect Protein Markers, 10-225 kDa, Millipore Sigma, Catalog # 69079). C, The lanes of the gel were quantified by densitometry and normalized from 0-1 based on the lowest and highest values. Error bars represent the range of two replicates.

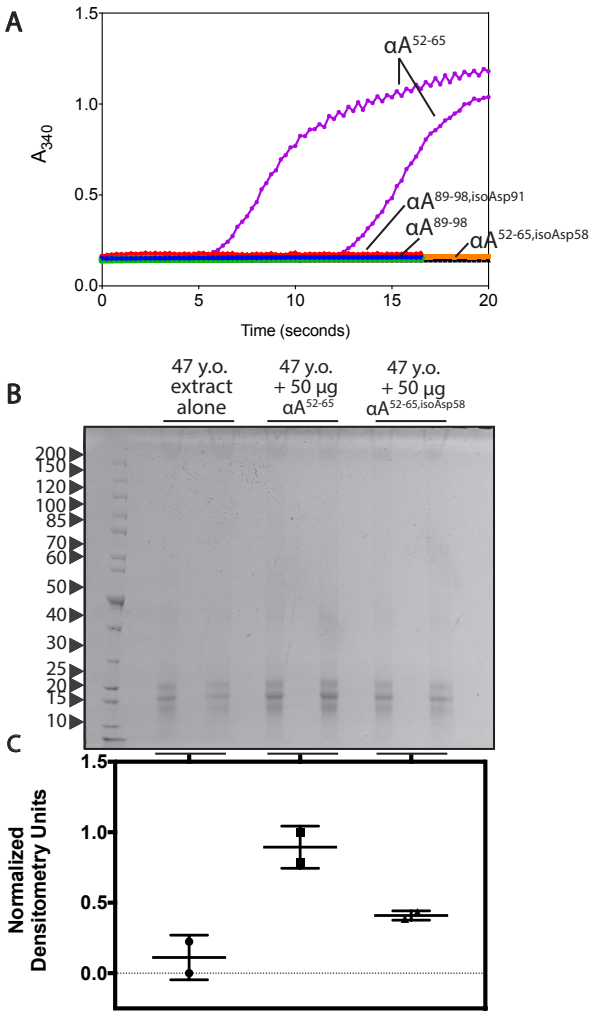


FIG. 8. **The native form of WS-HMW derived synthetic peptide αA$^{52-65}$ shows enhanced self-aggregation and protein precipitation compared to the isomerized form**. A, The isomerized and native αA$^{52-65}$ peptides were dissolved in 50 mM Tris-HCl, pH 7.5 and 1% DMSO at 3 mg/ml. In order to solubilize the native αA$^{89-98}$ peptide, the isomerized and native forms were first dissolved in 3% ammonia in water as described "Experimental Procedures" and subsequently diluted into in 50 mM Tris-HCl, pH 7.5 at a final concentration of 2 mg/ml. Aggregation was analyzed by light scattering at 340 nm. Assays were performed in duplicate and one line for each duplicate is shown. Lines pertaining to peptide solutions are labeled, the black line represents the buffer background for the αA$^{52-65}$ condition, while the green line represents the buffer background for the αA$^{89-98}$ condition. B, 47 year old lens extract (500 µg) was incubated with 50 µg of the αA^{52-}

^{65} peptides (from stocks made in water) for 14 h at 37 °C. Reactions were spun for 10 min at 3,000 rpm. The resulting pellet was separated by SDS-PAGE. Molecular weight markers (kDa) are indicated with arrows to the left of the gel (PageRuler Unstained Protein Ladder, 10-200 kDa, ThermoFisher Scientific, Catalog # 26614). C, The lanes of the gel were quantified by densitometry and normalized from 0-1 based on the lowest and highest values. Error bars represent the range of two replicates.

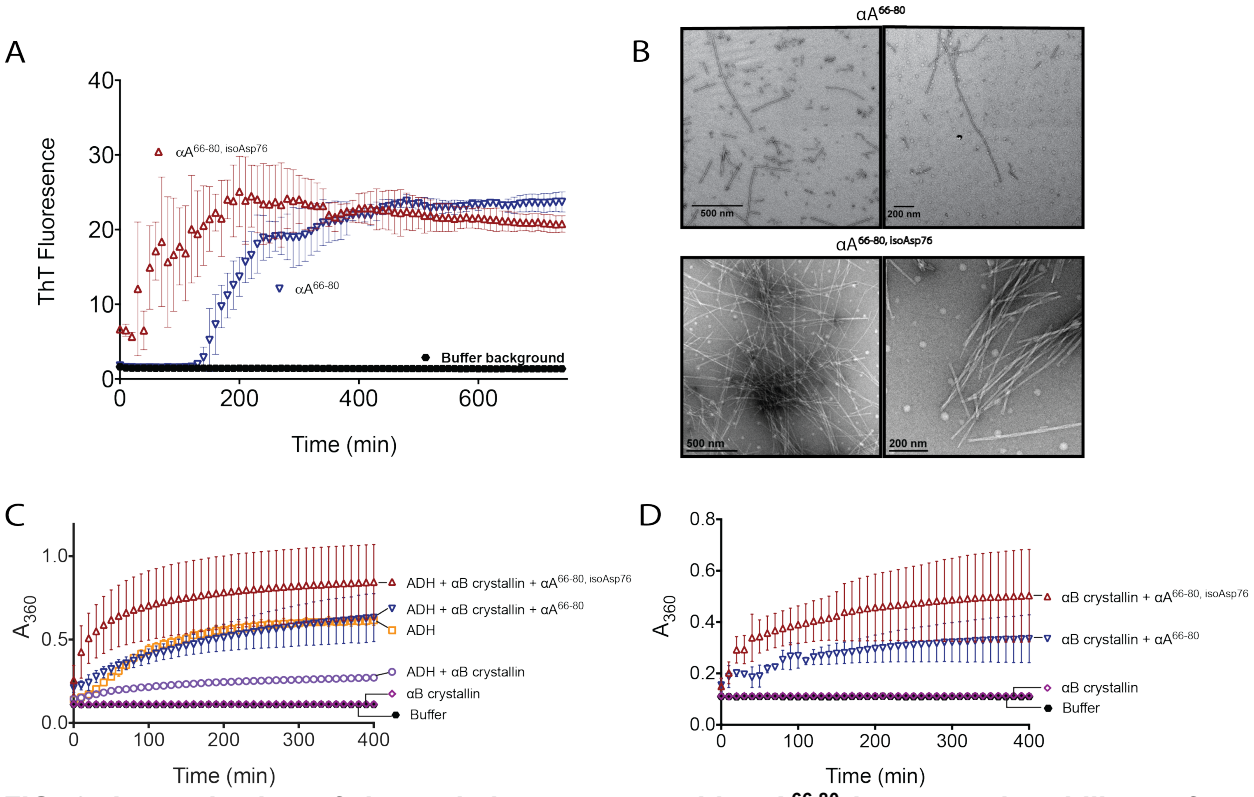


FIG. 9: Isomerization of the anti-chaperone peptide αA$^{66-80}$ increases its ability to form amyloid and enhances its anti-chaperone activity in vitro. A, αA$^{66-80}$ (solid line) and αA$^{66-80,\ isoAsp76}$ (dashed line) peptides were incubated with amyloid-binding dye Thioflavin T to monitor self-aggregation as described in the "Experimental Procedures" section. B, Peptides were incubated without ThT and the resulting aggregates were observed by electron microscopy as described in the "Experimental Procedures" section. The scale bar represents either 200 or 500 nm as indicated. Two representative fields are shown. C, Aggregation was measured for alcohol dehydrogenase (ADH, 150 μg) incubated under denaturing conditions (50 mM sodium phosphate, pH 7.0, 100 mM NaCl, and 10 mM phenanthroline) with and without αB crystallin (40 μg) and peptides (40 μg). Aggregation was monitored by light scattering at $A_{360\ nm}$. D, Control aggregation reactions were performed as in Panel C with peptides and αB crystallin alone. Data sets from panels C and D were from a single experiment and were separated onto two graphs for clarity, the αB crystallin alone and buffer background from panels C and D are duplicated.

SUPPLEMENTARY MATERIALS:

Sequence	Positions in Master Proteins	Theo. MH+ [Da]	Abundances 44 year old	Abundances 76 year old	Ion Score
LFRTVLDSGISEVR	P02489 [52-65]	1591.88023	2.2	197.8	70
VLDSGISEVR	P02489 [56-65]	1074.57896	12.9	187.1	65
SADGMLTFCGPKIQ	P02489 [134-147]	1524.71851	6	194	62
RTVLDSGISEVR	P02489 [54-65]	1331.72775	68.3	131.7	61
SLFRTVLDSGISEVR	P02489 [51-65]	1678.91226		200	61
SADGMLTFCGPKI	P02489 [134-146]	1396.65993		200	61
SADGMLTFCGPKIQ	P02489 [134-147]	1540.71343	21.3	178.7	61
LDVKHFSPEDLTVK	P02489 [75-88]	1627.869	0.8	199.2	58
RTVLDSGISEVRSD	P02489 [54-67]	1533.78672	191	9	57
TVLDSGISEVRSD	P02489 [55-67]	1377.68561	200		55
VQDDFVEIH	P02489 [89-97]	1101.52111	22.9	177.1	51
LDVKHFSPEDLTVKVQD	P02489 [75-91]	1970.02293	1.6	198.4	51
STISPYYRQSLF	P02489 [42-53]	1461.73726	0.9	199.1	49
DDFVEIH	P02489 [91-97]	874.39412	85.5	114.5	47
KVQDDFVEIH	P02489 [88-97]	1229.61608	183.7	16.3	45
DVKHFSPEDLTVKVQD	P02489 [76-91]	1856.93887	28.1	171.9	42
VQDDFVEIHG	P02489 [89-98]	1158.54258		200	42
IFLDVK	P02489 [73-78]	734.4447	181.3	18.7	37
SCSLSADGMLTFCGPKIQ	P02489 [130-147]	1971.89728	6.2	193.8	37
VKVQDDFVEIH	P02489 [87-97]	1328.68449	173.1	26.9	37
RLFDQFFGE	P02489 [21-29]	1158.55783	42.2	157.8	37
TISPYYRQSLF	P02489 [43-53]	1374.70523	10.3	189.7	36
RLFDQF	P02489 [21-26]	825.42536	75.9	124.1	35
LFRTVLDSGI	P02489 [52-61]	1120.63609		200	34
LDVKHFSPEDLTVKVQ	P02489 [75-90]	1854.99599	0.6	199.4	34
SADGMLTFCGPKIQT	P02489 [134-148]	1625.76619		200	34
LSSTISPYYRQ	P02489 [40-50]	1314.66884	200		34
LFRTVLDSGISEVRSD	P02489 [52-67]	1793.9392	24.8	175.2	33
HFSPEDLTVKVQ	P02489 [79-90]	1399.72161		200	33
LFRTVLD	P02489 [52-58]	863.49853	87.7	112.3	33
SLFRTVLDSGIS	P02489 [51-62]	1294.70014	51.8	148.2	33
SSTISPYYRQSLF	P02489 [41-53]	1548.76928	10.8	189.2	32
TVKVQDDFVEIH	P02489 [86-97]	1429.73217	200		32
DFVEIH	P02489 [92-97]	759.36718	41.4	158.6	29
LDVKHFSPEDLTVKV	P02489 [75-89]	1726.93741		200	29
LDVKHF	P02489 [75-80]	758.41955	5.6	194.4	28
RLFDQFF	P02489 [21-27]	972.49378	80.7	119.3	27
SLFRTVLDSGISEVRSD	P02489 [51-67]	1880.97123	27.9	172.1	26

TABLE S1: αA crystallin peptides. Lens peptides were isolated from the gel filtration HMW fraction and identified as described in “Experimental Procedures.” All peptides with an ion score

of 26 or higher are displayed with their corresponding theoretical mass and relative abundances in a 44-year-old and 76-year-old whole lens sample on a scale of 0-200.

Sequence	Positions in Master Proteins	Theo. MH+ [Da]	Abundances 44 year old	Abundances 76 year old	Ion Score
VLGDVIEVH	P02511 [93-101]	980.54112	57.2	142.8	73
GEHLLESDLFPT	P02511 [29-40]	1357.66342	9.8	190.2	61
RIPADVDPLTIT	P02511 [123-134]	1310.73144	2.2	197.8	49
RIPADVDPLTITS	P02511 [123-135]	1397.76347	2.7	197.3	46
GDVIEVH	P02511 [95-101]	768.38864		200	43
HFSPEELKVKVLGD	P02511 [83-96]	1597.85843		200	41
EELKVKVL	P02511 [87-94]	957.59791	8.9	191.1	39
RAPSWFDTGLSEMR	P02511 [56-69]	1652.78495		200	38
VLGDVIEVHGK	P02511 [93-103]	1165.65755		200	37
REEKPAVTAAPKK	P02511 [163-175]	1424.82199		200	37
RIPADVD	P02511 [123-129]	785.41519	1.8	198.2	37
HLLESDLFPT	P02511 [31-40]	1171.59937	5.6	194.4	36
RLFDQFFGE	P02511 [22-30]	1158.55783	42.2	157.8	36
SSLSSDGVLTVNGPRKQ	P02511 [135-151]	1744.9188	69.4	130.6	34
RLFDQF	P02511 [22-27]	825.42536	75.9	124.1	34
REEKPAVTAAPK	P02511 [163-174]	1296.72703		200	34
YLRPPSFL	P02511 [48-55]	992.55638	17.1	182.9	34
HFSPEELKVKVL	P02511 [83-94]	1425.81003	0.6	199.4	33
SLSSDGVLTVNGPRKQ	P02511 [136-151]	1657.88677	73.3	126.7	33
LSSDGVLTVNGPRK	P02511 [137-150]	1442.79617		200	31
FHRKYRIPADVDPLTIT	P02511 [118-134]	2042.11817		200	31
HFSPEELKVK	P02511 [83-92]	1213.65755		200	31
PEELKVKVL	P02511 [86-94]	1054.65067	6.2	193.8	30
RIPADVDPLTITSS	P02511 [123-136]	1484.7955		200	30
SPEELKVKVL	P02511 [85-94]	1141.6827	9.8	190.2	30
VSGPERTIPITREEKPAVTAAPKK	P02511 [152-175]	2575.45661	74.6	125.4	29
RFSVNLDVKHF	P02511 [74-84]	1361.73245	21.8	178.2	28
LDVKHFSPEELKVK	P02511 [79-92]	1668.93193		200	28
RIPADVDPL	P02511 [123-131]	995.55202	20.2	179.8	27
SSLSSDGVLTVNGPRKQV	P02511 [135-152]	1843.98722		200	27
LDVKHF	P02511 [79-84]	758.41955	5.6	194.4	27
KYRIPADVDPLTIT	P02511 [121-134]	1601.88973		200	26
RIPADVDPLTITSSLS	P02511 [123-138]	1684.91159		200	26
RLFDQFF	P02511 [22-28]	972.49378	80.7	119.3	26
VKHFSPEELKVKVL	P02511 [81-94]	1652.9734		200	26

TABLE S2: αB crystallin peptides. Lens peptides were isolated from the gel filtration HMW fraction and identified as described in “Experimental Procedures.” All peptides with an ion score of 26 or higher are displayed with their corresponding theoretical mass and relative abundances in a 44-year-old and 76-year-old whole lens sample on a scale of 0-200.

Sequence	Positions in Master Proteins	Theo. MH+ [Da]	Abundances 44 year old	Abundances 76 year old	Ion Score
GPWKITIYD	P05813 [29-37]	1092.57242	200		47
GWFNNEVGSMKIQ	P05813 [152-164]	1509.71547		200	28
TIFEKENFIGRQ	P05813 [127-138]	1481.7747	16	184	27
SKMTIF	P05813 [124-129]	726.38547	1.4	198.6	27
WEISDDYPSLQAM	P05813 [139-151]	1554.67809	0.4	199.6	26

TABLE S3: βA3 crystallin peptides. Lens peptides were isolated from the gel filtration HMW fraction and identified as described in “Experimental Procedures.” All peptides with an ion score of 26 or higher are displayed with their corresponding theoretical mass and relative abundances in a 44-year-old and 76-year-old whole lens sample on a scale of 0-200.

Sequence	Positions in Master Proteins	Theo. MH+ [Da]	Abundances 44 year old	Abundances 76 year old	Ion Score
TIFEQENFLGKK	P53673 [108-119]	1453.76856		200	55
IFEQENFLGKK	P53673 [109-119]	1352.72088	1.4	198.6	45
GELSDDYPSLQAM	P53673 [120-132]	1425.62024	1.8	198.2	45
SRLTIFEQENFLGKK	P53673 [105-119]	1809.98576		200	42
DSRLTIFEQENFLGKK	P53673 [104-119]	1925.0127		200	42
SGAWVCSQFPGYRGF	P53673 [146-160]	1718.77439	30.8	169.2	36
ERLTSFRPA	P53673 [89-97]	1076.58472		200	32
FQVQSIR	P53673 [186-192]	877.48903	24.1	175.9	32
GELSDDYPSLQ	P53673 [120-130]	1223.54264	6.3	193.7	31
SRLTIF	P53673 [105-110]	736.4352	8.8	191.2	30
SGAWVCSQFPGYRG	P53673 [146-159]	1571.70597	115.8	84.2	29
GWEGNEVGSFHVH	P53673 [133-145]	1454.64475	200		29
FLGKKGELSDDYPSLQAM	P53673 [115-132]	1998.9841		200	27
KGELSDDYPSLQAM	P53673 [119-132]	1553.7152		200	27

TABLE S4: βA4 crystallin peptides. Lens peptides were isolated from the gel filtration HMW fraction and identified as described in “Experimental Procedures.” All peptides with an ion score of 26 or higher are displayed with their corresponding theoretical mass and relative abundances in a 44-year-old and 76-year-old whole lens sample on a scale of 0-200.

Sequence	Positions in Master Proteins	Theo. MH+ [Da]	Abundances 44 year old	Abundances 76 year old	Ion Score
HLEGSFPVLA	P53674 [238-247]	1069.56767		200	62
HLEGSFPVLATEPPK	P53674 [238-252]	1621.85843	21	179	59
EGSFPVLATEPPK	P53674 [240-252]	1371.71546	2.3	197.7	55
SFPVLATEPPK	P53674 [242-252]	1185.6514	54.3	145.7	54
TLAPTTVPITSAK	P53674 [38-50]	1299.75184	22.2	177.8	50
RDKQWHLEGSFPVLATEPPK	P53674 [233-252]	2335.21934		200	48
RLRDKQWHLEGSFPVLATEPPK	P53674 [231-252]	2604.40452		200	45
VVFELENFQGR	P53674 [62-72]	1337.68483		200	44
GSFPVLATEPPK	P53674 [241-252]	1242.67286	39.7	160.3	39
WHLEGSFPVLATEPPK	P53674 [237-252]	1807.93775		200	39
MFILEKG	P53674 [113-119]	837.45389	18.7	181.3	38
NFRGEMFILEKG	P53674 [108-119]	1440.7304		200	30
TTLAPTTVPITSA	P53674 [37-49]	1272.70456	12.7	187.3	28
LEGSFPVLATEPPK	P53674 [239-252]	1484.79952	38.9	161.1	28
WHLEGSFPVLA	P53674 [237-247]	1255.64698		200	27
RLMSFRPIKM	P53674 [135-144]	1278.71733		200	26

TABLE S5: βB1 crystallin peptides. Lens peptides were isolated from the gel filtration HMW fraction and identified as described in “Experimental Procedures.” All peptides with an ion score of 26 or higher are displayed with their corresponding theoretical mass and relative abundances in a 44-year-old and 76-year-old whole lens sample on a scale of 0-200.

Sequence	Positions in Master Proteins	Theo. MH+ [Da]	Abundances 44 year old	Abundances 76 year old	Ion Score
KMEIIDDDVPSFHAH	P43320 [121-135]	1753.8214		200	51
KMEIIDDDVPSFHA	P43320 [121-134]	1616.76248		200	48
KMEIIDDDVPSFHAHG	P43320 [121-136]	1810.84286		200	41
MEIIDDDVPSFHAH	P43320 [122-135]	1625.72643		200	35
SVRVQSGTWVGYQYPGYRGL	P43320 [143-162]	2273.14618		200	34
KMEIIDDDVPSFH	P43320 [121-133]	1545.72537		200	33
SGTWVGYQYPGYRGL	P43320 [148-162]	1703.81763		200	33
KMEIIDDDVPSFHAHGYQEKVSSVRVQ	P43320 [121-147]	3114.53131	3.1	196.9	31
PKIIIF	P43320 [17-22]	730.48617	8.3	191.7	29
SHELNGPCPNLKETGVEKAGSVLVQ	P43320 [31-55]	2663.34574		200	26
EIIDDDVPSFHAH	P43320 [123-135]	1494.68595	5.4	194.6	26

TABLE S6: βB2 crystallin peptides. Lens peptides were isolated from the gel filtration HMW fraction and identified as described in “Experimental Procedures.” All peptides with an ion score of 26 or higher are displayed with their corresponding theoretical mass and relative abundances in a 44-year-old and 76-year-old whole lens sample on a scale of 0-200.

Sequence	Positions in Master Proteins	Theo. MH+ [Da]	Abundances 44 year old	Abundances 76 year old	Ion Score
GELAGPEDALARQVE	Q12934 [79-93]	1554.77583	66.3	133.7	77
FIETPIPLFTQSH	Q12934 [324-336]	1529.79986		200	66
RVELQAQTTTLEQAIK	Q12934 [239-254]	1829.0127	5.8	194.2	61
FIETPIPLFTQ	Q12934 [324-334]	1305.70892	1.1	198.9	53
AQRVELQAQTTTLEQAIK	Q12934 [237-254]	2028.10839			51
QQIIHTTPPASIVTS	Q12934 [191-205]	1592.86425	2.6	197.4	50
SLAALQGLGERVA	Q12934 [40-52]	1284.72703	12.3	187.7	48
GELAGPEDALARQV	Q12934 [79-92]	1425.73323		200	48
QQIIHTTPPASIVT	Q12934 [191-204]	1505.83222	6.1	193.9	47
TTTLEQAIKSA	Q12934 [246-256]	1162.63139	6.5	193.5	45
LDAFQRLGEL	Q12934 [72-81]	1161.62625		200	45
GELAGPEDALAR	Q12934 [79-90]	1198.60624		200	42
ISILQQIIH	Q12934 [187-195]	1064.64626	12	188	42
FIETPIPLFT	Q12934 [324-333]	1177.65034	4.5	195.5	42
SLAALQGLGERVAAH	Q12934 [40-54]	1492.82305		200	40
KTVEVVESIEKIS	Q12934 [620-632]	1460.82065	2.4	197.6	39
QQIIHTTPPASIVTSGM	Q12934 [191-207]	1780.9262		200	38
IETPIPLFTQ	Q12934 [325-334]	1158.6405	0.4	199.6	37
RLQLEAQ	Q12934 [157-163]	857.48394	2.2	197.8	36
FIETPIPLFTQSHG	Q12934 [324-337]	1586.82132		200	35
NKEADEALLHNL	Q12934 [145-156]	1366.69612	6.4	193.6	34
LTEREVAAL	Q12934 [213-221]	1001.56259		200	34
RQLDAFQRL	Q12934 [70-78]	1146.63782	12.8	187.2	33
GMREEKLLTEREVAAL	Q12934 [206-221]	1844.98986		200	33
IETPIPLFTQSH	Q12934 [325-336]	1382.73144		200	33
LDAFQRL	Q12934 [72-78]	862.47813	13.3	186.7	33
AFQRLGELAGPE	Q12934 [74-85]	1287.66918	177.5	22.5	32
IETPIPLF	Q12934 [325-332]	929.53425		200	32
MREEKLLTEREVAAL	Q12934 [207-221]	1787.96839		200	31
VIVETMIGKTK	Q12934 [644-654]	1218.71262	200		31
FIETPIPL	Q12934 [324-331]	929.53425		200	31
FIETPIPLF	Q12934 [324-332]	1076.60266		200	30
QLDAFQRL	Q12934 [71-78]	990.53671		200	30
RIIEIEGNRLT	Q12934 [311-321]	1313.75357	9.2	190.8	28
AALQGLGERVAAHVQ	Q12934 [42-56]	1519.83395	39.6	160.4	27
REEKLLTEREVAAL	Q12934 [208-221]	1656.92791	6.1	193.9	27
LEEGREVLSHL	Q12934 [225-235]	1281.67974		200	27
RVELQAQTTTLEQAIKS	Q12934 [239-255]	1916.04473		200	27
SLAALQGLGERVAAHVQ	Q12934 [40-56]	1719.95004		200	26

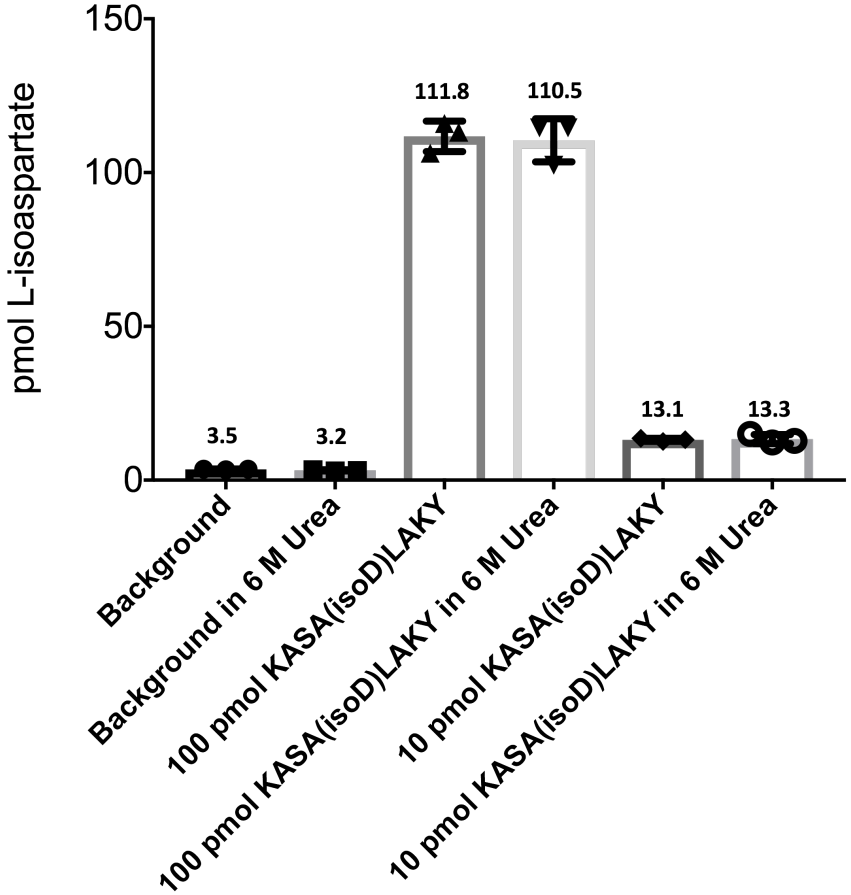
TABLE S7: Filensin peptides. Lens peptides were isolated from the gel filtration HMW fraction and identified as described in “Experimental Procedures.” All peptides with an ion score of 26 or higher are displayed with their corresponding theoretical mass and relative abundances in a 44-year-old and 76-year-old whole lens sample on a scale of 0-200.

Sequence	Positions in Master Proteins	Theo. MH+ [Da]	Abundances 44 year old	Abundances 76 year old	Ion Score
YILPQGEYPEYQRWM	P22914 [60-74]	1972.9262		200	45
IFEKGDFSGQ	P22914 [98-107]	1127.53677		200	43
GEYPEYQRWM	P22914 [65-74]	1358.5834	13.2	186.8	41

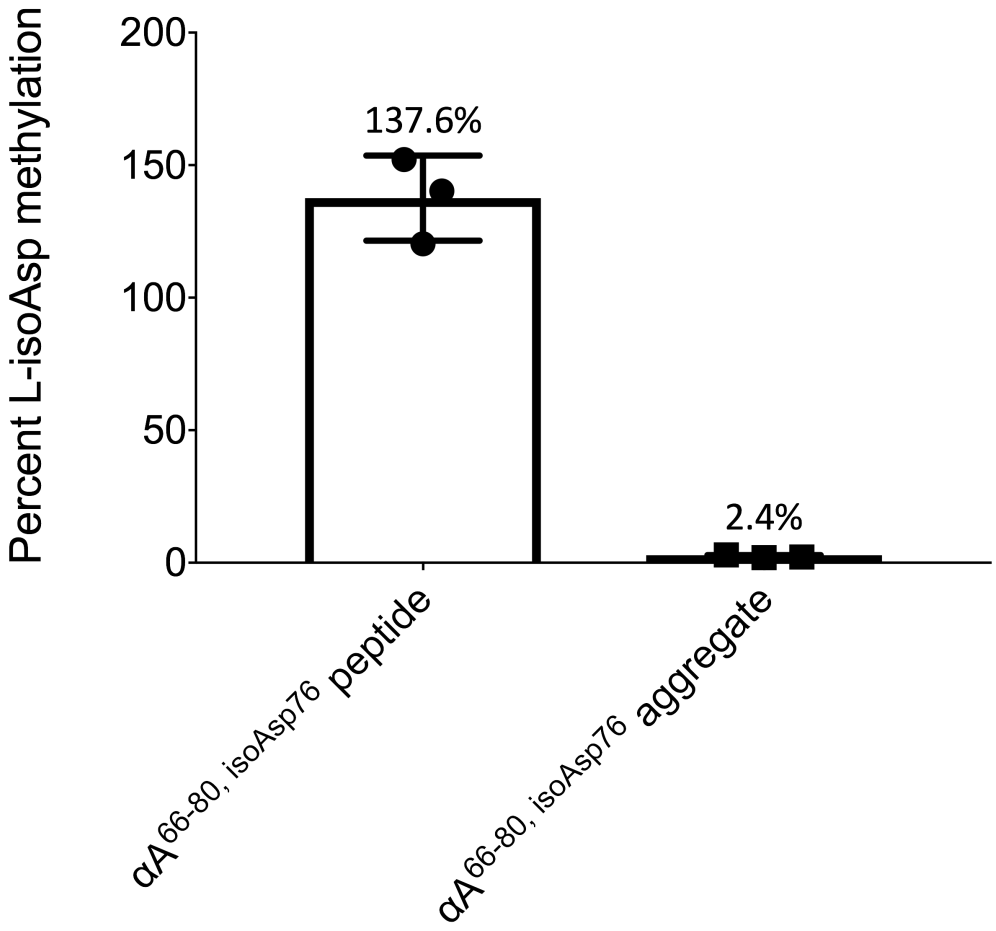
TABLE S8: γS crystallin peptides. Lens peptides were isolated from the gel filtration HMW fraction and identified as described in “Experimental Procedures.” All peptides with an ion score of 26 or higher are displayed with their corresponding theoretical mass and relative abundances in a 44-year-old and 76-year-old whole lens sample on a scale of 0-200.

Sequence	Positions in Master Proteins	Theo. MH+ [Da]	Abundances 44 year old	Abundances 76 year old	Ion Score
QQVGEAVLENARL	Q13515 [162-174]	1426.76487		200	60
WASSCQQVGEAVLENARL	Q13515 [157-174]	2017.976		200	57
KVIDEANLTKM	Q13515 [211-221]	1261.68205	6.3	193.7	56
TGLDDILETIRIQ	Q13515 [266-278]	1486.81115		200	56
QQVGEAVLENARLM	Q13515 [162-175]	1557.80535	4	196	55
RASWASSCQQVGEAVLENARL	Q13515 [154-174]	2332.14625		200	53
SSCQQVGEAVLENARLM	Q13515 [159-175]	1891.90006	2.1	197.9	51
RASWASSCQQVGEAVLENARLM	Q13515 [154-175]	2463.18674		200	50
KVIDEANLTKMDLESQIE	Q13515 [211-228]	2076.05291		200	48
SWASSCQQVGEAVLENARL	Q13515 [156-174]	2105.00803		200	47
WASSCQQVGEAVLENARLM	Q13515 [157-175]	2149.01648	0.2	199.8	45
QVGEAVLENARLM	Q13515 [163-175]	1429.74677		200	44
SCQQVGEAVLENARLM	Q13515 [160-175]	1804.86803	1.2	198.8	44
QVGEAVLENARL	Q13515 [163-174]	1298.70629	23.7	176.3	41
SVFLQGLRSSGL	Q13515 [82-93]	1263.70556	7	193	40
SLYKVIDEANLTKM	Q13515 [208-221]	1624.86147	12.4	187.6	39
YHALLD	Q13515 [405-410]	731.37227		200	30
GLRSSGLATVPAPGLER	Q13515 [87-103]	1680.93914	20.9	179.1	30
SWASSCQQVGEAVLENARLM	Q13515 [156-175]	2236.04851		200	30
HALLDREE	Q13515 [406-413]	982.49523	2.9	197.1	29
SCQQVGEAVLENARL	Q13515 [160-174]	1673.82754		200	29
YHALLDREE	Q13515 [405-413]	1145.55856	4.4	195.6	29
SVFLQGLRSSGLATVPAPGLERD	Q13515 [82-104]	2370.27758		200	29
HMSQTQEEKLAAAL	Q13515 [304-317]	1556.77372		200	28
SVFLQGLR	Q13515 [82-89]	919.53598	9	191	28
AEEEINSLY	Q13515 [202-210]	1067.48915		200	28
ALLDREESG	Q13515 [407-415]	989.48981		200	27
DILETIRIQ	Q13515 [270-278]	1100.631	15.2	184.8	27
GAVVGRL	Q13515 [364-370]	671.41989		200	26
SVFLQGLRSSGLA	Q13515 [82-94]	1334.74268	2.4	197.6	26
AVLENARLM	Q13515 [167-175]	1016.55573	11	189	26
SVFLQGLRSS	Q13515 [82-91]	1093.60003	31.3	168.7	26
SVFLQGLRSSGLATVPAPGLER	Q13515 [82-103]	2255.25064		200	26

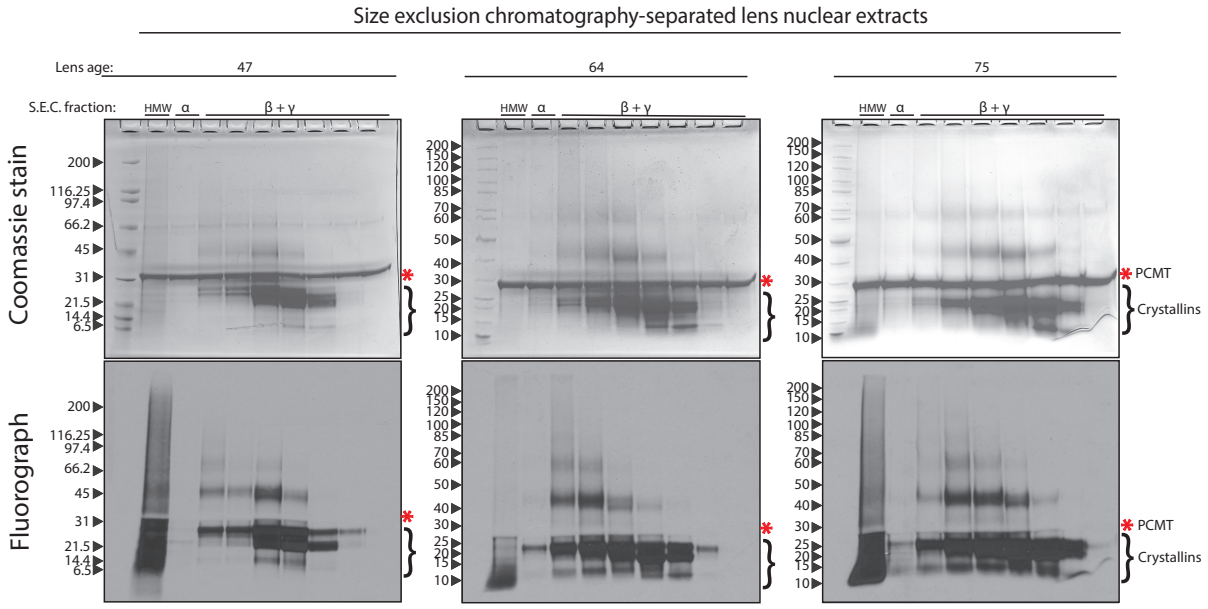
TABLE S9: Phakinin peptides. Lens peptides were isolated from the gel filtration HMW fraction and identified as described in “Experimental Procedures.” All peptides with an ion score of 26 or higher are displayed with their corresponding theoretical mass and relative abundances in a 44-year-old and 76-year-old whole lens sample on a scale of 0-200.



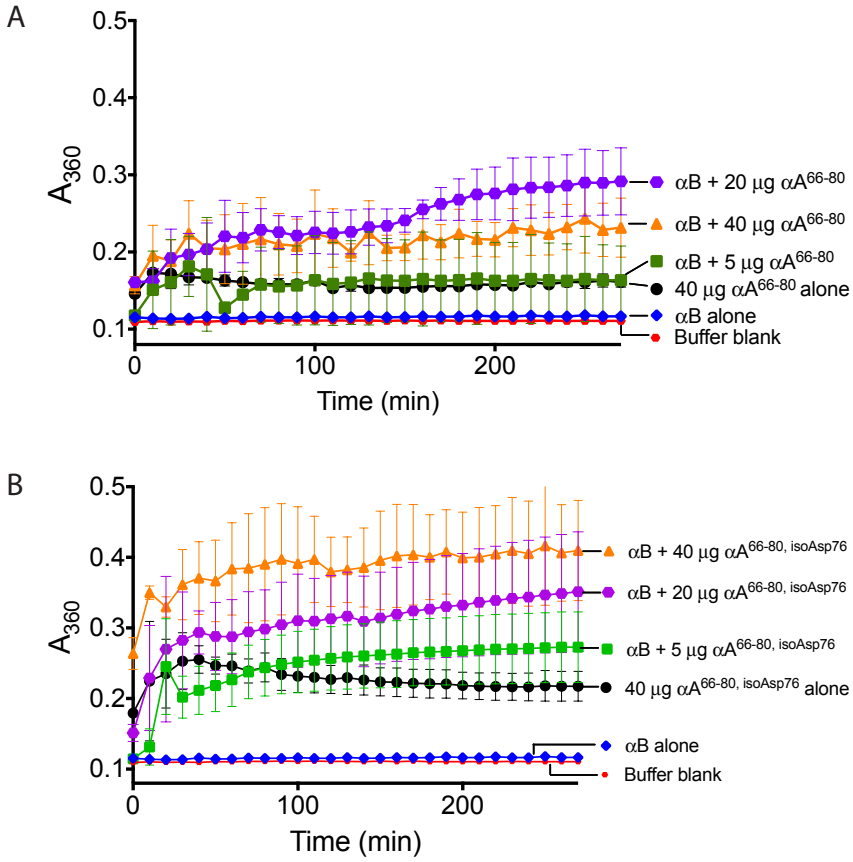
SUPPORTING FIG. 1: Exogenous PCMT1 labeling of L-isoAsp is not affected by urea. KASA(isoD)LAKY peptide was resuspended in 50 mM Tris-HCl, pH 7.9, 150 mM NaCl, and 6 M urea to match buffer conditions of re-solubilized water-insoluble lens extracts ("in 6 M urea" indicated on the x-axis), or in 50 mM Tris-HCl, pH 7.9, 150 mM NaCl buffer lacking urea to match buffer conditions of water-soluble lens extracts. These buffers alone were used in the "Background" reactions. Buffer alone, 100 pmol of KASA(isoD)LAKY, and 10 pmol of KASA(isoD)LAKY were analyzed for L-isoAsp content by PCMT1 methylation as described in the Experimental Procedures section. Reactions were performed in triplicate, and the average value is displayed above each column. Error bars represent the standard deviation.



Supporting Fig. 2: Exogenous PCMT1 cannot methylate aggregated αA$^{66\text{-}80,\ isoAsp76}$ peptide. Aggregates of αA$^{66\text{-}80,\ isoAsp76}$ were formed at 1 mg/mL in 50 mM Tris-HCl, pH 7.6, 150 mM NaCl (TBS). Soluble peptide solutions of αA$^{66\text{-}80,\ isoAsp76}$ were prepared by dissolving αA$^{66\text{-}80,\ isoAsp76}$ in TBS with 10% DMSO and filtering immediately prior to PCMT1 methylation assays, carried out as described in the Experimental Procedures section. Values of "Percent L-isoAsp methylation" were calculated by dividing the pmol of L-isoAsp methylated by the pmol of LisoAsp present in the reaction (70 pmol soluble peptide analyzed, 1000 pmol aggregate analyzed).



SUPPORTING FIG. 3: Analysis of L-isoAsp damage in the native size-exclusion fractions of human nuclear lens extracts by SDS-PAGE reveals the HMW gel filtration fraction contains significant amounts of LMW L-isoaspartyl-containing species. Fractions from size exclusion chromatography as in Fig. 4 were trichloroacetic acid-precipitated, resuspended in 100 mM Tris-HCl, pH 7.9, 150 mM NaCl, 0.1% SDS, 6 M Urea and L-isoAsp sites within the fractions were radiolabeled by PCMT1 and [^{3}H] AdoMet as described in the Fig. 3 legend. Labeled fractions were then separated by SDS-PAGE on a 4-12% gradient gel. The Coomassie-stained gel is shown in the upper panels; the fluorograph (one week exposure) is shown in the lower panels. In the experiment with the 47-year-old extract, molecular weight markers shown in Fig. 3 were used; in the 64- and 75-year-old-samples the molecular weight standards include synthetic polypeptides of the designated molecular weights (Thermo-Scientific, PageRuler Unstained Protein Ladder, catalog # 26614). The bands corresponding to PCMT1 in the Coomassie-stained gel are indicated by asterisks. The migration positions of intact crystallins are indicated by brackets on the right hand side of the gels. Labels above lanes represent the corresponding peaks from the size exclusion run (see Fig. 4). The age of lens sample in years is designated above the corresponding gel.



SUPPORTING FIG. 4: The $\alpha A^{66\text{-}80,\ isoAsp76}$ **peptide aggregates purified αB crystallin more than the native** $\alpha A^{66\text{-}80}$ **peptide.** αB crystallin (40 μg) was incubated with designated amounts of the $\alpha A^{66\text{-}80}$ (top panel) or $\alpha A^{66\text{-}80,\ isoAsp76}$ (bottom panel) peptide in 50 mM sodium phosphate, pH 7.0, 100 mM NaCl, and 10 mM phenanthroline as described in the "Experimental Procedures" section. Aggregation was monitored by light scattering at A_{360}. Symbols represent the mean and error bars represent the standard deviation of three technical replicates.

CHAPTER 4

Structure of amyloid-β (20-34) with Alzheimer's-associated isomerization at Asp23 reveals a distinct protofilament interface

Structure of amyloid-β (20-34) with Alzheimer's-associated isomerization at Asp23 reveals a novel protofibril interface

Rebeccah A. Warmack, David R. Boyer, Chih-Te Zee, Logan S. Richards, Michael R. Sawaya, Duilio Cascio, Tamir Gonen, David S. Eisenberg, and Steven G. Clarke

ABSTRACT:

Amyloid-β (Aβ) harbors numerous post-translational modifications (PTMs) that may affect Alzheimer's disease (AD) pathogenesis. Here we present the 1.1 Å resolution MicroED structure of an Aβ 20-34 fibril with and without the disease-associated PTM, L-isoaspartate, at position 23 (L-isoAsp23). Both wild-type and L-isoAsp23 protofilaments adopt β-helix-like folds with tightly packed cores, resembling the cores of full-length fibrillar Aβ structures, and both self-associate through two distinct interfaces. One of these is a unique Aβ interface strengthened by the isoaspartyl modification. Powder diffraction patterns suggest a similar structure may be adopted by protofilaments of an analogous segment containing the heritable Iowa mutation, Asp23Asn. Consistent with its early onset phenotype in patients, Asp23Asn accelerates aggregation of Aβ 20-34, as does the L-isoAsp23 modification. These structures suggest that the enhanced amyloidogenicity of the modified Aβ segments may also reduce the concentration required to achieve nucleation and therefore help spur the pathogenesis of AD.

INTRODUCTION

A prevalent theory for the biochemical basis of Alzheimer's disease is the amyloid cascade hypothesis, which describes the aggregation of the Aβ peptide into oligomer or fibrous structures that then trigger the formation of neurotoxic tau neurofibrillary tangles[1-3]. The Aβ peptide is subject to a number of posttranslational modifications (PTMs) that may affect its aggregation *in vivo*[4]. Specifically, Aβ phosphorylation (Ser8, Ser26), pyroglutamylation (Glu3, Glu11), nitration (Tyr10), and racemization/isomerization (Asp1, Asp7, Asp23, Ser26) have been shown *in vitro* to increase the aggregation propensity or neurotoxicity of the Aβ 1-42 peptide[5-11], while other modifications, such as dityrosine crosslinking (Tyr10), have been shown to increase the stability of the Aβ aggregates[12].

Isomerized products of aspartic acid residues perturb protein structure by rerouting the peptide backbone through the side chain β-carbonyl. This age-dependent modification introduces a methylene group within the polypeptide backbone and thus may have a significant effect on the structure of Aβ oligomers or fibrils[13-15]. Additionally, the isopeptide bond is resistant to degradation, potentially increasing the concentration of the isomerized Aβ form. Despite the presence of a repair enzyme in the brain, the L-isoaspartate (D-aspartate) *O*-methyltransferase (PCMT1) for L-isoaspartate, the isomerization of Aβ Asp1, Asp7, and Asp23 has been identified within AD brain parenchyma[16-17]. In the cases of the heritable early onset AD Iowa mutation (Asp23Asn), 25-65% of Asn23 residues have been shown to be isomerized in frontal lobe tissues[18], consistent with the increased rates of spontaneous deamidation/isomerization of asparagine relative to aspartate[19]. *In vitro* studies demonstrate that L-isoaspartate at Asp23 (L-isoAsp23) significantly accelerates Aβ 1-42 fibril formation, while L-isoAsp7 alone does not[11, 20]. Subsequent studies using peptides with multiple sites of isomerization showed only minor accelerated aggregation of the tri-isomerized species (1, 7, and 23), over the di-isomerized species (7 and 23)[18]. Taken together, these results suggest that amongst the known sites of

Asp isomerization in Aβ, L-isoAsp23 is primarily responsible for the increase in aggregation propensity *in vitro*.

Given the relevance of the isomerization of Asp23 to both sporadic and hereditary Iowa mutant forms of AD, we sought to discover the structural basis for its acceleration of fibril formation[10,17,18]. As a platform for evaluating this modification, we chose synthetically generated 15-mer peptides encoding residues 20-34 of the Aβ peptide ($A\beta^{20-34}$) with and without an L-isoAsp modification at position 23 and spanning the core of known Aβ fibril structures[21-29]. Challenged by the small size of crystals formed by this segment, we employed the cryoEM method MicroED to determine the structures. The structures of $A\beta^{20-34}$ and $A\beta^{20-34,\ isoAsp23}$, determined to 1.1 Å resolution by direct methods, reveal with atomic detail a conserved kinked β-helix-like-turns with complex features similar to those observed previously at lower resolution in the cores of fibrillar Aβ 1-42, as well as a distinct pair of protofilament interfaces. Our results suggest that the L-isoAsp23 residue facilitates the formation of a more stable form of this unique interface, promoting enhanced fiber formation and stability. The length of these peptide segments, four residues longer than any other crystallographically determined amyloid structures[30-33], is key in facilitating its complex fold - a conformation more representative of the full-length Aβ fibrils.

RESULTS

Fibril formation and characterization of $A\beta^{20-34}$ peptides

Six early-onset hereditary Alzheimer's mutations and two PTMs, including the isomerized Asp23, are localized in the Aβ 1-42 peptide to a region spanning six residues from Ala21 to Ser26 near the center of the peptide (**Fig. 1a, b**)[17,34-35]. The amyloid-forming propensity of segments in this region of Aβ was assessed using a computational method of predicting steric zippers by a threading protocol (ZipperDB[36]). This method highlights a region of Aβ from Asn28 to Gly37 with high aggregation propensity near the site of Asp23 isomerization (**Fig. 1b**). To characterize segments containing an isomerized Asp residue at position 23, we utilized synthetic 15 residue peptides spanning the Aβ residues 20-34 ($A\beta^{20-34}$) in which Asp23 was substituted with either an L-Asn residue (Iowa mutant; $A\beta^{20-34,\ Asp23Asn}$), or an L-isoAsp residue ($A\beta^{20-34,\ isoAsp23}$).

To evaluate the effect of these variations on this 15-residue segment of Aβ, we assayed its capacity to form fibrils as measured by light scattering at 340 nm (**Fig. 1c**). Both the peptide based on the Iowa mutant ($A\beta^{20-34,\ Asp23Asn}$) and the peptide based on L-isoAsp23 ($A\beta^{20-34,\ isoAsp23}$) demonstrated significantly enhanced fibril formation over that of $A\beta^{20-34}$, with the Iowa mutant peptide displaying the fastest initial rate of fibril formation (**Fig. 1c**). Fibers of the native peptide at this concentration (1.6 mM) were not observed by light scattering or electron microscopy. We further discovered that only 34% of these $A\beta^{20-34,\ isoAsp23}$ aggregates could be methylated by the L-isoAsp repair protein carboxyl methyltransferase (PCMT1) *in vitro* (**Supplementary Figure 1**). These data suggest that a majority of the L-isoAsp sites are occluded from the normal repair pathway once in the aggregate form.

To determine the ability of these modified forms to accelerate the aggregation of native peptide, seeding of 3.2 mM $A\beta^{20-34}$ was performed using 10 μM final concentrations of pre-aggregated seeds of $A\beta^{20-34}$, $A\beta^{20-34,\ Asp23Asn}$, and $A\beta^{20-34,\ isoAsp23}$ (**Fig. 1d**). The addition of each of the preformed aggregates caused significant acceleration in the onset of fiber formation. The

largest shift occurred with the native $A\beta^{20-34}$ peptide, followed by isomerized $A\beta^{20-34, isoAsp23}$ and $A\beta^{20-34, Asp23Asn}$. Powder diffraction performed on the final aggregates revealed nearly identical sets of reflections, suggesting that the three seeds have similar enough structures to template wild-type $A\beta^{20-34}$ aggregates whose diffraction resembles unseeded fibrils (**Supplementary Figure 2a**). Fibrillization experiments of full-length Aβ 1-40 with and without the L-isoAsp modification at residue 23 reveals that the isomerized species displays a shorter lag time, consistent with the results obtained with the corresponding $A\beta^{20-34}$ peptides (**Supplementary Figure 2b; Fig. 1c**). Thus, while the isomerized form may be only a minor component of the *in vivo* Aβ population, it aggregates at a faster rate and this species can cross-seed the native form efficiently *in vitro*.

In contrast to the results obtained with 1.6 mM $A\beta^{20-34}$, increasing the concentration to 3.2 mM $A\beta^{20-34}$ did yield aggregates ~77nm in width (**Fig. 1c, d**). Importantly, light scattering under these conditions for this native peptide is not detected until 3.5 h at the earliest, while shifts in light scattering for the 1.6 mM isomerized and mutated peptides were detected by 1.5 h and 0.5 h, respectively (**Fig. 1d**). Direct comparisons of formation rates were complicated by the insolubility of the $A\beta^{20-34, Asp23Asn}$ peptide at high concentrations, but the delayed onset of even the 3.2 mM $A\beta^{20-34}$ incubation compared to the 1.6 mM $A\beta^{20-34, Asp23Asn}$ and $A\beta^{20-34, isoAsp23}$ incubations also support the increased rates of aggregation of the mutated and isomerized peptides (**Fig. 1c, d**).

Fibrils of each segment were also investigated for their resistance to dissociation by dilution into increasing concentrations of sodium dodecyl sulfate (SDS) at 70 °C as measured by light scattering at 340 nm (**Fig. 2**). Fibrils of the native 15-residue Aβ segment appeared to partially dissolve upon dilution into the SDS-free buffer, although remaining aggregates were found by electron microscopy, but were completely dissolved upon incubation with 1% SDS and higher concentrations (**Fig. 2**). In contrast, the isomerized peptide showed increased resistance to dissolution compared to the native peptide, and still showed light scattering at a concentration

of 2% SDS, though no more aggregates were seen at 5% SDS (**Fig. 2b**). The fibrils of the Iowa mutant appeared to be largely unaffected by dilution even at the highest concentrations of SDS, with no significant changes observed in the levels of light scattering. However, the aggregates in 5% SDS seen by electron microscopy appeared to be less bundled than at lower concentrations (**Fig. 2b**). These results show that alterations of the structure at Asp23 strongly contribute to fibril formation and stability.

Crystallization and data collection of the Aβ$^{20-34}$ segments

To understand the atomic structural basis for changes in the properties of the isomerized peptide, we sought to crystallize it in the amyloid state. Vapor diffusion screening yielded no crystals large enough for analysis by conventional x-ray crystallography for either segment. Instead ordered nanocrystals of the native segment were obtained with continuous shaking at 1200 rpm, and ordered nanocrystals of the isomerized segment were generated with constant mixing using an acoustic resonant shaker[37-38] for analysis by microcrystal electron diffraction (MicroED[39-40]) as described in the Methods section. Nanocrystals obtained in varying buffer conditions were evaluated by morphology and diffraction via light and electron microscopy, respectively. Those formed under the most promising conditions were used as seeds for additional rounds of batch crystal formation. The optimal crystallization condition for the isomerized segment was 50 mM Tris, pH 7.6, 150 mM NaCl, and 1% DMSO for 48 h with 2% seeds. Crystals of the native segment grew in 50 mM Tris, pH 7.5, 150 mM NaCl, and 1% DMSO for 30 h without seeding. Isomerized crystal trials produced densely bundled nanocrystals that could not be disaggregated by sonication and freeze thawing. However, washing crystal solutions with a 0.75% (*w/v*) solution of β-octyl glucoside in TBS, pH 7.6 yielded a higher number of single crystals for subsequent data collection. Dilution one to one in buffer yielded sufficient single crystals of the native segment for data collection (**Fig. 3a, d**). Data were collected on a Thermo Fisher TALOS Arctica microscope operating at 200 kV using a bottom mount CetaD CMOS detector. Each Aβ$^{20-34}$ nanocrystal could be rotated continuously up to 140

degrees during data collection. A 1.1 Å resolution structure was obtained by direct methods for each segment as described in the Methods section; refinement statistics for the structures are shown in **Table 1**.

MicroED structures of $A\beta^{20-34}$ and $A\beta^{20-34,\ isoAsp23}$ segments

The structures of both the $A\beta^{20-34}$ and the $A\beta^{20-34,isoAsp23}$ protofilaments reveal parallel, in-register architectures in which individual peptide chains stack through backbone hydrogen bonds every 4.8 Å and 4.9 Å along the protofilament axis, respectively (**Fig. 3**). In cross-section, both protofilaments appear triangular owing to sharp turns (β-arches) at Gly25 and Gly29, which divide each chain into three short, straight segments (**Fig. 3b, e and Supplementary Figure 3a**). When compared with the structures in the protein databank, the three-sided $A\beta^{20-34,isoAsp23}$ structure aligns best with a β-helical antifreeze protein from *Marinomonas primoryensis,* but lacks linker regions between each stacked chain. We thus designate this amyloid motif as a β-helix-like turn (**Supplementary Figure 4**)[41-42]. At the central core of both $A\beta^{20-34}$ and the $A\beta^{20-34,isoAsp23}$ protofilaments are the buried side chains of Phe20, Ala21, Val24, Asn27, and Ile31 in a zipper-like "intraface" that is completely dry. The side chain of Asn27 further stabilizes the assembly by forming a ladder of hydrogen bonds (polar zipper) along the length of the protofilament[43] (**Supplementary Figure 3b**).

Each protofilament self-associates with neighboring protofilaments in the crystals through two distinct interfaces. Interface A in both structures resembles a canonical steric zipper – with intersheet distances of 8.3 Å and 9.1 Å for the native and isomerized, respectively (**Fig. 3b, e**). Both are lined by the hydrophobic side chains of Ala30, Ile32, and Leu34 that are related by 2_1 screw symmetry (steric zipper symmetry class 1[44]). Interface A is completely dry owing to a high S_c of 0.73 in the native and 0.62 in the isomerized. This interface buries approximately 130 $Å^2$ per chain in the native form, and 131 $Å^2$ in the isomerized form.

Unlike the dry steric zipper Interface A, six water molecules line the second $A\beta^{20-34}$ interface, which we designate the "L-Asp Interface B" (Fig. 3b, e). Here the protofilaments are also related by a 2-fold screw symmetry axis. Nearest this central axis, Gly25 and Ser26 contact their symmetry partners across the interface, separated by only 3.5 Å. Furthest from the axis, Asp23 and Lys28 from opposing protofilaments form charged pairs. In between these two regions is a solvent channel with the six ordered waters, yielding low shape complementarity (S_c=0.43) to this interface overall. In contrast, in the $A\beta^{20-34,isoAsp23}$ "L-isoAsp Interface B" the truncated side chain of the L-isoAsp23 residue no longer forms a charged pair with Lys28, and instead the isomerized protofilaments form a completely dry interface along the methylene group of L-isoAsp23, Val24, Gly25, and Ser26 with high surface complementarity (S_c=0.81; **Fig. 3e**). This interface is tightly mated over its entire surface with an average distance of 4.0 Å between the backbones. Interface B buries approximately 139 $Å^2$ and 122 $Å^2$ per chain for the native and isomerized forms, respectively. The exclusion of water molecules from the L-isoAsp interface B likely results in a favorable gain in entropy for the structure, and there are attractive van der Waals forces along the tightly mated residues L-isoAsp23-Ser26.

Powder diffraction studies of Aβ peptides

X-ray powder diffraction patterns revealed that the fibrils of $A\beta^{20-34}$ segments appear largely isomorphous, sharing major reflections at ~4.7, 10, 12.2,14, and 29-31 Å (**Fig. 4a, b**). The similarity among the powder diffraction patterns of $A\beta^{20-34,\ isoAsp23}$, $A\beta^{20-34,\ Asp23Asn}$, and $A\beta^{20-34}$ indicates that $A\beta^{20-34,\ Asp23Asn}$ mimics the structures of the native and isomerized segments. We modeled an L-Asn residue at position 23 of the $A\beta^{20-34,\ isoAsp23}$ structure to see if the native L-amino acid could be accommodated in the dry L-isoAsp Interface B (**Fig. 4c, right panel**). The L-Asn residue was integrated into the $A\beta^{20-34,\ isoAsp23}$ Interface B scaffold without significant clashes. However, this Asn model lacks a backbone hydrogen bond extending between the isoAsp23 amide carboxyl to the Val24 amide nitrogen of the adjacent protofilament that is present in our $A\beta^{20-34,\ isoAsp23}$ structure (**Fig. 4c**). The residue at site 23 has to adopt an allowed,

but unusual left-handed helical conformation to form the L-isoAsp interface B. Both the methylene of the isoAsp residue and the isoAsp23 to Val24 main chain hydrogen bond may help stabilize this structure. This backbone hydrogen bond is present in the native $A\beta^{20\text{-}34}$ structure (**Fig. 4c**). In this native structure, the Asp main chain adopts a more canonical β-sheet conformation, but the side chain protrudes towards the opposite protofilament, prohibiting a tight, dry interface along residues Asp23-Ser26 as in the L-isoAsp interface B.

The L-Asn side chain in the L-isoAsp Interface B model may be able to compensate for this loss by forming another ladder of hydrogen bonds along the protofilament axis (**Fig. 4c, right panel**). Thus, this second interface packing may be achievable for a $A\beta^{20\text{-}34,\ Asp23Asn}$ structure as shown in the L-isoAsp Interface B model, however, the X-ray fiber diffraction reveals that the native $A\beta^{20\text{-}34}$ and mutated $A\beta^{20\text{-}34,\ Asp23Asn}$ peptides share more similarities than the isomerized $A\beta^{20\text{-}34,\ isoAsp23}$ and the $A\beta^{20\text{-}34,\ Asp23Asn}$ peptide. Both the native and heritable Iowa mutant forms lack more defined peaks at 22.9, 24.7, 29.4, and 32.5, while both have more broad peaks at 30.9 Å (**Fig. 4a, b**). These similarities between the $A\beta^{20\text{-}34}$ and $A\beta^{20\text{-}34,\ Asp23Asn}$ fiber diffractions patterns, and the lack of a methylene group in the normal L-residues, may suggest that the Iowa mutant $A\beta^{20\text{-}34,\ Asp23Asn}$ peptide will assume a structure more similar to the native $A\beta^{20\text{-}34}$ structure, as modeled in **Fig. 4c (left panels)**. This model maintains the backbone hydrogen bond between Asn23 and Val24, the ordered core of the $A\beta^{20\text{-}34}$ structure, and allows for the additional polar zipper between stacked Asn23 residues. The added network of hydrogen bonds along the asparagine side chain may explain in part the increased fiber formation rates and stability of $A\beta^{20\text{-}34,\ Asp23Asn}$ against SDS and heat denaturation. While the isomorphous powder diffraction patterns seen between $A\beta^{20\text{-}34,\ isoAsp23}$, $A\beta^{20\text{-}34,\ Asp23Asn}$, and $A\beta^{20\text{-}34}$ do support the models in which $A\beta^{20\text{-}34,\ Asp23Asn}$ mimics the native and isomerized structures, it cannot be ruled out that $A\beta^{20\text{-}34,\ Asp23Asn}$ forms a distinct structure, perhaps lacking either the L-Asp or the L-isoAsp novel interface B, with the ordered core simply stabilized further by the Asn polar zipper.

Importantly, the powder diffraction of full-length Aβ and the shorter peptide segments all display cross-β patterns with strong reflections at ~4.7 and 9-10 Å (**Fig. 4a**), and the crystal structures of $A\beta^{20\text{-}34}$ and $A\beta^{20\text{-}34,\ isoAsp23}$ form parallel, in-register beta-sheets similar to other full-length Aβ structures. Thus we hypothesized that $A\beta^{20\text{-}34,\ isoAsp23}$ structure could form the core of a distinct isomerized Aβ polymorph. To visualize a potential full-length fiber with the $A\beta^{20\text{-}34,\ isoAsp23}$ structure as its core, we added the remaining residues of Aβ 1-42 onto the ends of the $A\beta^{20\text{-}34,\ isoAsp23}$ protofilaments, and energy minimized the entire model as described in the "Methods" section. The resulting model demonstrates that the remainder of the residues of Aβ 1-42 can be accommodated in a favorable conformation with the isomerized segment as a core with interface A or B as the primary interface (**Fig. 5**).

Comparison of segment structures to known Aβ structures

The structures presented here are the longest segments of an amyloid peptide determined by crystallography – four residues longer than the previous amyloid spines determined by MicroED[30-33]. This extension is significant due to the fact that as the number of residues in a segment grows, the packing of idealized β-strands in a lattice becomes more difficult owing to the strain created by the natural twist of the β-sheet/strand. This strain hypothesis is consistent with observations that as the number of residues in an amyloid segment grows, the crystals that can be grown are correspondingly smaller[45]. In the literature to date, the crystal structures of shorter segments of amyloid proteins have revealed that the dominant forces stabilizing protofilaments occur between different peptide chains[46]. In the native and modified $A\beta^{20\text{-}34}$ structures we are not only able to see interactions between protofilaments, such as the interfaces A and B, but we also see folding of the peptide to produce a β-helix-like turn with a hydrophobic core of interacting residues within the same chain.

While not all full-length native structures contain β-arches, such as the peptide dimer structure shown in Schmidt et al., 2015[47] (PDB code: 5AEF), all do include ordered cores involving steric zippers similar to those found in shorter amyloid peptide structures, and a

majority of the known Aβ structures do display β-helix-like turns as seen in the segment structures (**Fig. 6** and **Supplementary Figure 5**). The native $A\beta^{20-34}$ structure aligns well with a number of these full-length Aβ structures, and both the native and isomerized structures presented here have the lowest total atom RMSD with a structure of the Aβ Osaka mutant[29], E22Δ, at 2.741 Å and 2.963 Å, respectively. A tree representing the structural relationships between residues 20-34 of eight full length Aβ structures and our $A\beta^{20-34}$ structure based on total atom RMSD values shows that six of the eight structures contain turns about the Gly25 and Gly29 residues[21,23-24,26,28-29], which align well with interface B of our L-Asp $A\beta^{20-34}$ structure. Four[21,23,26,29] of these structures correspond to both the $A\beta^{20-34}$ segment structures with regards to the placement of charged residues Glu22, Asp23, and Lys28 outside the hydrophobic core; and yield total atom RMSD values of ~4 Å or lower with $A\beta^{20-34}$ (**Fig. 6** and **Supplementary Figure 5**). These strong overlaps between our segment structure and other full-length Aβ structures support the validity of this segment as an atomic resolution structure of an Aβ core. Importantly, in each of the full-length structures shown here, the putative interface B is accessible as a possible secondary nucleation site (**Fig. 6**). This interface is stabilized within our structures by the L-isoAsp modification which mates more tightly between protofilaments than the L-Asp interface B and excludes waters. Thus a full-length structural polymorph with this interface may be isolated more readily with this modification.

The increased structural complexity afforded by extending from 11 to 15 residues is appreciated best in comparing the crystal structures of $A\beta^{20-34}$ to the shorter Aβ 24-34 crystal structure, 5VOS[32] (**Fig. 7**). The four extra N-terminal residues in both native and modified $A\beta^{20-34}$ facilitate formation of kinks at Gly25 and Gly29, creating an internal core, whereas the Aβ 24-34 peptide assumes a linear β-strand. Despite Aβ 24-34 lacking these kinks, there is remarkable alignment between residues Gly29 to Leu 34 and interface A of the $A\beta^{20-34}$ crystals**,** yielding a total atom RMSD of 0.70 Å and 0.68 Å with the native and isomerized forms, respectively (**Fig. 7**). An inhibitor was previously developed to the human islet amyloid polypeptide (hIAPP) steric

zipper interface analogous to this interface of the 5VOS Aβ 24-34 segment, and was shown to be effective against fibril formation of both hIAPP and full length Aβ[32]. Given the striking alignment between our $A\beta^{20-34}$ interface A and the 5VOS Gly29-Leu34 segment, as well as the distinct lack of modifications and mutations in the region of Asn27-Gly33, this interface may be an ideal scaffold for Aβ inhibitor design in both its homotypic steric zipper form as shown here, or in the heterotypic zippers displayed in many of the full-length Aβ structures (**Fig. 6**).

DISCUSSION

The typical age of onset for sporadic Alzheimer's disease is after 65 years, suggesting that slow spontaneous processes such as the accumulation of age-dependent PTMs in Aβ may be contributing factors to aggregation and toxicity[4]. The spontaneous isomerization of aspartate (isoAsp) has been identified at all three aspartate residues within the Aβ 1-42 peptide – 1, 7, and 23. However, immunohistochemical studies have shown that while native Aβ and isoAsp7 Aβ are present in senile plaques from four non-disease patient controls, isoAsp23 Aβ was identified only in one of the four non-disease patient controls, as well as in the senile plaques from all AD patient samples, indicating that the isoAsp23 may be more specifically associated with AD pathology than native Aβ and the L-isoAsp7 form[10]. This implied pathogenicity of isoAsp23 correlates with *in vitro* studies, which have demonstrated accelerated amyloid formation of the isoAsp23 Aβ 1-40 and 1-42 peptides compared to native Aβ[10,11,17-18,20]. These results suggest that the change in the structure of Aβ accompanying isomerization at Asp23 may represent a route to the pathogenesis of AD.

In this work we present the 1.1 Å structures of segments spanning residues 20-34 of the Aβ peptide containing either an Asp or an isoAsp residue at site 23. These 15-residue segments, crystallized at physiological pH, maintain a topology seen in the core of Aβ fibrils, a β-helix-like turn (**Fig. 6**). The length of these peptides facilitates their similar overall fold to previous wild-type Aβ fibril structures and demonstrates that amyloid cores are rigid and ordered enough to form crystals. These structures reveal a previously unseen protofilament interface (B)

involving residues Asp23-Lys28 in the native structure, and residues L-isoAsp23-Ser26 in the isomerized structure. The native interface (L-Asp interface B) has low surface complementarity and contains six water molecules encased between charged residue pairs Asp23 and Lys28 on opposing sheets. In contrast, the isomerized interface (L-isoAsp interface B) is a dry tightly mated sheet with high surface complementarity. Our data suggests that the changes in the structure along this interface, namely the exclusion of water molecules and van der Waals attractive forces associated with the high S_c, are likely responsible in part for the increases in fiber formation rate and stability observed for the isomerized peptide. The modified interface may provide a better site for secondary nucleation of amyloid formation resulting in the observed enhancements in aggregation. However, it cannot be ruled out from the data presented here that the flexibility imparted by the methylene group of the L-isoAsp residue promotes amyloid formation by allowing an ordered nucleus for primary nucleation to form at a faster rate than the native peptide.

Our models of Asn23 in the $A\beta^{20\text{-}34,\ isoAsp23}$ "L-isoAsp interface B" indicate that the completely dry interface may be possible for native residues (**Fig. 4**). However, the native $A\beta^{20\text{-}34}$ structure did not preferentially adopt this interface, and instead forms a hydrated L-Asp Interface B. Similar to our $A\beta^{20\text{-}34}$ peptide structure, alignments of previous Aβ structures onto the $A\beta^{20\text{-}34}$ and $A\beta^{20\text{-}34,\ isoAsp23}$ protofilaments show the native Asp23 side chain carboxyl group protruding into the putative interface B region (**Supplementary Figure 5**). The hereditary Iowa mutant NMR structure (**Fig. 6 and Supplementary Figure 5** (PDB: 2MPZ[22])), kinks at Gly25 and Asn27, rather than Gly25 and Gly29, and thus there is no equivalent interface A. Yet, our preparations of crystals in TBS of the $A\beta^{20\text{-}34}$, $A\beta^{20\text{-}34,\ isoAsp23}$, and $A\beta^{20\text{-}34,\ Asp23Asn}$ constructs appear nearly identical by X-ray powder diffraction, suggesting the structure of an Iowa mutant protofilament would resemble the native and isomerized structures presented here (**Fig. 4**), barring minor differences due to packing polymorphisms or different environmental conditions.

It is clear that both the isomerization and Iowa mutation at residue 23 accelerate aggregation and increase stability of Aβ fibrils. Our structures of $A\beta^{20-34}$ and $A\beta^{20-34,\ isoAsp23}$ reveal a potential mechanism for the increases in fiber formation rate and fiber stability within the isoAsp23 form: the addition of a completely dry interface with high surface complementarity. This analysis leads to the hypothesis that the Asp23 isomerization *in vivo* could lead to the accelerated formation of Aβ fibrils, thereby contributing to the aggregation of Aβ and AD pathology. The hereditary Iowa mutation Asp23Asn may work in a similar manner either by forming the same fold as the isomerized Asp23, or since Asn undergoes isomerization more rapidly relative to Asp, it may also produce an isomerized Aβ with accelerated aggregation and increased stability. The isomerized structure may also provide insight into the mechanisms behind the A21G, E22G, and E22Δ hereditary mutations that introduce flexibility into the same region of the backbone. Importantly, we have also found that the only known repair pathway for L-isoAsp, the enzyme PCMT1, is unable to fully methylate and repair aggregates of $A\beta^{20-34,\ isoAsp23}$ *in vitro*, thus once the modified aggregates have formed *in vivo* they may be difficult to repair and clear (**Supplementary Figure 1**).

Recent structures of tau isolated from Alzheimer's disease patients have revealed distinct structural polymorphs[48]. Both the paired helical filaments (PHFs) and the straight filaments (SFs) of tau display β-arches in their sheets, which is a feature also shared by the native and isomerized $A\beta^{20-34}$ structures (**Fig. 6 and Supplementary Figure 4**). This similarity not only suggests that our structure's β-helix-like turn may be a common amyloid motif, but it also identifies a potential cross-seeding site between Aβ and the tau protein of Alzheimer's disease. This discovery emphasizes the need for atomic-resolution structures of disease-associated amyloid, as these core segments are critical for structure-based drug design and protein prediction efforts[49-52]. These crystal structures can be used in conjunction with full-length cryo-EM structures to obtain a high-resolution view of the interactions mediating amyloid filament formation[53]. High-resolution structures are also valuable when looking at the effect

PTMs may have on amyloid structure as seen here and elsewhere[54]. Therefore, the combination of increasing peptide length and high resolution makes the $A\beta^{20-34}$ and $A\beta^{20-34,\ isoAsp23}$ structures an important step forward for the structural characterization of amyloid proteins and their role in disease.

METHODS:

***Materials*:** $A\beta^{20-34}$ peptides corresponding to the human sequence were purchased from and validated by Genscript at a purity of 98% or higher as the trifluoroacetic acid (TFA) salt and were stored at -20 °C. Peptides were validated by electrospray ionization (ESI) mass spectrometry performed by Genscript. Aβ 1-42 was purchased from Bachem Americas, Inc. (Catalog #, H-1368).

Aggregation of $A\beta^{20-34}$ peptides for fibril formation rates: Peptides were dissolved at 1.6 mM in 50 mM Tris-HCl, pH 7.6, 150 mM NaCl (TBS) with 2.5% DMSO unless otherwise designated in the figure legend. Peptides solutions were filtered through 0.22 µm cellulose acetate Costar Spin-X centrifuge tube filters (Corning Inc., product #8161). Filtered peptide solutions in a final volume of 100 µL/well in a 96 well plate (Fisherbrand, 12565501) were read at 340 nm in a Varioskan plate reader at 37 °C with continuous shaking at 1200 rpm. Readings were recorded every 15 min.

Seeding of $A\beta^{20-34}$ segment

Seeds were formed shaking continuously on an acoustic resonant shaker at 37 °C at a frequency setting of 37^{33-34}. Seeds of $A\beta^{20-34}$ were formed at 5 mg/ml in 50 mM Tris-HCl, pH 7.6, 150 mM NaCl (TBS) with 1% DMSO; seeds of $A\beta^{20-34,\ isoAsp23}$ were formed at 2.5 mg/ml in 50 mM Tris-HCl, pH 7.6, 150 mM NaCl (TBS) with 1% DMSO; and seeds of $A\beta^{20-34,\ Asp23Asn}$ were formed at 2.5 mg/ml in 100 mM Tris-HCl, pH 7.5, 10% isopropanol, and 200 mM sodium acetate. All seeds were diluted to 200 µM stocks and 5 µL were added to 3.2 mM $A\beta^{20-34}$ in a final volume of 100 µL. Not all wells of the unseeded 3.2 mM $A\beta^{20-34}$ condition aggregated within the time course of this assay (**Fid. 1d**). Solutions were read in a 96 well plate at 340 nm in a Varioskan

plate reader at 37 °C with continuous shaking at 1200 rpm. Readings were recorded every 15 min.

Synthesis and purification of native Aβ 1-40

The syntheses of Aβ (1-40) WT and Aβ (1-40) IsoAsp23 were completed in a CEM Liberty BlueTM Microwave Peptide Synthesizer. The crude peptides were purified using an Interchim puriFlash® 4125 Preparative Liquid Chromatography System.

The purified Aβ (1-40) WT has an estimated purity of 93% by HPLC and was characterized by ESI-MS via direct injection into a Q-Exactive™ Plus Hybrid Quadrupole-Orbitrap TM Mass Spectrometer. The calculated average mass for $C_{194}H_{295}N_{53}O_{58}S$: 4327.148 g/mol, m/z calculated: [M+3H] 3+ = 1443.39; [M+4H] 4+ = 1082.79; [M+5H] 5+ = 866.44; [M+6H] 6+ = 722.20. Observed: 1443.3913; 1082.7955; 866.4374; 722.1991. The purified Aβ (1-40) IsoAsp23 has an estimated purity of 97% by HPLC and was characterized by ESI-MS via direct injection into a Q-Exactive™ Plus Hybrid Quadrupole-Orbitrap TM Mass Spectrometer. The calculated average mass for $C_{194}H_{295}N_{53}O_{58}S$: 4327.148 g/mol, m/z calculated: [M+3H]3+ = 1443.39; [M+4H]4+ = 1082.79; [M+5H]5+ = 866.44. Observed: 1443.3929; 1082.7973; 866.4385.

Crystallization of the segments: $A\beta^{20-34}$ was resuspended at a concentration of 3.2 mM in 50 mM Tris-HCl, pH 7.5, 150 mM NaCl (TBS) with 1% DMSO in a final volume of 100 µL. The peptide solution was then shaken continuously for 30 h at 1200 rpm at 37 °C. $A\beta^{20-34,isoAsp23}$ was resuspended at a concentration of 1.6 mM in 50 mM Tris-HCl, pH 7.6, 150 mM NaCl (TBS) with 1% DMSO in a final volume of 200 µL. The filtered peptide solution was then shaken for two days on an acoustic resonant shaker at 37 °C at a frequency setting of 37^{24-25}. 4 µL of this suspension was then used to seed 196 µL of a second peptide solution (1.6 mM) as a two percent seed on the acoustic resonant shaker at 37 °C. Crystals were obtained within 48 hours. The presence of crystals was verified by electron microscopy, using a standard holder, with no

negative stain. Crystals of the native and isomerized segments were on average ~77 and ~71 nm in width, respectively, and were typically more than 2 μm in length.

MicroED sample preparation: Quantifoil R1.2/1.3 cryo-EM grids (Electron Microscopy Sciences, product # Q325CR1.3) were glow discharged for 30 s on either site, and 1.5 μL of a 1:1 dilution of $A\beta^{20-34}$ crystals in 50 mM Tris-HCl, pH 7.5, 150 mM NaCl (TBS) with 1% DMSO was pipetted on both sides. 20 μL of $A\beta^{20-34,isoAsp23}$ crystal suspensions were spun down at 5,000g for 5 min, the supernatant was removed and pelleted crystals were resuspended in 50 μL TBS + 0.75% (*w/v*) β-octyl-glucoside (VWR, P-1110), and rotated at 4 °C for 1 hour. These detergent treated crystals were then spun down a second time. Pelleted crystals were resuspended in 50 μL water. 1.5 μL of the washed crystal solution was then applied to both sides of a glow discharged Quantifoil R1.2/1.3 cryo-EM grid (Electron Microscopy Sciences, product # Q325CR1.3). All grids were plunge frozen into supercooled ethane using a Vitrobot Mark 4 instrument.

MicroED Data Collection and Processing

MicroED data was collected in a manner similar to previous studies[42]. Briefly, plunge-frozen grids were transferred to an FEI Talos Arctica electron microscope and diffraction data were collected using a bottom-mount CetaD 16M CMOS camera with a sensor size of 4,096 x 4,096 pixels, each 14 x 14 μm. Diffraction patterns were recorded by operating the detector in continuous mode with 2 × 2 pixel binning, producing data sets with frames 2,048 × 2,048 pixels in size. The exposure rate was set to <0.01 e^-/A^2/second. The exposure time per frame was set at 3 seconds while the rotation speed was set to was set to 0.3 deg/s resulting in a final oscillation range of 0.9 deg/exposure for the $A\beta^{20-34}$ data collection, and to 0.443 deg/s resulting in a final oscillation range of 1.329 deg/exposure for the $A\beta^{20-34,isoAsp23}$ data collection. This rotation rate was optimized to allow a maximum amount of reciprocal space to be sampled before crystal decay was observed while also slow enough to prevent overlapping diffraction spots in the diffraction images. Diffraction movies typically covered a 50 to 140 deg wedge of

reciprocal space and were taken of crystals randomly orientated on the grid with respect to the incident beam. These crystals had a highly preferred orientation on the grid, resulting in a systematic missing cone and hence lower completeness along the c* axis; however, this did not preclude structure determination, with a high overall completeness of over 80% for both structures (see **Table 1**).

Structure Determination

Diffraction datasets were converted to SMV format to be compatible with X-ray data processing software[55]. Data were indexed and integrated using XDS[56]. The parameters controlling the raster size during indexing and integration were optimized to reduce contributions by background and to exclude intensities that conform poorly to the lattice determined during indexing. The number of diffraction images used per crystal was aggressively pruned to maximize I/sigma. The resulting outputs from XDS were sorted and merged in XSCALE. To produce a final merged dataset, partial datasets were selected based on their effects on the Rmerge values. In total, for the $A\beta^{20-34}$ structure, ten partial datasets, containing 404 diffraction images, were merged to produce a final dataset with high completeness up to 1.1 Å. An *ab initio* solution was achieved using SHELXD[57]. In total, for the $A\beta^{20-34,\ isoAsp23}$ structure, five partial datasets, containing 159 diffraction images, were merged to produce a final dataset with high completeness up to 1.1 Å and an *ab initio* solution was also achieved using SHELXD. The phases obtained from both $A\beta^{20-34}$ coordinates produced by SHELX were used to generate maps of sufficient quality for subsequent model building in Coot[58]. The resulting models were refined with Phenix[59], using electron scattering form factors, against the measured data.

Powder Diffraction Sample Preparation and Data Collection: Designated aggregates of Aβ 1-42 and $A\beta^{20-34}$ peptides were prepared in buffers as described in the figure legends. Aggregates were spun at 20,000g for 5 min. The pellet was resuspended in water and re-spun. Pelleted fibrils were resuspended in 5 uL water and pipetted between two facing glass rods that

were 2 mm apart and allowed to dry overnight at room temperature. These glass rods with ordered fibrils were secured to a brass pin and mounted for diffraction at room temperature using 1.54 Å X-rays produced by a Rigaku FRE+ rotating anode generator equipped with an HTC imaging plate. Patterns were collected at a distance of 200 mm and analyzed using the ADXV software package[60].

Sodium dodecyl sulfate (SDS) dissolution of aggregates: Aggregates of $A\beta^{20-34}$, $A\beta^{20-34,\ isoAsp23}$, and $A\beta^{20-34,\ Asp23Asn}$ were all prepared in TBS, with 1%, 2.5%, and 2.5% DMSO, respectively. Both $A\beta^{20-34,\ isoAsp23}$ and $A\beta^{20-34,\ Asp23Asn}$ were prepared at a peptide concentration of 2.5 mg/ml, while $A\beta^{20-34}$ was prepared at 5 mg/ml, shaking at 1200 rpm at 25 °C. The $A\beta^{20-34}$ was diluted to 2.5 mg/ml prior to the denaturation assay. Suspensions of $A\beta^{20-34}$ aggregates were diluted 1:1 in 2%, 3%, 4%, and 10% SDS stocks in TBS, and heated for 15 min at 70 °C in a PTC-100 Peltier thermal cycler as described by Guenther et al., 2018[54]. Measurements at 340 nm were recorded on a Nanodrop 2000 instrument. 2 uL of each solution was analyzed by electron microscopy for remaining aggregates on glow discharged Formvar/Carbon 400 mesh, Copper grids (Ted Pella, catalog # 01754-F).

Analysis of S_a and surface S_c in $A\beta^{20-34}$ structures: The structures of $A\beta^{20-34}$ and $A\beta^{20-34,\ isoAsp23}$ were used to measure buried surface area (S_a) and (S_c) from an assembly consisting of 2 sheets generated by translational symmetry each consisting of ten stacked β-strands. S_a was calculated as an average of the buried surface area per chain, and the difference between the sum of the solvent accessible surface area of the two sheets and the solvent accessible surface area of the entire complex, divided by the total number of strands in both sheets using the CCP4 suite.

Modeling modified and full-length Aβ and RMSD calculations: Residues 1-42 of Aβ were modeled onto the N- and C-termini of the $A\beta^{20-34,\ isoAsp23}$ structure and alternate residues were mutated in coot and the resulting structures were energy minimized in the Crystallography & NMR System (CNS)[61].

Distance matrices for RMSD relationships between $A\beta^{20\text{-}34,\ isoAsp23}$ and residues 20-34 from native structures were generated in the LSQKAB program of CCP4, and resulting matrices were used to generate the tree shown in Figure 4.

Data availability: Atomic coordinates and structure factors for the $A\beta^{20\text{-}34}$ structure have been deposited in the Protein Data Bank under accession code 6OIZ. The map for this structure has been deposited in the EMDB with accession code EMD-20082. Atomic coordinates and structure factors for the $A\beta^{20\text{-}34,\ isoAsp23}$ structure have been deposited in the Protein Data Bank under accession code 6NB9. The map for this structure has been deposited in the EMDB with accession code EMD-0405. The source data underlying Figs 1c, 1d, 2, 4b, Supplementary Figures 1 and 2 are provided as a Source Data file.

Acknowledgements:

We thank Dr. Jose Rodriguez for his advice throughout the course of this work. This work was supported in part by grants from the National Science Foundation (MCB-1714569 to S.G.C. and MCB-1616265 to D.S.E.), the National Institutes of Health (AG 054022 to D.S.E.), and the Howard Hughes Medical Institute (to D.S.E. and T.G.). R.A.W. was supported by a USPHS National Research Service Award GM007185 and a UCLA Dissertation Year Fellowship and Pauley Fellowship Award. D.R.B. was supported by a National Science Foundation Graduate Research Fellowship. L.R. was supported by a USPHS National Research Service Award 5T32GM00849. Work in the S.G.C. laboratory was also supported by funds from the Elizabeth and Thomas Plott Chair in Gerontology of the UCLA Longevity Center, a grant from the Life Extension Foundation, and funds from a UCLA Faculty Research Award. We acknowledge the use of instruments at the Electron Imaging Center for Nanomachines supported by UCLA and by instrumentation grants from NIH (1S10RR23057 and 1U24GM116792) and NSF (DBI-1338135). We thank Michael Collazo at the UCLA-DOE Macromolecular Crystallization Core Technology Center for crystallization support. This work was supported in part by NIGMS grant No. R35GM128867 awarded to J.A. Rodriguez. The mass analysis of Aβ(1-40) wild-type and

IsoAsp23 was supported by the National Institutes of Health under instrumentation grant 1S10OD016387-01.

REFERENCES

1. Selkoe, D.J. The molecular pathology of Alzheimer's disease. *Neuron*. **6**, 487-498 (1991).
2. Hardy, J.A., Higgins, G.A. Alzheimer's disease: the amyloid cascade hypothesis. *Science*. **256**, 184-185 (1992).
3. Selkoe, D.J. and Hardy, J. The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Mol Med*. **8,** 595-608 (2016).
4. Roher, A.E., Kokjohn, T.A., Clarke, S.G., Sierks, M.R., Maarouf, C.L., Serrano, G.E., Sabbagh, M.S., Beach, T.G. APP/Aβ structural diversity and Alzheimer's disease pathogenesis. *Neurochem Int*. **110,** 1-13 (2017).
5. Kummer, M.P., Heneka, M.T. Truncated and modified amyloid-beta species. *Alzheimers Res Ther*. **6,** 28 (2014).
6. Rezaei-Ghaleh, N., Amininasab, M., Kumar, S., Walter, J., Zweckstetter, M. Phosphorylation modifies the molecular stability of β-amyloid-deposits. *Nat Comm*. **7,** 11359 (2016).
7. Kumar, S., Wirths, O., Stüber, K., Wunderlich, P., Koch, P., Theil, S., Rezaei-Ghaleh, N., Zweckstetter, M., Bayer, T.A., Brüstle, O., Thal, D.R., Walter, J. Phosphorylation of the amyloid β-peptide at Ser26 stabilizes oligomeric assembly and increases neurotoxicity. *Acta Neuropathol*. **131,** 525-537 (2016).
8. Mandler, M., Walker, L., Santic, R., Hanson, P., Upadhaya, A.R., Colloby, S.J., Morris, C.M., Thal, D.R., Thomas, A.J., Schneeberger, A., Attems, J. Pyroglutamylated amyloid-β is associated with hyperphosphorylated tau and severity of Alzheimer's disease. *Acta Neuropathol.* **128,** 67-79 (2014).

9. Kummer, M.P., Hermes, M., Delekarte, A., Hammerschmidt, T., Kumar, S., Terwel, D., Walter, J., Pape, H.C., König, S., Roeber, S., Jessen, F., Klockgether, T., Korte, M., Heneka, M.T.. Nitration of tyrosine 10 critically enhances amyloid-β aggregation and plaque formation. *Neuron.* **71,** 833-844 (2011).
10. Shimizu, T., Fukuda, H., Murayama, S., Izumiyama, N., Shirasawa, T. Isoaspartate formation at position 23 of amyloid beta peptide enhanced fibril formation and deposited onto senile plaques and vascular amyloids in Alzheimer's disease. *J Neurosci Res.* **70,** 451-461 (2002).
11. Fukuda, H., Shimizu, T., Nakajima, M., Mori, H., Shirasawa, T. Synthesis, aggregation, and neurotoxicity of the Alzheimer's Abeta1-42 amyloid peptide and its isoaspartyl isomers. *Bioorg Med Chem Lett.* **9,** 953-956 (1999).
12. Al-Hilaly, Y.K., Williams, T.L., Stewart-Parker, M., Ford, L., Skaria, E., Cole, M., Bucher, W.G., Morris, K.L., Sada, A.A., Thorpe, J.R., Serpell, L.C. A central role for dityrosine crosslinking of Amyloid-β in Alzheimer's disease. *Acta Neuropathol Com*. **1**, 1-17 (2013).
13. Geiger, T., Clarke, S.G. Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linke reactions that contribute to protein degradation. *J Biol Chem.* **262,** 785-794 (1987).
14. Fujii, N., Takata, T., Fujii, N. Quantitative analysis of isomeric (l-α-, l-β-, D-α-, D-β-_ aspartyl residues in proteins from elderly donors. *J Pharm Biomed Anal.* **10,** 25-33 (2015).
15. Fujii, N., Takata, T., Fujii, N., Aki, K., Sakaue, H. D-Amino acids in protein: The mirror of life as a molecular index of aging. *Biochim Biophys Acta Proteins Proteom.* **1866,** 840-847 (2018).
16. Roher, A.E., Lowenson, J.D., Clarke, S., Woods, A.S., Cotter, R.J., Gowing, E., Ball, M.J. beta-Amyloid-(1-42) is a major component of cerebrovascular amyloid deposits:

implications for the pathology of Alzheimer disease. *Proc Natl Acad Sci USA*. **90,** 10836-10840 (1993).

17. Shimizu, T., Watanabe, A., Ogawara, M., Mori, H., Shirasawa, T. Isoaspartate formation and neurodegeneration in Alzheimer's disease. *Arch Biochem Biophys.* **381,** 225-234 (2000).
18. Tomikodoro, Y., Rostagno, A., Neubert, T.A., Lu, Y., Rebeck, G.W., Frangione, B., Greenberg, S.M., Ghiso, J. Iowa variant of familial Alzheimer's disease: accumulation of postranslationally modified AbetaD23N in parenchymal and cerebrovascular amyloid desposits. *Am J Pathol.* **176,** 1841-1854 (2010).
19. Radkiewizc, J.L., Zipse, H., Clarke, S., Houk, K.N. Neighboring Side Chain Effects on Asparaginyl and Aspartyl Degradation: An Ab Initios Study of the Relationship between Peptide Conformation and Backbone NH Acidity. *J. Am Chem Soc*. **123,** 3499-3506 (2001).
20. Fossati, S., Todd, K., Sotolongo, K., Ghiso, J., Rostagno, A. Differential contribution of isoaspartate post-translational modifications to the fibrillization and toxic properties of amyloid-β and the Asn23 Iowa mutation. *Biochem J*. **456,** 347-360 (2013).
21. Wälti, M.A., Ravotti, F., Arai, H., Glabe, C.G., Wall, J.S., Böckmann, A., Güntert, P., Meier, B.H., Riek, R. Atomic-resolution structure of a disease-relevant Aβ(1-42) amyloid fibril. *Proc Natl Acad Sci USA.* **113,** E4976-4984 (2016).
22. Sgourakis, N.G., Yau, W.M., Qiang, W. Modeling an in-register, parallel "iowa" aβ fibril structure using solide-state NMR data from labeled samples with rosetta. *Structure.* **23,** 216-227 (2015).
23. Xiao, Y., Ma, B., McElheny, D., Parthasarathy, S., Long, F., Hoshi, M., Nussinov, R., Ishii, Y. Aβ(1-42) fibril structure illuminates self-recognition and replication of amyloid in Alzheimer's disease. *Nat Struct Mol Biol.* **22,** 499-505 (2015).

24. Paravastu, A.K., Leapman, R.D., Yau, W.M., Tycko, R. Molecular structural basis for polymorphism in Alzheimer's beta-amyloid fibrils. *Proc Natl Acad Sci USA.* **105,** 18349-18354 (2008).

25. Lührs, T., Ritter, C., Adrian, M., Riek-Loher, D., Bohrmann, B., Döbeli, H., Schubert, D., Riek, R. 3D structure of Alzheimer's amyloid-beta(1-42) fibrils. *Proc Natl Acad Sci USA.* **102,** 17342-17347 (2005).

26. Colvin, M.T., Silvers, R., Ni, Q.Z., Can, T.V., Sergeyev, I., Rosay, M., Donovan, K.J., Michael, B., Wall, J., Linse, S., Griffin, R.G., Atomic Resolution Structure of Monomorphic Aβ42 Amyloid Fibrils. *J Am Chem Soc.* **138,** 9663-9674 (2016).

27. Lu, J.X., Qiang, W., Yau, W.M., Schwieters, C.D., Meredith, S.C., Tycko, R. Molecular structure of b-amyloid fibrils in Alzheimer's disease brain tissue. *Cell.* **154**, 1257-1268 (2013).

28. Gremer, L., Schölzel, D., Schenk, C., Reinartz, E., Labahn, J., Ravelli, R.B.G., Tusche, M., Lopez-Iglesias, C., Hoyer, W., Heise, H., Willbold, D., Schröder, G.F. Fibril structure of amyloid-β(1-42) by cryo-electron microscopy. *Science*. **358,** 116-119 (2017).

29. Schütz, A.K., Vagt, T., Huber, M., Ovchinnikova, O.Y., Cadalbert, R., Wall, J., Güntert, P., Böckmann, A., Glockshuber, R., Meier, B.H. Atomic-resolution three-dimensional structure of amyloid β fibrils bearing the Osaka mutation. *Angew Chem Int Ed Engl.* **54,** 331-335 (2015).

30. Guenther, E.L., Ge, P., Trinh, H., Sawaya, M.R., Cascio, D., Boyer, D.R., Gonen, T., Zhou, Z.H., Eisenberg, D.S. Atomic-level evidence for packing and positional amyloid polymorphism by segment from TDP-43 RRM2. *Nat Struct Mol Biol.* **25,** 311-319 (2018).

31. Rodriguez, J. A., Ivanova, M.I., Sawaya, M.R., Cascio, D., Reyes, F.E., Shi, D., Sangwan, S., Guenther, E.L., Johnson, L.M., Zhang, M., Jiang, L., Arbing, M.A., Nannenga, B.L., Hattne, J., Whitelegge, J., Brewster, A.S., Messerschmidt, M., Boutet,

S., Sauter, N.K., Gonen, T., Eisenberg, D.S. Structure of the toxic core of α-synuclein from invisible crystals. *Nature*. **525,** 486–490 (2015).

32. Krotee, P., Griner, S.L., Sawaya, M.R., Cascio, D., Rodriguez, J.A., Shi, D., Philipp, S., Murray, K., Saelices, L., Lee, J., Seidler, P., Glabe, C.G., Jiang, L., Gonen, T., Eisenberg, D.S. Common fibrillar spines of amyloid-β and human islet amyloid polypeptide revealed by microelectron diffraction and structure-based inhibitors. *J Biol Chem.* **293,** 2888-2902 (2018).
33. Krotee, P., Rodriguez, J.A., Sawaya, M.R., Cascio, D., Reyes, F.E., Shi, D., Hattne, J., Nannenga, B.L., Oskarsson, M.E., Philipp, S., Griner, S., Jiang, L., Glabe, C.G., Westermark, G.T., Gonen, T., Eisenberg, D.S. Atomic structures of fibrillar segments of hIAPP suggest tightly mated β-sheets are important for cytotoxicity. *eLife.* **6,** e19273 (2017).
34. Hatami, A., Monjazeb, S., Milton, S., Glabe, C.G. Familial Alzheimer's Disease Mutations within the Amyloid Precursor Protein Alter the Aggregation and Conformation of the Amyloid-B Peptide. *J Biol Chem.* **292,** 3172-3185.
35. Milton, N.G. Phosphorylation of amyloid-beta at the serine 26 residue by human cdc2 kinase. *Neuroreport.* **12,** 3839-3844, 2001.
36. Goldschmidt, L., Teng, P.K., Reik, R., Eisenberg, D. Identifying the amylome, proteins capable of forming amyloid-like fibrils. *Proc Natl Acad Sci USA*. **107,** 3487-3492 (2010).
37. Matta, L.M., Zhu, C., Jagoda, J., Zinn, T. Mixing by resonant acoustic driving in a closed chamber. *J Propul Power.* **12,** 366-370 (1996).
38. Leung, D. A new and improved method for the preparation of drug nanosuspension formulations using acoustic mixing technology. *Intl J Pharm.* **473,** 10-19 (2014).
39. Shi, D., Nannenga, B.L., Iadanza, M.G., Gonen, T. Three-dimensional electron crystallography of protein microcrystals. *eLife.* **2,** e01345 (2013).

40. Nannenga, B.L., Gonen, T. Protein structure determination by MicroED. *Curr Opin Struct Biol.* **27,** 24-31 (2014).

41. Holm, L., Laasko, L. Dali server update. Nucleic acids research. 44, W351-355 (2016).

42. Garnham, C.P., Campbell, R.L., Davies, P.L. Anchored clathrate waters bind antifreeze proteins to ice. *Proc Natl Acad Sci.* **108,** 7363-7367 (2011).

43. Perutz, M. Polar zippers: Their role in human disease. Prot Sci. 3, 1629-1637 (1994).

44. Sawaya, M.R., Sambashivan, S., Nelson, R., Ivanova, M.I., Sievers, S.A., Apostol, M.I., Thompson, M.J., Balbirnie, M., Wiltzius, J.J., McFarlane, H.T., Madsen, A.Ø., Riekel, C., Eisenberg, D. Atomic structures of amyloid cross-β spines reveal varied steric zippers. *Nature.* **447,** 453-457 (2007).

45. Colletier, J., Laganowsky, A., Landau, M., Zhao, M., Soriaga, A.B., Goldschmidt, L., Flot, D., Cascio, D., Sawaya, M.R., Eisenberg, D. Molecular Basis for Amyloid-β Polymorphism. *Proc Natl Acad Sci USA*. **108,** 16938-16943 (2011).

46. Eisenberg, D.S., Sawaya, M.R. Structural studies of Amyloid Proteins at the Molecular Level. *Annu Rev Biochem.* **86,** 69-95 (2017).

47. Schmidt, M., Rohou, A., Lasker, K., Yadav, J.K., Schiene-Fischer, C., Fändrich, M., Grigorieff, N. Peptide dimer structure in an Aβ(1-42) fibril visualized with cryoEM. *Proc Natl Acad Sci USA*. **112,** 11858-11863.

48. Fitzpatrick, A.W.P, Falcon, B., He, S., Murzin, A.G., Murshudov, G., Garringer, H.J., Crowther, R.A., Ghetti, B., Goedert, M., Scheres, S.H.W. Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature.* **547,** 185-190 (2017).

49. Sievers, S.A., Karanicolas, J., Chang, H.W., Zhao, A., Jiang, L., Zirafi, O., Stevens, J.T., Münch, J., Baker, D., Eisenberg, D. Structure-based design of non-natural amino-acid inhibitors of amyloid fibril formation. *Nature.* **475,** 96-100 (2011).

50. Seidler, P.M., Boyer, D.R., Rodriguez, J.A., Sawaya, M.R., Cascio, D., Murray, K., Gonen, T., Eisenberg, D.S. Structure-based inhibitors of tau aggregation. *Nat Chem*. **10,** 170-176 (2018).

51. Saelices, L., Chung, K., Lee, J.H., Cohn, W., Whitelegge, J.P., Benson, M.D., Eisenberg, D.S.. Amyloid seeding of transthyretin by ex vivo cardiac fibrils and its inhibition. *Proc Natl Acad Sci USA.* **115,** e6741-6750 (2018).

52. Saelices, L., Sievers, S.A., Sawaya, M.R., Eisenberg, D.S. Crystal Structures of Amyloidogenic Segments of Human Transthyretin. *Protein Sci*. **27,** 1295-1303 (2018).

53. Li, B., Ge, P., Murray, K.A., Sheth, P., Zhang, M., Nair, G., Sawaya, M.R., Shin, W.S., Boyer, D.R., Ye, S., Eisenberg, D.S., Zhou, Z.H., Jiang, L. Cryo-EM of full-length a-synuclein reveals fibril polymorphs with a common structural kernel. *Nat Commun*. **9,** 3609 (2018).

54. Guenther, E.L., Cao, Q., Trinh, H., Lu, J., Sawaya, M.R., Cascio, D., Boyer, D.R., Rodriguez, J.A., Hughes, M.P., Eisenberg, D.S. Atomic structures of TDP-43 LCD segments and insights into reversible or pathogenic aggregation. *Nat Struct Mol Biol*. **25,** 463-471 (2018).

55. Hattne, J., Reyes, F.E., Nannenga, B.L., Shi, D., de la Cruz, M.J., Leslie, A.G.W., Gonen, T. MicroED data collection and processing. *Acta Crystallogr A Found Adv.* **71,** 353-360 (2015).

56. Kabsch, W. XDS. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 125–132 (2010).

57. Sheldrick, G. M. A short history of SHELX. *Acta Crystallogr. A.* **64,** 112–122 (2008)

58. Emsley, P., Lohkamp, B., Scott, W.G., Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*. **66,** 486-501 (2010).

59. Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimoc, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., Adams, P.D. Towards

automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr*. **68,** 352-367 (2012).

60. Porebski, B.T., Ho, B.K., Buckle, A.M. Interactive visualization tools for the structural biologist. *J. Appl. Cryst.* **46,** 1518-1520 (2013).
61. Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., Read, R.J., Rice, L.M., Simonson, T., Warren, G.L. Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr.* **54**, 905-921 (1998).

Table 1. Data collection and refinement statistics

	$A\beta^{20-34}$	$A\beta^{20-34, isoAsp23}$
Data collection		
Space group	$P2_1$	$P2_1$
Cell dimensions		
a, *b*, *c* (Å)	33.17, 4.78, 30.33	29.20, 4.87, 32.44
α, β, γ (°)	90.00, 111.10, 90.00	90.00, 101.90, 90.00
Resolution (Å)	1.10	1.05 (1.20-1.05)*
R_{sym} or R_{merge} (%)	18.9	19.7
$I / \sigma I$	5.41 (3.28)	3.76 (1.38)
Completeness (%)	85.2	82.7 (53.0)
Redundancy	6.67 (6.14)	4.19 (3.10)
Refinement		
Resolution (Å)	7.74-1.10 (1.13-1.10)	5.96-1.05 (1.20-1.05)
No. reflections	3544 (1141)	3943 (1167)
R_{work} / R_{free} (%)	19.4/21.3 (21.3/26.9)	19.7/24.6 (27.0/32.4)
No. atoms		
Protein	210	204
Ligand/ion	0	0
Water	7	4
B-factors		
Protein	6.50	8.29
Ligand/ion	-	-
Water	20.78	27.70
R.m.s. deviations		
Bond lengths (Å)	0.56	1.04
Bond angles (°)	0.68	0.90

*Ten crystals were used in determining the $A\beta^{20-34}$ structure

*Five crystals were used in determining the $A\beta^{20-34, isoAsp23}$ structure

*Values in parentheses are for highest-resolution shell.

FIGURES:

Figure 1

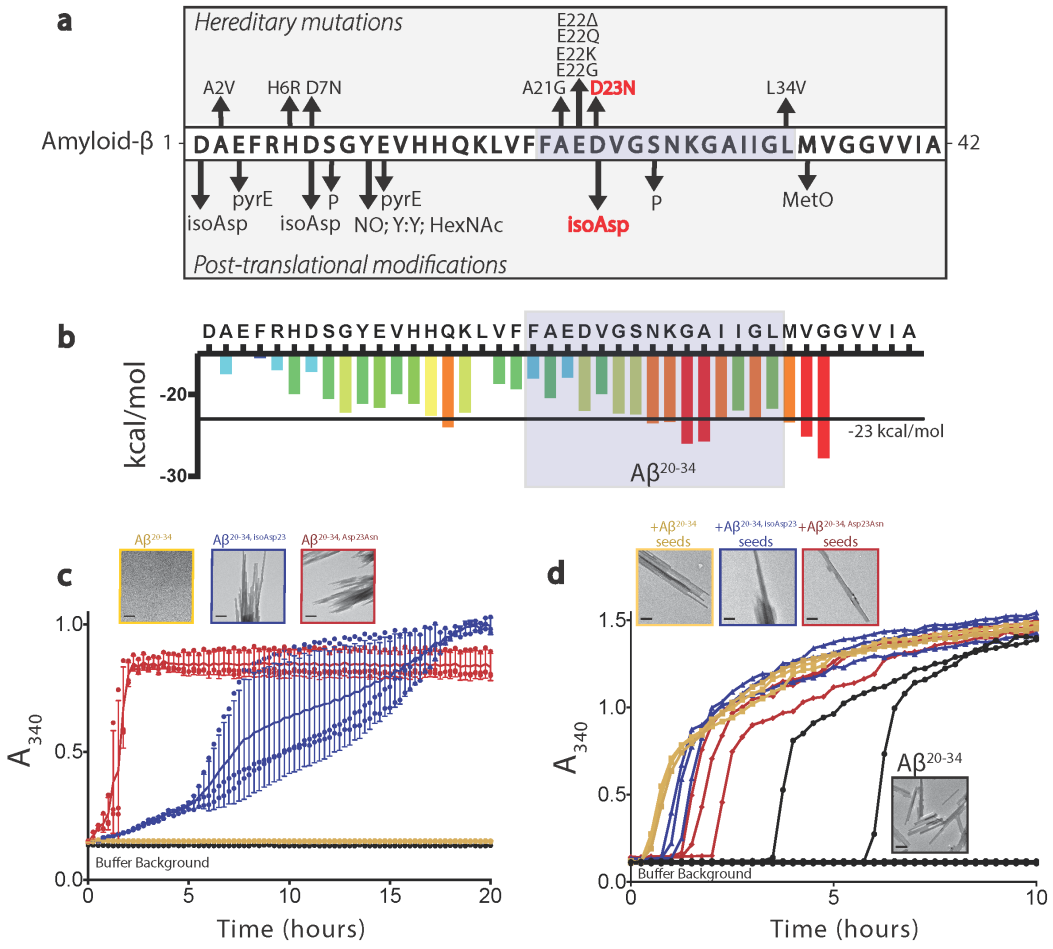


FIG. 1. L-isoAsp in Aβ$^{20-34}$ accelerates fiber formation and can seed native segment. a, Sequence of human Aβ including known early-onset hereditary mutations and PTMs (pyrE = pyroglutamate; P = phosphorylation; NO = nitration; Y:Y = dityrosine crosslink; HexNAc = glycosylation; MetO = oxidation) **. b,** ZipperDB[22] amyloid propensity profile for the human Aβ sequence with the Aβ$^{20-34}$ sequence highlighted in light blue. **c,** 1.6 mM of Aβ$^{20-34}$, Aβ$^{20-34, Asp23Asn}$, and Aβ$^{20-34, isoAsp23}$ peptide aggregation was monitored by turbidity at 340 nm. Each data point is shown as a round symbol, the solid line represents the mean value, and error bars represent SD of 3 replicates. EM images of aggregates are shown at the top left of the graph, scale bars

represent 0.5 μm in each image. **d**, Aggregation of 3.2 mM $A\beta^{20\text{-}34}$ in 50 mM Tris, pH 7.5, 150 mM NaCl, and 1% DMSO was monitored by turbidity at 340 nm alone (black lines), or with 10 μM preaggregated seeds of $A\beta^{20\text{-}34}$ (yellow), $A\beta^{20\text{-}34,\ isoAsp23}$ (blue), $A\beta^{20\text{-}34,\ Asp23Asn}$ (red). Each line represents a replicate well. Electron micrographs of aggregates are shown at the top left of the graph, scale bars shown at the lower left represent 0.5 μm in each image.

Figure 2

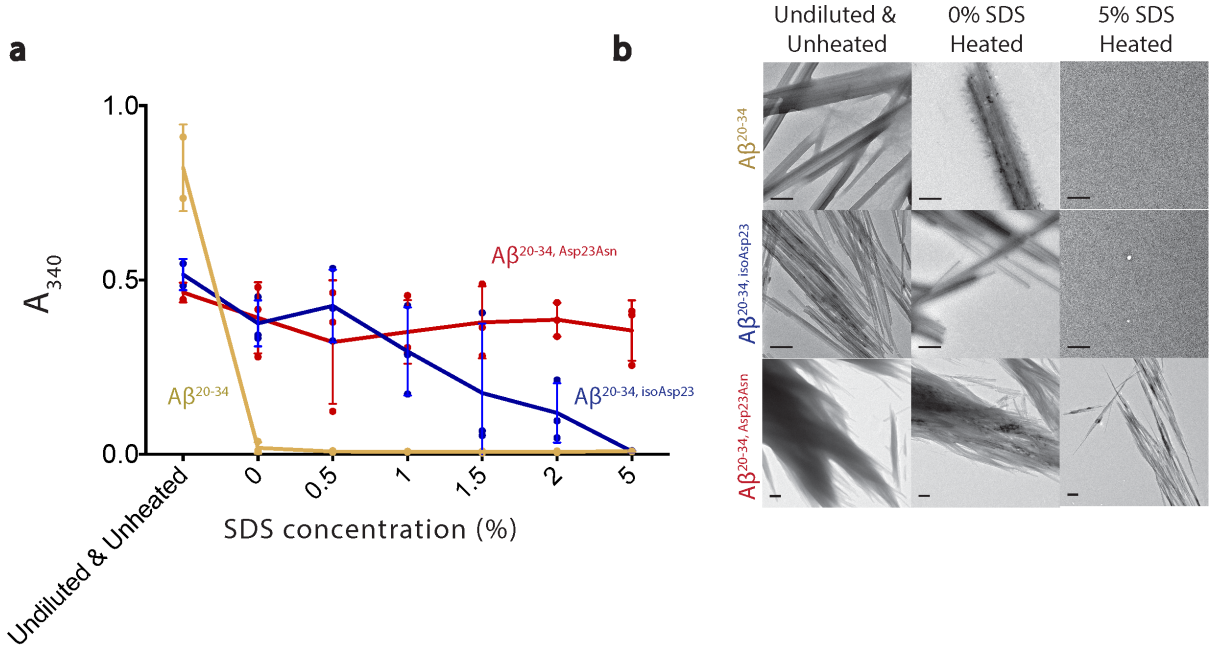


FIG. 2: **Modified fibers have increased resistance to SDS disaggregation. a**, Fiber stocks (Undiluted & Unheated initial points are two readings of the fiber stocks) were mixed 1:1 in buffer (0% SDS final) and increasing concentrations of SDS (1, 1.5, 2, 5% final) as described in the Methods. Each data point is shown as a round symbol, the solid line represents the mean value, and error bars represent the SD of three technical replicates. **b,** EM images of disaggregated fibers, scale bars in the lower left represent 0.5 μm.

Figure 3

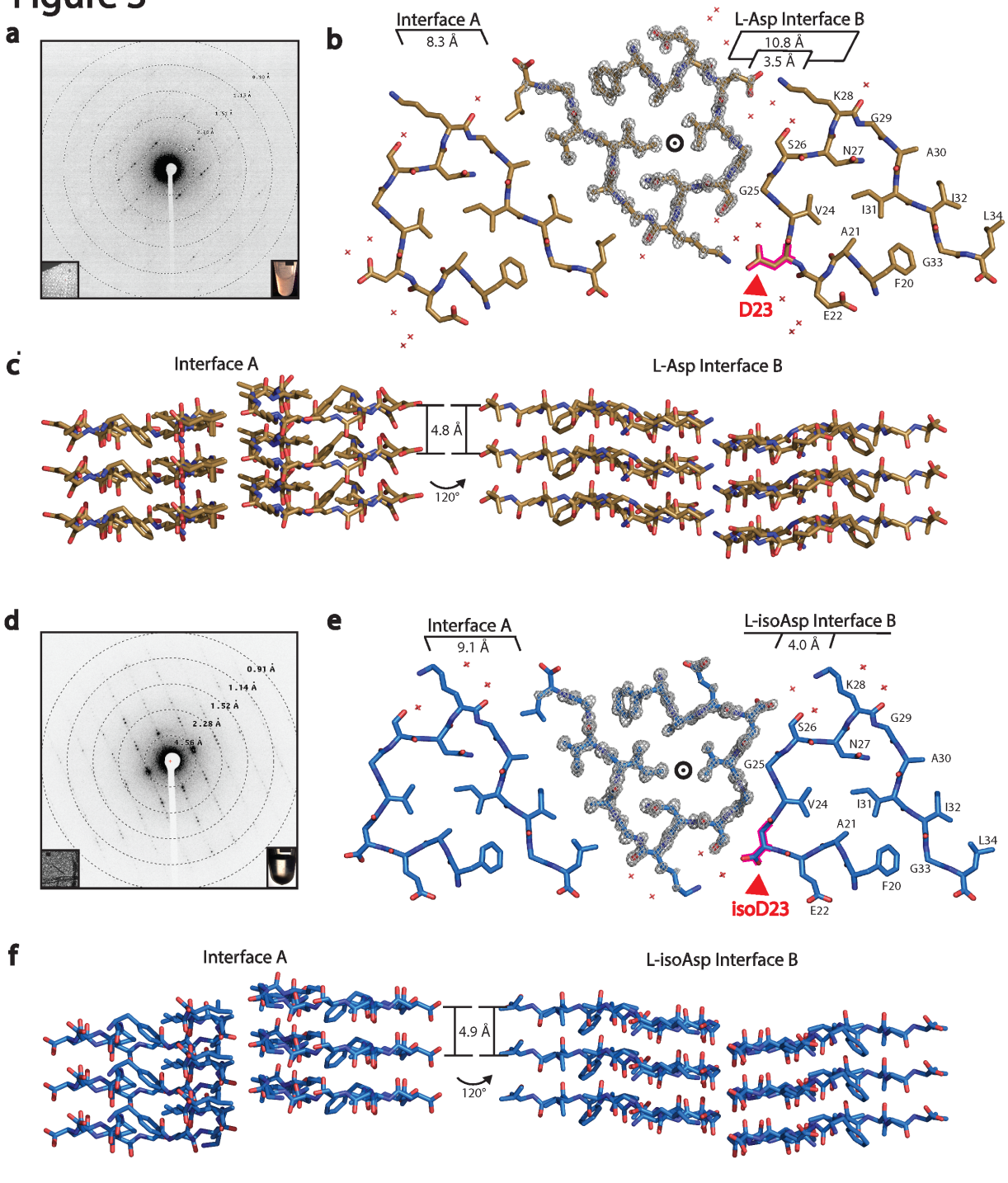


FIG. 3: Aβ$^{20\text{-}34,\ isoAsp23}$ structure contains an altered protofilament interface. a, Representative single crystal electron diffraction pattern of Aβ$^{20\text{-}34}$ with resolution rings obtained during MicroED data collection. Left inset shows the diffracting crystal (lower left scale bar represents 1 μm). Right

inset shows light microscope image of microcrystal sediment in a 1.6 mL microfuge tube (scale bar represents 3 mm). **b**, One layer of the $A\beta^{20-34}$ crystal structure viewed down the fibril axis highlighting two distinct steric zipper interfaces. Interface distances are labeled (center black circle indicates fibril axis). Waters are represented by red crosses. The $2F_o$-F_c density is shown as a grey mesh at 2σ on the center protofilament. The Asp-23 residue is outlined in magenta and shown by the red arrow. **c**, Three layers of the $A\beta^{20-34}$ structure viewed perpendicular to the fibril axis (indicated by arrows; Left panel – Interface A, right panel – L-Asp Interface B). **d**, Representative single crystal electron diffraction pattern of $A\beta^{20-34,\ isoAsp23}$ with resolution rings obtained during MicroED data collection. Left inset shows the diffracting crystal (lower left scale bar represents 1 μm). Right inset shows a light microscope image of microcrystal sediment in a 1.6 mL microfuge tube (scale bar represents 3 mm). **e**, One layer of the $A\beta^{20-34,\ isoAsp23}$ crystal structure viewed down the fibril axis highlighting two distinct steric zipper interfaces. Interface distances are labeled (center black circle indicates fibril axis). Waters are represented by red crosses. The $2F_o$-F_c density is shown as a grey mesh at 2σ on the center protofilament. The L-isoAsp-23 residue is outlined in magenta and shown by the red arrow. **f**, Three layers of the $A\beta^{20-34,\ isoAsp23}$ structure viewed perpendicular to the fibril axis (indicated by arrows; Left panel – Interface A, right panel – L-isoAsp Interface B).

Figure 4

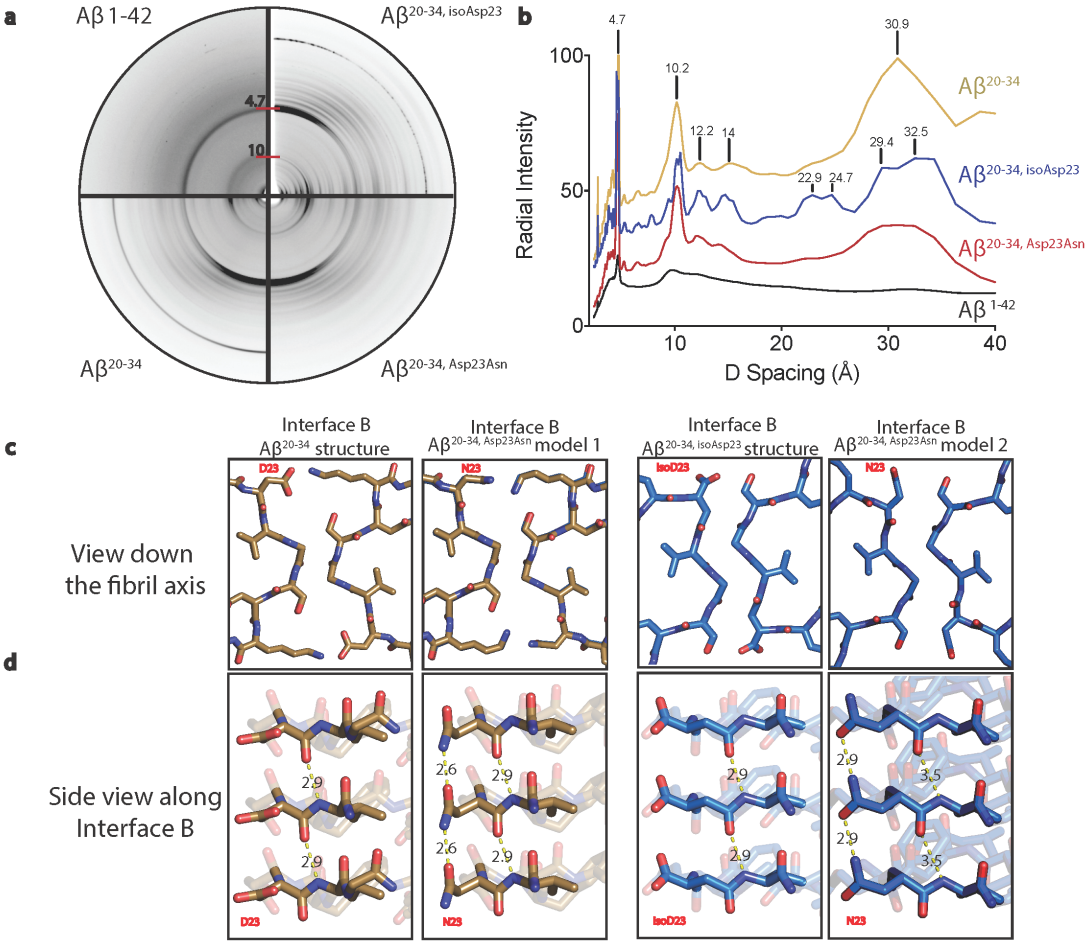


FIG. 4: A putative model of the heritable Iowa mutation in interface B. a, Fiber diffraction patterns of Aβ 1-42, $Aβ^{20-34}$, $Aβ^{20-34, isoAsp23}$, and $Aβ^{20-34, Asp23Asn}$**.** All fibers including Aβ 1-42 were prepared in 50 mM Tris, pH 7.6, 150 mM NaCl, and 1% DMSO, excepting $Aβ^{20-34, Asp23Asn}$, in which the DMSO concentration was raised to 5%. **b**, Intensities of reflections from fiber diffraction of the segments were plotted against D spacing. Radial intensity values for staggered for visibility of peaks. **c**, From left to right, interface B down the fibril axis of $Aβ^{20-34}$ structure, a model $Aβ^{20-34, Asp23Asn}$ on the backbone of the $Aβ^{20-34}$ structure, $Aβ^{20-34, isoAsp23}$ structure, and a model $Aβ^{20-34, Asp23Asn}$ on the backbone of the $Aβ^{20-34, isoAsp23}$ structure. **d,** A view perpendicular to

the fibril axis of residues 23-24 of each structure. Yellow dashed lines represent measured distances in Å between the amide carboxyl of residue 23 and the amide nitrogen of Val24 on the adjacent strand.

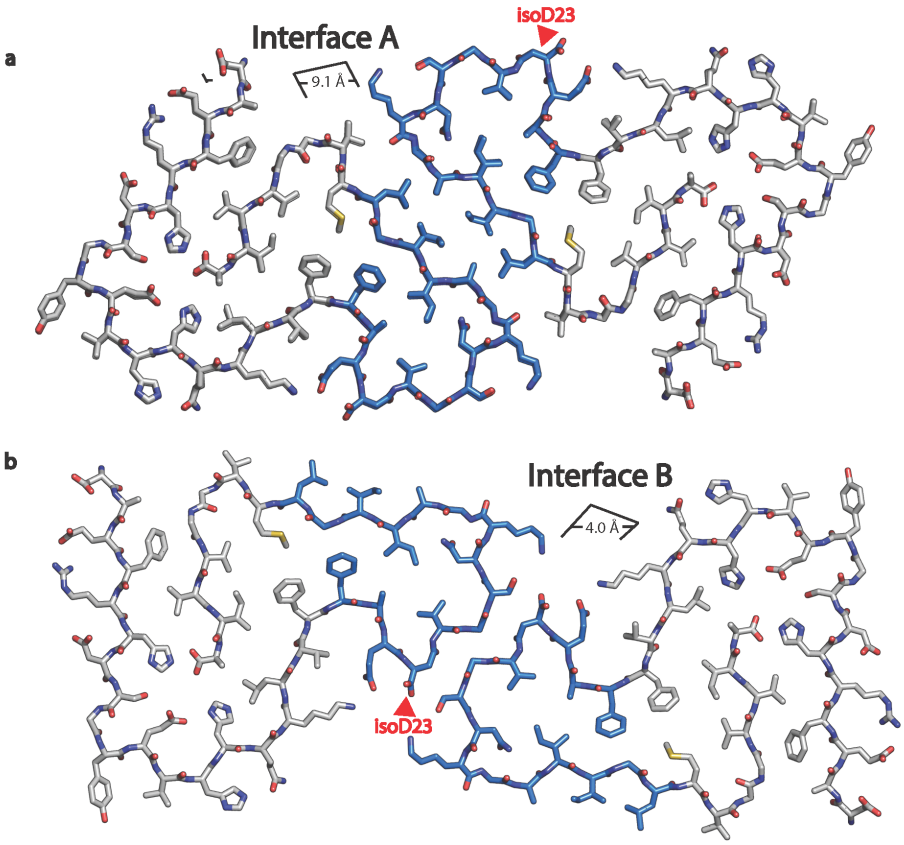


FIG. 5: Model of Aβ$^{20\text{-}34,\ isoAsp23}$ as the core of an Aβ 1-42 modified polymorph. a, Model of Aβ$^{1\text{-}42,\ isoAsp23}$ centered on interface A. **c**, Centered on interface B. Blue residues correspond to the crystal structure core (Aβ$^{20\text{-}34,\ isoAsp23}$), gray sticks correspond to the modeled extension (1-19 and 35-42).

Figure 6

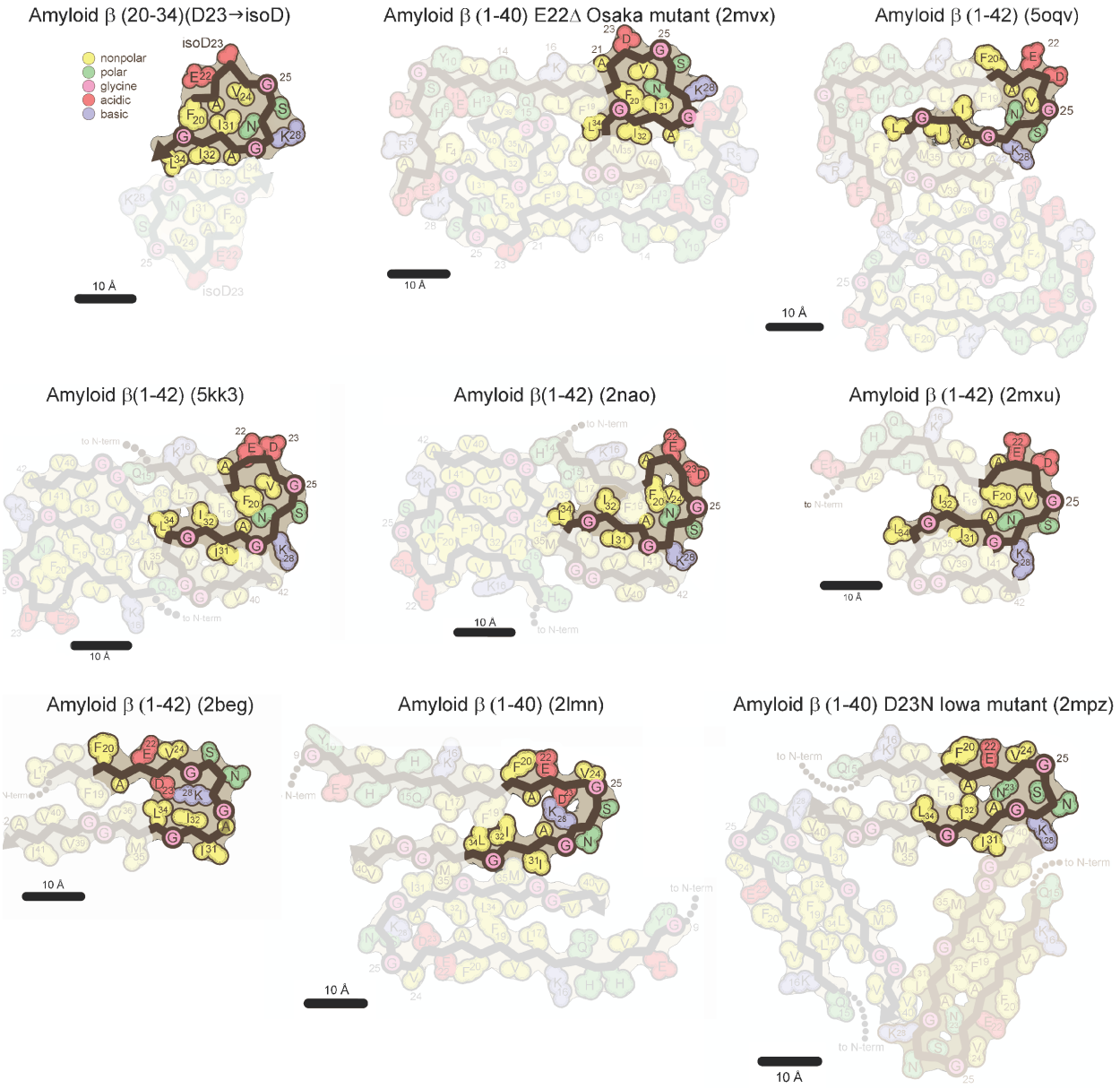


FIG. 6: $A\beta^{20\text{-}34,\ isoAsp23}$ core assumes a similar fold to full-length native Aβ structures. Schematic diagrams of residues 20-34 of previously solved Aβ structures[21-26,28-29]. The most structurally divergent structures from $A\beta^{20\text{-}34,\ isoAsp23}$ are shown in the bottom panel. Residues are colored according to general chemical properties (legend – top left).

Figure 7

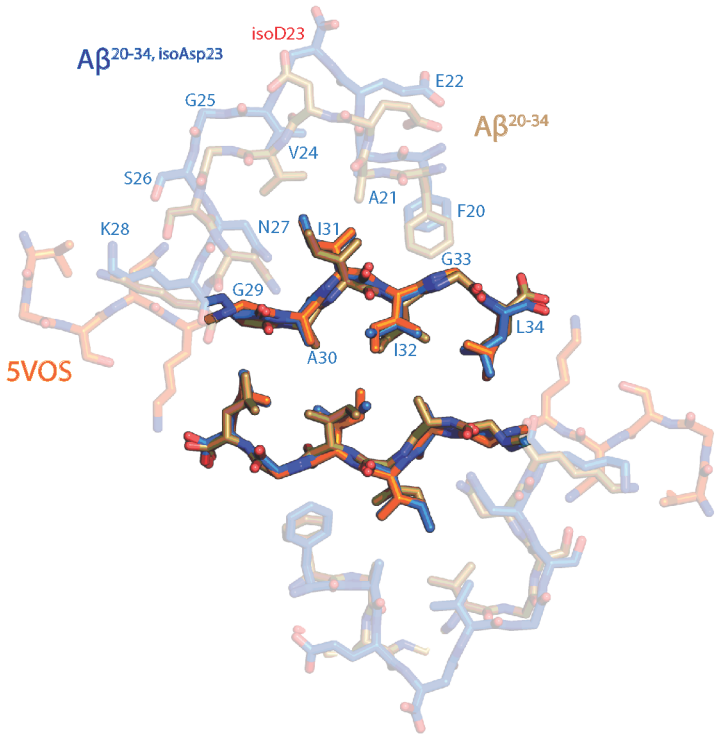


FIG. 7: Aβ 24-34 peptide structure shares similarities with the Aβ$^{20-34}$ structures. The Aβ steric zipper structure with the lowest total atom RMSD, 5VOS (orange), is shown aligned with Interface A of Aβ$^{20-34}$ (gold) and Aβ$^{20-34, isoAsp23}$.

SUPPLEMENTARY INFORMATION

SUPPLEMENTARY METHODS:

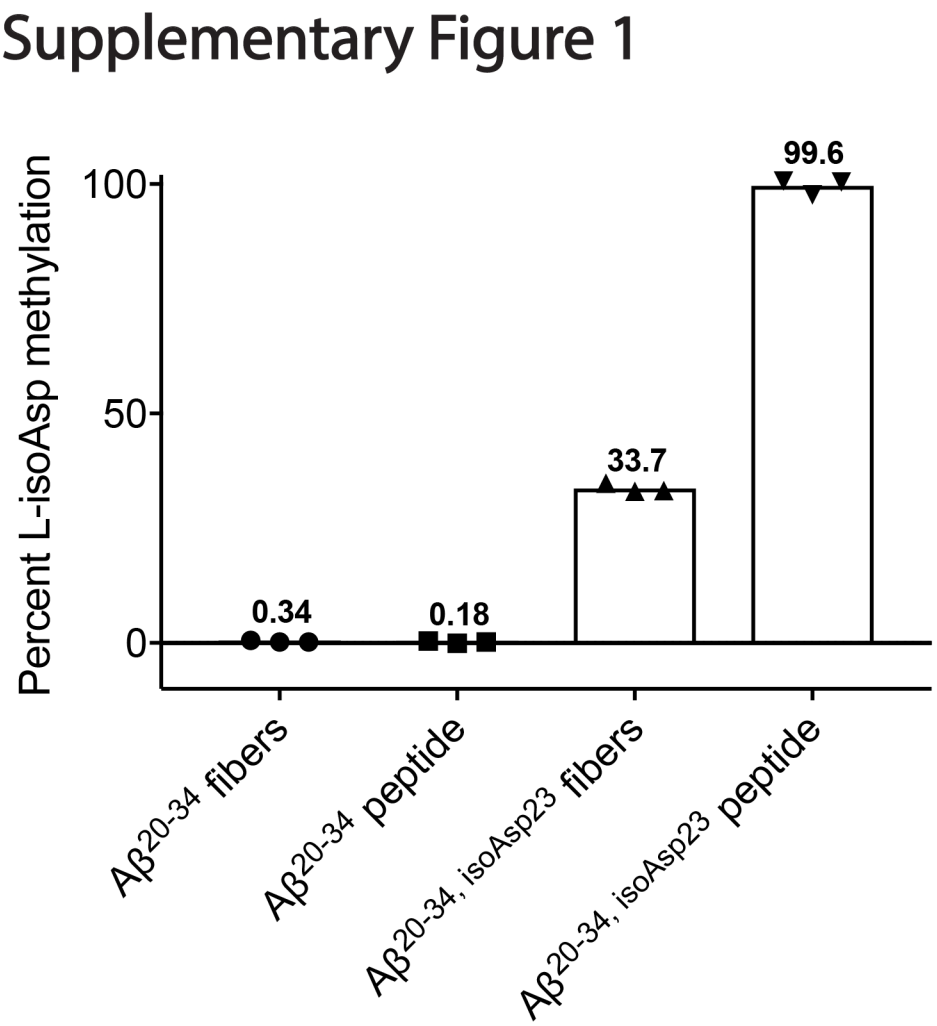
Determination of L-isoaspartate levels by the PCMT1 methanol vapor diffusion assay

PCMT1 was used as an analytical reagent to quantify L-isoAsp levels in $A\beta^{20-34}$ peptide solutions or aggregates. Aggregates of $A\beta^{20-34}$ or $A\beta^{20-34,\ isoAsp23}$ were formed at 3.2 or 1.6 mM, respectively in 50 mM Tris-HCl, pH 7.5, 150 mM NaCl with 1% DMSO. In a final volume of 100 µL, 130 pmol of either these aggregates or freshly dissolved, filtered peptide solutions were incubated for 2 h at 37 °C with 5 µg PCMT1 (purified as a His-tagged enzyme from *Escherichia coli (E. coli)* containing the expression plasmid #34852 available from Addgene.com as described by Patananan et al., 2014[1] with a specific activity at 37 °C of 5,300 pmol of methyl esters formed on KASA(isoD)LAKY/min/mg of enzyme). Final concentrations in the reactions included 135 mM Bis-Tris-HCl, pH 6.4, and 10 µM S-adenosyl-l-[methyl3H]methionine ([3H]AdoMet) (prepared by a 1600-fold isotopic dilution of a stock of 72 Ci/mmol [3H]AdoMet (PerkinElmer Life Sciences, NET155H00) with nonisotopically labeled AdoMet (p-toluenesulfonate salt; Sigma-Aldrich A2408)). The reaction was stopped by adding 10 µL of 2 M sodium hydroxide, and 100 µL of the 110 µL mixture was transferred to a 9 by 2.5 cm piece of folded thick filter paper (Bio-Rad; catalog number 1650962) wedged in the neck of a 20-mL scintillation vial above 5 mL scintillation reagent (Safety Solve, Research Products International, catalog number 121000), tightly capped, and incubated at room temperature. After 2 h, the folded filter papers were removed, the caps replaced, and the vials were counted thrice for 5 minutes each in a Beckman LS6500 scintillation counter. Background radioactivity in a reaction containing no substrate was determined by incubating the recombinant human PCMT1, 135 mM Bis-Tris-HCl buffer, and 10 µM [3H] AdoMet as described above and was subtracted from the value obtained in experimental samples. Samples were analyzed in triplicate.

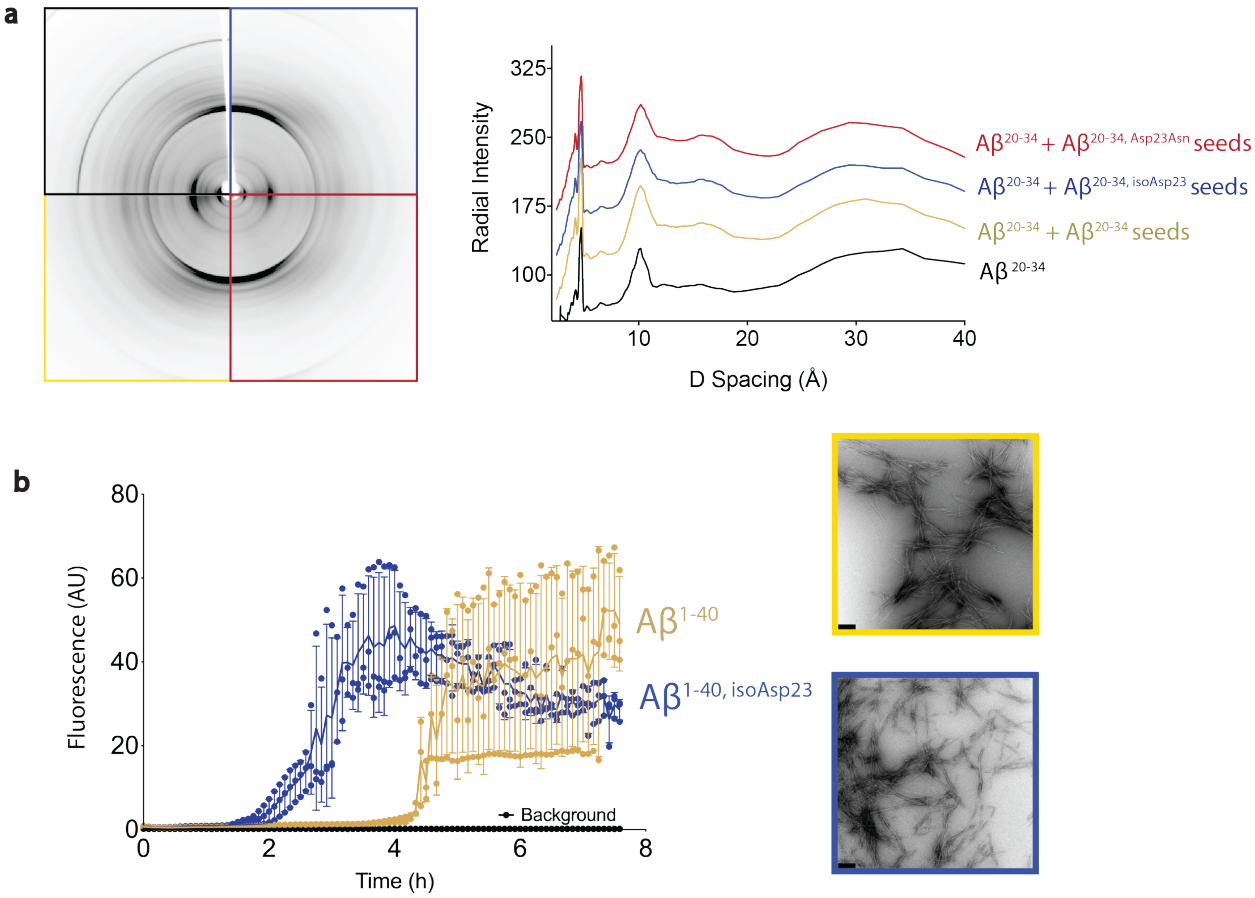
Fibril formation of full length Aβ 1-40

Wild-type and L-isoAsp23 Aβ 1- 40 were resuspended at a final concentration of 40 μM with 10 μM ThT in 10 mM phosphate, 127 mM NaCl, and 2.7 mM KCl, pH 7.4 (PBS). Fibrils were formed at 37 °C with continuous shaking at 600 rpm in a Varioskan plate reader. Fluorescence of three replicate wells was monitored in a 96 well plate, with readings taken every 5 min (excitation – 440 nm, emission - 482 nm, bottom read). Concentrations of all peptide solutions were verified by absorbance at 280 nm. EM images were recorded on an FEI Tecnai G_2 TF20 TEM.

SUPPLEMENTARY FIGURES:

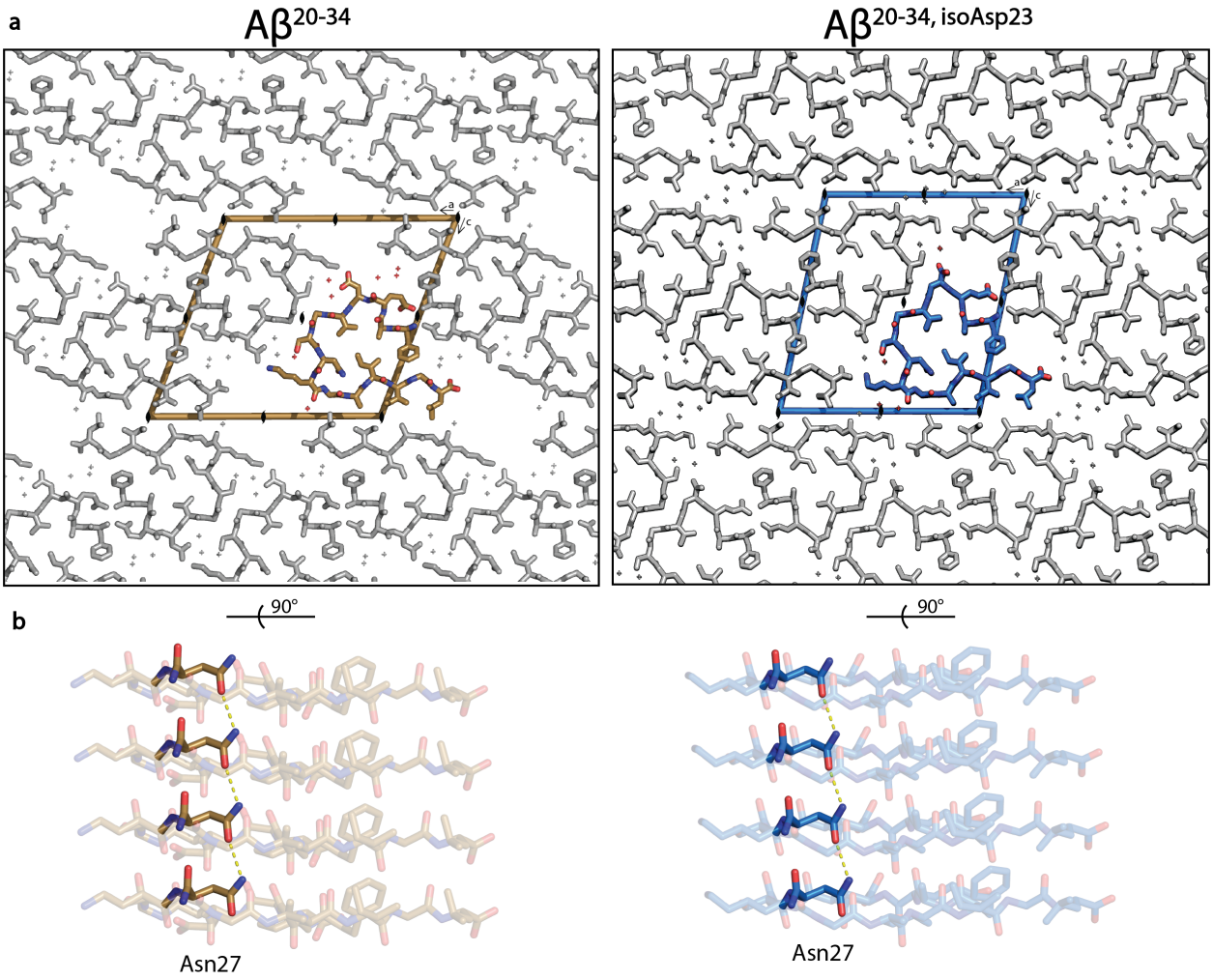


SUPPLEMENTARY FIGURE 1. PCMT1 is unable to fully methylate aggregated Aβ$^{20\text{-}34,\ isoAsp23}$. Methylation of free peptide and aggregated fibers of Aβ$^{20\text{-}34}$ and Aβ$^{20\text{-}34,\ isoAsp23}$ was detected as described in the experimental procedures. Levels of detected L-isoAsp were normalized between 0-100% methylation. The normal L-Asp Aβ$^{20\text{-}34}$ fibers and peptide were included as negative controls and were not methylated by PCMT1, as shown by the averages of 0.34% and 0.18% methylation, respectively.



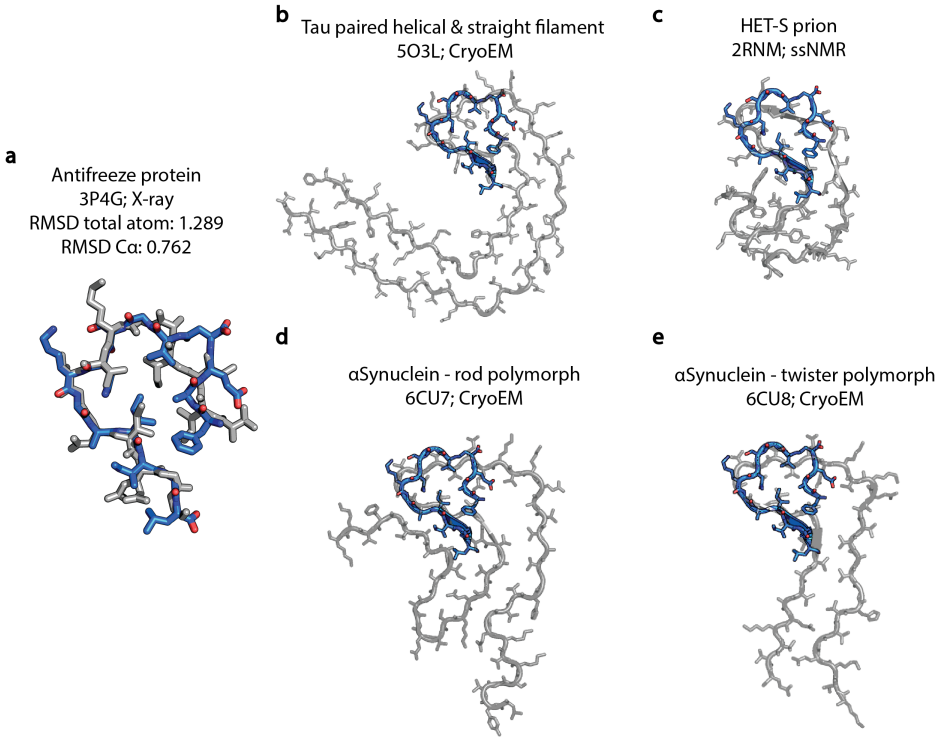
SUPPLEMENTARY FIGURE 2: Seeded Aβ$^{20-34}$ aggregates display identical fiber diffraction patterns, and the full-length aggregation matches the segment seeding. a, Aggregates from the seeding assay of Aβ$^{20-34}$ shown in the main text Fig. 1d were ordered between glass capillaries and fiber diffraction data was collected as described in the "Methods." Radial intensity of the reflections was plotted against D Spacing (right). **b**, 20 μM wild-type Aβ$^{1-40}$ (gold lines) or Aβ$^{1-40, isoAsp23}$ (blue lines) were incubated at 37 °C. Fiber formation was monitored by Thioflavin T fluorescence, readings were recorded every 5 min. Each data point is shown as a round symbol, the solid line represents the mean value, and error bars represent SD of three technical replicates. A representative EM image of each condition is shown on the right, scale bars at the lower left represent 200 nm. Source data are provided as a Source Data file.

Supplementary Figure 3



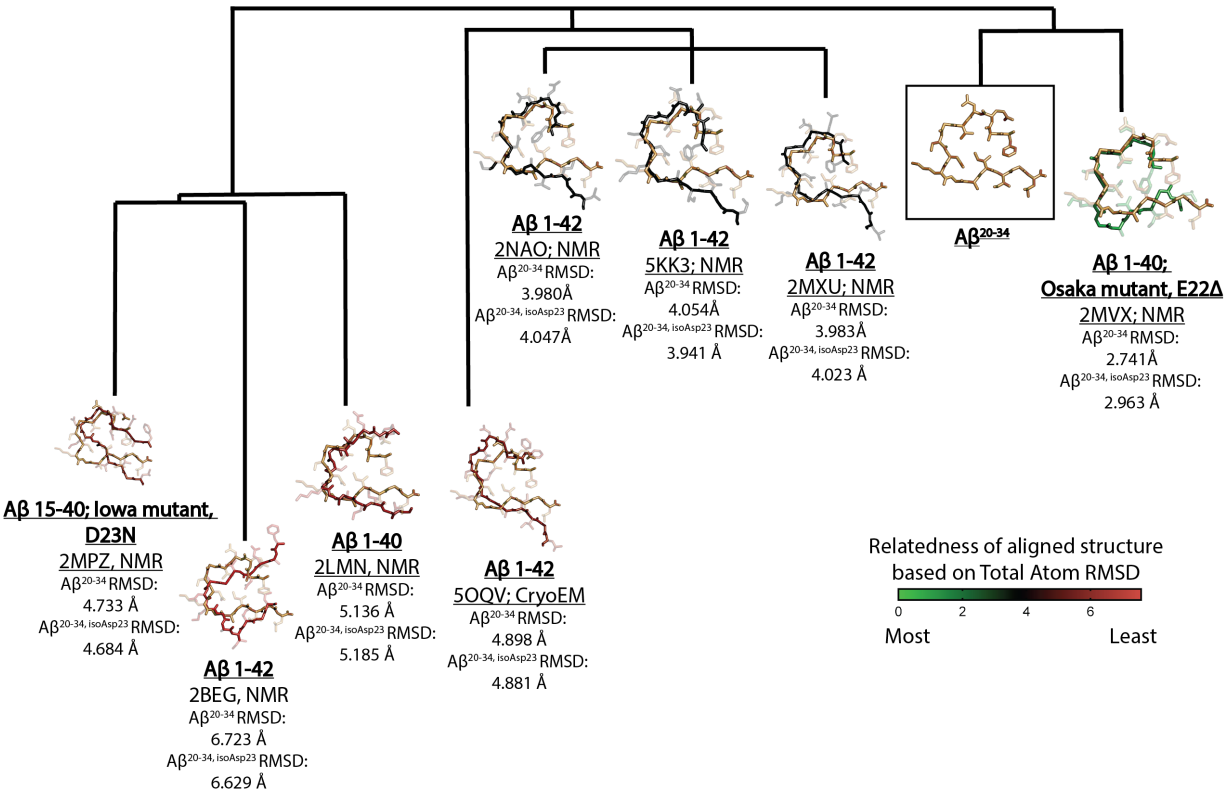
SUPPLEMENTARY FIGURE 3. Aβ$^{20-34}$ structures reveal two interfaces and a polar zipper. a, The crystal structures are shown here along the 2_1 axis of the unit cell. Black symbols ( ) represent the 2_1 axis of symmetry. **b**, The asparagine ladder motif is shown with yellow dashed lines between strands, the structure is shown perpendicular to the protofilament axis, along the face of residues Lys28-Leu34. The dashed lines within the Aβ$^{20-34}$ structure correspond to 2.8 Å, and the dashed lines within the Aβ$^{20-34, isoAsp23}$ structure correspond to 2.9 Å.

Supplementary Figure 4



SUPPLEMENTARY FIGURE 4. $A\beta^{20\text{-}34,\ isoAsp23}$ matches β-helical structures and other amyloid fibrils. a, The top match for the $A\beta^{20\text{-}34,\ isoAsp23}$ structure from comparison to all PDB structures via a DALI search[2] was the β-helical antifreeze protein 3P4G[3]. **b-d**, Other amyloid fibril structures, including tau[4] and αSynuclein[5] show similar tight turns (β-arches) between steric zippers.

Supplementary Figure 5



SUPPLEMENTARY FIGURE 5. Total atom RMSD relationships between Aβ structures. The $A\beta^{20\text{-}34,\ isoAsp23}$ (gold, second from right) was aligned to residues 20-34 of full-length Aβ structures[6-13]. Backbone and total atom RMSDs were calculated using CCP4. Branches of the evolutionary tree represent total atom RMSD relatedness between the structures and were generated as described in the Methods section. The backbones of the aligned structures are also colored on a scale of red to green according to how closely matched their total RMSD values are to $A\beta^{20\text{-}34}$ (legend shown on bottom right).

CHAPTER 5

The PCMTDs are a novel L-isoAsp maintenance pathway and interact with cullin-RING ligase proteins

ABSTRACT:

The damaged L-isoaspartate (L-isoAsp) modification has often been observed to have a negative effect on protein function, but is known to be repaired by the protein-L-isoaspartyl (D-aspartyl) *O*-methyltransferase (PCMT1). Studies have demonstrated that L-isoAsp levels unexpectedly plateau in *Pcmt1* knockout mice that are transgenic with a mouse *Pcmt1* cDNA under a neuron-specific promoter, while levels of L-isoAsp containing peptides in the urine increased. Thus it has been hypothesized that a secondary maintenance pathway exists which recognizes L-isoAsp damaged proteins and initiates their degradation. Herein we described the protein carboxyl methyltransferase domain-containing protein 1 (PCMTD1) that contains L-isoAsp binding motifs as well as SOCS box ubiquitin ligase recruitment motifs. We show that this protein interacts with components of an E3 cullin-RING ligase complex, which is necessary for PCMTD1 binding of *S*-adenosylmethionine binding. The PCMTD1 in combination with the E3 ligase complex is able to reduce the levels of L-isoAsp detected by PCMT1 *in vitro*. Thus we propose that the L-isoAsp modification acts as an “off-switch” for aged proteins, with the PCMTD proteins as L-isoaspartyl specific E3 ligase adaptor proteins that target the damaged proteins for degradation.

INTRODUCTION:

Proteins can accumulate a number of post-translational modifications (PTMs) over time that alter normal enzymatic function, and threaten protein stability. These include oxidation, carbonylation, glycation, deamidation, and isomerization, which can occur by inter- or intra-molecular reactions (1). Even with this variety of damaging alterations, only a few protein repair mechanisms exist. Methionine sulfoxide reductases (Msrs) can catalyze the reduction of methionine sulfoxide back to the methionine-reduced form (2). Protein deglyclation can be accomplished by phosphorylation of the third carbon of the ketoamine, which creates an unstable product that can be released as a keto-deoxyaldolase (3). Finally, the protein-L-isoaspartyl (D-aspartyl) *O*-methyltransferase (PCMT1) is a well-characterized repair enzyme for L-isoaspartate and D-aspartate residues, returning them to their normal L-aspartate state (4).

Even these three mechanisms do not address all cases of methionine oxidation, glycation, and aspartate isomerization, and many of these modified proteins, along with a majority of other damaged proteins are funneled to the lysosome or the proteasome (5). The lysosome is a powerful protein degradation system, capable of breaking down whole aggregates, organelles, and aged cytosolic proteins through various forms of autophagy. The ubiquitin-proteasome system is more regulated, involving over 600 E3 ligases, which recognize specific substrates (6), thus only these proteins are ubiquitinated and directed towards degradation by the proteasome. The RING ligase family is the largest family of E3 ligases, so called because of the RING domain, which recruits the appropriate E2 enzyme (7).

One of the subcategories of E3 ligase families – the Elongin-C-Cullin-SOCS-box (ECS) family – is characterized by the presence of a SOCS box motif in the C-terminus of the substrate recognition protein (the SOCS box protein; ref. 8). This motif consists of a “Cul-box” and a “BC-box,” which interact with the RING complex core proteins Cullin 2 or Cullin 5, and the Elongin B and Elongin C proteins, respectively. The Elongins further mediate interactions between the SOCS box protein and the core Cullin protein. The Elongins and the SOCS box

protein interact with the N-terminal domain of the Cullin, while the C-terminal domain of the Cullin binds the RING protein, Rbx1 or Rbx2, which then binds the activated E2 ligase. The approximately circular shape of this complex brings the substrate bound to the SOCS box protein in close proximity to the activated E2 ligase, which can then transfer ubiquitin to a lysine on the surface of the substrate protein (9).

A degradation pathway for L-isoaspartate-damaged proteins was hypothesized when elevated levels of L-isoaspartate-containing peptides were detected in the urine of transgenic mice lacking the L-isoaspartate repair enzyme PCMT1 in all cells excepting neurons (10). Herein we describe the characterization of the PCMT domain-containing protein 1 (PCMTD1). This protein (along with the homologous PCMTD2) not only contains the same L-isoaspartate recognition motifs as the PCMT1 enzyme, but also contains an extended C-terminal domain with the SOCS box motif (originally identified in Suppressor of Cytokine Signaling proteins). We have shown that PCMTD1 co-purifies with Elongins B and C and interacts with Cullin 5 (Cul5). We have further shown that while the PCMTD1 enzyme can bind the methyltransferase co-factor *S*-adenosylmethionine, it has no detectable methyltransferase activity. However, it is able to lower the levels of detectable L-isoaspartate in wild-type and *Pcmt1* knockout mouse lysates and that this is bolstered by the presence of Cul5-RING ligase components. Although we do not identify specific substrates, we believe that the PCMTD1 protein represents a novel, protein damage-specific E3 ligase.

RESULTS:

PCMTD1 conserves functional binding sites from PCMT1 and SOCS-box proteins

Figure 1 displays a sequence alignment of the human amino acid sequences of PCMT1 (227 residues), PCMTD1 (357 residues), and the homologous protein PCMTD2 (361 residues). The sequences are ~26% similar to PCMT1, with ~130 residues inserted between Thr217 and Asp218 of the equivalent PCMT residues. The red boxes in Fig. 1 show conservation of the co-factor *S*-adenosylmethionine (AdoMet) binding domains, and the blue boxes show residues

involved in recognition and binding of L-isoaspartyl substrates. Within the additional 130 amino acids of the PCMTD proteins, we display in green two sequence motifs that comprise the "SOCS box," the BC-box and the Cul-box. The BC-box is a 12 residue motif that recruits the proteins Elongin-B and Elongin-C, which mediate interactions of SOCS-box proteins with the cullin-RING core proteins, the cullins. A leucine at the +4 position within this motif has been shown to be critical for this interaction (11, 12, 13), which is conserved within the PCMTD sequences. The PCMTD proteins are very unique in comparison to all other known SOCS-box containing proteins, as in these proteins the Cul-box is usually +11 residues from the BC-box. In the PCMTDs, the Cul-box is separated from the BC-box by ~90 residues. This Cul-box helps recruit the cullin-RING ligase core protein cullin 2 or 5 to SOCS-box proteins.

Importantly, it has been shown that the SOCS-box is not fully folded in the absence of Elongins B and C (12). Consistent with this, we found the purification of His-PCMTD1 was greatly improved by co-expression with Elongins B and C. The binding of the Elongins to PCMTD1 can be seen by their co-elution by gel filtration, with unbound Elongins eluting in later fractions (**Fig. 2A**). The ability of this protein to bind AdoMet was tested by incubating radiolabeled AdoMet (S-adenosyl-l-[methyl^{3}H]methionine ([^{3}H]AdoMet)) with PCMT1 as a positive control, PCMTD1, and PCMTD1-Elongins. Each of these incubations were exposed to short-wave ultraviolet light to crosslink [^{3}H] AdoMet bound to proteins, separated by SDS-PAGE, and exposed to film. As mentioned above, the yield of the His-PCMTD1 alone purification was very poor, thus less than <0.2 µg was able to be loaded onto the gel, the most visible band is indicated by an asterisk (**Fig. 2B**). In Lanes 1-3, [^{3}H] AdoMet can be seen bound to both PCMT1 and the PCMTD1-Elongins complex. If the AdoMet is being specifically bound, it is known that the unmethylated equivalent, *S*-adenosylhomocysteine (AdoHcy) can compete with the AdoMet, but adenosine-triphosphate (ATP) does not. Non-radioactive ATP and AdoHcy were added to the incubations in lanes 4-6 and 7-9, respectively. The PCMT1-[^{3}H] AdoMet signal was not affected by the presence of ATP, and while the signal of the PCMTD1-Elongins-[^{3}H] AdoMet

appeared to decrease, there was still significant signal. In the presence of AdoHcy, [^{3}H] AdoMet binding to both PCMT1 and PCMTD1-Elongins appears to be abolished. These experiments suggest that the PCMTD1-Elongins complex is capable of specifically binding the methyltransferase cofactor AdoMet.

No signal was detected in the lanes 2, 5, or 8, corresponding to PCMTD1 alone. This may indicate that the Elongins are necessary for folding of the PCMTD1 protein, and it does not properly form the AdoMet binding site in their absence, but it also cannot be ruled out that the concentration of the PCMTD1 protein alone is too low in this instance to view binding. To test whether or not the C-terminus is necessary for AdoMet binding, or if the concentrations of the full-length PCMTD1 protein presented issues, a truncated construct corresponding to residues 1-231, which include the region of the PCMTD1 protein that is homologous to PCMT1 including all AdoMet binding motifs, was purified (designated as PCMTD1Δ232-357). This truncated product expressed and purified extremely well. In Figure 2C we show that this truncated PCMTD1 form, now at the same concentration as PCMT1, still does not crosslink with the [^{3}H] AdoMet. Thus it appears that even with all of the AdoMet binding motifs present in PCMTD1, the C-terminus is necessary for [^{3}H] AdoMet binding under the conditions used here, and that the Elongins likely play a crucial role in ordering this C-terminal region.

The PCMTD1-Elongins complex does not display methyltransferase activity

The conservation of the AdoMet and L-isoaspartyl-binding sites from PCMT1 to PCMTD1 suggests that this protein may retain similar L-isoaspartyl-methylation activity. To probe for L-isoaspartyl methylation, we took advantage of the labile nature of the methyl-ester that is created during the reaction. Methylation of L-isoaspartyl residues with [^{3}H] AdoMet generates a labile [^{3}H] methyl ester, which can be released as [^{3}H] methanol upon the addition of base, and detected by a scintillator. Using both a known peptide and protein substrate of PCMT1, we tested for L-isoaspartyl methylation by PCMTD1-Elongins (**Fig. 3A**). The canonical methyltransferase PCMT1 alone (red) is able to methylate L-isoaspartate within both the peptide

(KASA(isoD)LAKY) and the protein (ovalbumin). Incubation of substrates with the PCMTD1-Elongins, however, did not show any methylation of L-isoaspartate above the enzyme-alone background (blue columns).

It is possible that the KASA(isoD)LAKY peptide and the ovalbumin protein are not preferred substrates for PCMTD1, or that it is a methyltransferase, but it is not methylating L-isoaspartate residues. To test these hypotheses, PCMT1 and PCMTD1-Elongins were incubated with both wild-type and *Pcmt1* knockout mouse brain lysates with [^{3}H] AdoMet. All types of methylation using the [^{3}H] AdoMet can then be detected by SDS-PAGE separation of the radiolabeled lysates, and exposure to film. In the lanes containing extract and PCMT1, significant signal as a result of methylation was observed, and as expected the signal was higher in the *Pcmt1* knockout lysate, as more L-isoaspartate substrates have been allowed to accumulate (lanes 4 and 7; **Fig. 3B**). However, in lanes 5 and 8 in which the PCMTD1-Elongins were incubated with lysates, there is no signal observed above the lysate-alone lanes (3 and 6). Thus, while the PCMTD1-Elongins complex is able to bind AdoMet, it does not exhibit any type of methyltransferase activity.

The PCMTD1-Elongins associate with Cullin 5 and lower detectable L-isoaspartate levels in lysates

Mahrour et al., performed pull down experiments using FLAG-tagged PCMTD2 as bait in HEK293 cells and demonstrated that Elongins B and C, as well as Cullin 5 are associated (Cul5) (14). PCMTD1 was also able to be isolated when FLAG-Cul5 was used as bait. To see if this association could be reconstituted *in vitro*, Cul5 was co-expressed as a separate N-terminal and C-terminal region alongside Rbx2 similar to previous recombinant expression schemes for Cullin proteins (15, 16). Purified PCMTD1-Elongins and the purified Cul5-Rbx2 complexes were sequentially loaded onto Ni-resin, and co-purified by immobilized affinity metal chromatography (IMAC). To verify the association, the IMAC fractions containing the proteins were separated by gel filtration. Gel filtration profiles of the IMAC-purified PCMTD1-Elongins alone were compared

to PCMTD1-Elongins-Cul5-Rbx2 and protein standards runs (**Fig. 4A**). The A_{280} trace coupled with SDS-PAGE revealed that while the PCMTD1-Elongins (~66,000, red asterisk) alone eluted between the 158,000 Da bovine γ-globulin and the 44,000 ovalbumin standards, the PCMTD1-Elongins-Cul5-Rbx2 complex (blue asterisk) eluted before the γ-globulin standard, in accordance with the ~170,000 Da molecular weight of the whole complex. Electron microscopy revealed circular species ~30 nm in diameter on the grid, which may correspond to a group of several PCMTD1-Elongins-Cul5-Rbx2 complexes, as a single complex would be expected to be approximately 10-15 nm (**Fig. 4B**).

Due to the ability of PCMTD1 to complex to cullin-RING ligase components, the conservation of the L-isoaspartyl-recognition motifs, and in light of its lack of methyltransferase activity, it is plausible that the protein is able to maintain L-isoaspartate levels in the cell through mechanisms other than methylation. To investigate this possibility, PCMTD1-Elongins were mixed with wild-type and *Pcmt1* knockout brain lysate with and without ubiquitin ligase components (Ub. reaction components), including E1, E2, MgATP, ubiquitin, and Cul5/Rbx2. Purified PCMT1 and [^{3}H] AdoMet were then added to label remaining L-isoaspartate residues. These radiolabeled extracts were separated by SDS-PAGE and signal from radiolabeled L-isoaspartate was detected on film (**Fig. 5A**). As seen in Figure 3, there were significantly higher levels of L-isoaspartate signal in the Δ*Pcmt1* lysate than the wild-type lysate as expected. Strikingly, the combination of PCMTD1-Elongins with the Ub. reaction components decreased the L-isoaspartate signal throughout the lane by approximately 50% as quantified by densitometry, while the Ub. reaction components alone did not appear to lower signal (**Fig. 5B**). The lane in which PCMTD1-Elongins were incubated alone with the extract decreased overall signal by ~25%. Thus the PCMTD1-Elongins in combination with cullin-RING ligase components are able to lower PCMT1-detectable levels of L-isoaspartate *in vitro*. Interestingly, no large changes in the coomassie staining of the gel were seen, which might have been expected if there were increases in polyubiquitination by an active PCMTD1-Elongins-cullin-

RING ligase complex (**Fig. 5A**). However, the L-isoaspartate modification is typically low stoichiometry, and femtomolar amounts can be detected with the high sensitivity of the PCMT1-radiolabeled fluorography, and thus may not correlate to large shifts detectable by coomassie staining.

To test the effects of individual cullin-RING components on this system, we again incubated the PCMTD1-Elongins with Δ*Pcmt1* lysate, and sequentially removed the E1/E2 enzymes, Cul5, and ubiquitin (**Fig. 6A**). Strikingly, reactions that lacked each of these components had decreased levels of L-isoaspartate signal comparable to or lower than that of the lane containing all of the cullin-RING components and PCMTD1-Elongins. These results may support the hypothesis that *in vitro* when the PCMTD1-Elongins are able to associate with all cullin-RING components, and ubiquitination allowed to take place, the released ubiquitinated substrate may still be recognized and methylated by PCMT1 and [^{3}H] AdoMet before being degraded by proteasomal proteins present in the extract. However, when components of the cullin-RING complex are lacking the PCMTD1 protein may not release the unmodified substrate, fully occluding PCMT1 radiolabling (**Fig. 6B**).

The PCMTD1-Elongins catalyze the ubiquitination of carbonic anhydrase

The association of the PCMTD1-Elongins with cullin-RING ligase components suggests that the PCMTD1 protein may serve as an E3 substrate adaptor protein to recruit substrates for active ubiquitination. The high sequence homology between the N-terminus of PCMTD1 and the PCMT1 enzyme suggests that there may be some overlap in substrate specificity between these proteins (**Fig. 1**). To evaluate the hypothesis that PCMTD1 may facilitate ubiquitination of a substrate protein, we used the protein carbonic anhydrase which has previously been shown to be an endogenous substrate for the PCMT1 protein in mouse brain (17). A solution of carbonic anhydrase was allowed to age for two weeks at 37 °C in order to accumulate L-isoAsp residues. Unaged and aged carbonic anhydrase were then tested as substrates for PCMTD1-dependent ubiquitination as probed by recognition by an anti-ubiquitin antibody (**Fig. 7**). The

accumulation of ubiquitinated species was time dependent, as shown by the limited amount of signal observed in the 0 minute reactions lanes 1 and 4. The ubiquination detected increased to the greatest levels in the 30 minute reactions in which all E3 ligase components were present along with carbonic anhydrase and the PCMTD1-Elongins (lanes 2 and 5). Strikingly, when only the PCMTD1-Elongins are removed, the overall signal significantly decreases (lanes 3 and 6). The background ubiquitination seen in lanes 3 and 6 corresponds to Cul5 autoubiquitination seen in other experiments (data not shown). These results demonstrate that PCMTD1 can facilitate the ubiquitination of a protein substrate.

Given the ability of PCMTD1 to interact with the carbonic anhydrase substrate, it supports the hypothesis that PCMTD1 is recognizing similar motifs as PCMT1. This would further indicate that the PCMTD proteins are likely to recognize L-isoAsp residues in a similar manner as PCMT1. Interestingly, there were no significant differences observed between the unaged and aged solutions of carbonic anhydrase, despite the fact that the aged population should contain a higher number of L-isoAsp residues (**Fig. 7A**). The levels of L-isoAsp within the unaged and aged carbonic anhydrase were analyzed and revealed that while there was a rise increase in the modification within the aged population it was at only 3% stoichiometry (**Fig. 7B**). It is clear that the PCMTD1 protein is able to recognize at least a portion of the unaged population in order to catalyze ubiquitination. However, perhaps we would see much greater signal with a substrate containing high levels of L-isoAsp.

DISCUSSION:

In Mahrour et al., 2008 it was suggested that the PCMTD proteins may represent a link between methylation and ubiquitination based on the conservation of the PCMT1 L-isoaspartyl and AdoMet binding motifs. Here we do show that PCMTD1 is capable of associating with AdoMet, and that this binding may be dependent on certain cullin-RING ligase components *in vitro* (**Fig. 3**). However, even in the presence of a full complement of ligase components, [^{3}H] AdoMet, and mouse brain lysate, the PCMTD1 protein does not raise levels of detectable

methylation (**Fig. 4 and 5**). Instead, when the PCMTD1-Elongins are incubated with ligase components and extract, levels of PCMT1-methylated L-isoaspartate decrease (**Fig. 5**).

The majority of PCMT1 methyltransferases recognize both L-isoaspartyl and D-aspartyl residues, although some species variants, like the *Thermatoga maritima* PCMT enzyme can only recognize L-isoaspartate (18). Structural analyses of these enzymes have revealed that these enzymes primarily bind the substrate by interactions with the L-isoaspartate residue and the residues C-terminal to the modified amino acid (19). A crystal structure of the *Pyrococcus furiosus* PCMT1 enzyme with an L-isoaspartate containing peptide showed the Ser75 (Ser59 in human PCMT1) of Pre-motif I interacting with the carboxyl side chain of the L-isoaspartate residue. The Pre-motif I is highly similar in the PCMTD proteins, conserving the serine residue, which would interact directly with an L-isoaspartate side chain. The remaining residues do not appear likely to produce major clashes with an L-isoaspartate substrate, though it is difficult to predict how the structural accommodation of a some mutations may indirectly affect the active site. The PCMT1 structures also reveal that these enzymes conserve three charged amino acid residues buried in the protein's interior in or near the active site (20). Within the human enzyme, these are residues Arg36, Asp83, and Asp109. In both PCMTD proteins, the arginine is conserved, however, both aspartate residues are replaced by asparagine residues (**Fig. 1**). Asp109 of Post-motif I in particular represents a highly conserved site in AdoMet dependent methyltransferases, wherein an aspartate or glutamate residue hydrogen bonds with the ribose hydroxyl groups of the AdoMet (21, 22). The replacement of the aspartate with asparagine could orient the AdoMet cofactor in a similar position.

Thus sequence comparisons do not yield obvious explanations for the lack of methyltransferase activity within the PCMTD proteins. It is possible that a necessary cofactor is still missing, as we say in Figure 2 that the PCMTD1 protein did not purify and bind AdoMet in the absence of Elongins B and C. However, the results presented here support an alternate function for the PCMTD proteins in which these proteins maintain the isoaspartyl-binding motifs

of PCMT1 in order to recognize this age-related protein modification within proteins to catalyze their ubiquitination and eventual degradation. This is supported by the association of PCMTD1 with both Elongins B and C, as well as Cul5 and Rbx2 (**Fig. 2 and Fig. 4**) and the lowered levels of detectable L-isoaspartate in incubations with *Pcmt1* knockout mouse brain lysates (**Fig. 5**). Ubiquitination assays within extracts yield high background, making it difficult to probe for active ubiquitination by the PCMTD1-Elongins-Cul5 complex. However, experiments using a known PCMT1 substrate, carbonic anhydrase, revealed that the PCMTD1-Elongins can facilitate ubiquitination (Fig. 7). As the levels of L-isoAsp within the substrate were low, it is difficult to predict what region of the substrate the PCMTD1 adaptor protein is recognizing. Defining this interaction will require further experimentation, and will be crucial for our understanding of this newly identified E3 ligase adaptor protein.

The recognition of an age-related protein modification by an E3 ligase represents a unique discovery in the context of protein turnover. Certain E3 ligases have been found to recognize modified forms, such as the F-box proteins of the SCF E3 ligase family, which requires phosphorylation of substrates, and the von Hippel-Lindau protein, which recognizes hydroxylated hypoxia-inducible factor-1 (23, 24). However, to the authors' knowledge, no E3 ligases have been identified to specifically recognize aged protein for ubiquitination and degradation. This finding represents another layer of complexity to cellular protein homeostasis.

METHODS:

Reagents and plasmids

Full-length (1-357) and truncated (1-231) PCMTD1 were expressed as His-tagged constructs in pMAPle vectors designed and generated in the UCLA DOE Protein Expression Technology Core (PETC). The vector sequences include a pBR322 origin, lacI, MBP, TEV protease site, C-terminal His6, and co-express TEV protease. The Elongins were expressed on the ELONGIN BC plasmid (XX01TCEB1A-c001), which was a gift from Nicola Burgess-Brown (Addgene plasmid # 110274; http://n2t.net/addgene:110274 ; RRID:Addgene_110274). The Uba1 or E1

ubiquitin conjugating enzyme was expressed in pET28-mE1 - a gift from Jorge Eduardo Azevedo (Addgene plasmid # 32534; http://n2t.net/addgene:32534 ; RRID:Addgene_32534, ref. 25). The E2 enzyme was expressed in pET15-UbE2D1 was a gift from Wade Harper (Addgene plasmid # 15782; http://n2t.net/addgene:15782 ; RRID:Addgene_15782; ref. 26). In all ubiquitination and methylation assays, full-length untagged Cul5/Rbx2 purchased from Ubiquigent was utilized (catalog # 63-1002-025). For Cul5/Rbx2 co-purification with PCMTD1 fragments for the Cul5 N-terminus (1-384) and His-Rbx2-Cul5-C-terminus (385-780) were cloned into pET24+ and PACYDuet vectors, respectively, and co-transformed into *E. coli* BL21 cells. The plasmid encoding the recombinant human l-isoaspartyl protein methyltransferase (rhPIMT) with an N-terminal polyhistidine tag was a generous gift of Dr. Bruce Downie. Ubiquitin was purchased from Boston Biochem (catalog # U-100H).

Sequence alignment

Multiple sequence alignment was done with T-Coffee (27). UniProt accession numbers P22061 (PCMT1), Q96MG8 (PCMTD1), and Q9NV79 (PCMTD2).

Ethics Statement

This study was performed in accordance with animal use protocols approved by the UCLA Animal Research Committee (Protocol 1993-109-64). Mice were scheduled to be euthanized if they met any early removal criteria (kyphosis, lack of grooming behavior). However, this did not occur with any of the animals in our study.

Animal Husbandry

Pcmt1−/− animals were generated through breeding of Pcmt1+/− animals and maintained as reported previously (10, 28). These animals have been interbred for twenty years to obtain a genetically homogeneous population. Pcmt1−/− and Pcmt1+/+ offspring were used in this study. Mice were kept on a 12-hour light/dark cycle and allowed ad libitum access to water and NIH-31 7013 pellet chow (18% protein, 6% fat, 5% fiber, Harlan Teklad, Madison, WI).

Preparation of wild-type and Pcmt1 *knockout mouse tissue lysates*

Tissues were a kind gift from Dr. Jonathan Lowenson. Briefly, 52-day-old wild-type and *Pcmt1* knockout mice were euthanized in a CO_2 chamber. Brain tissue was removed and weighed, and 5 mL/gram tissue of ice-cold lysis buffer (250 mM sucrose, 10 mM Tris-HCl, pH 7.4,1 mM ethylenediaminetetraacetic acid (EDTA), 1mM phenylmethylsulfonyl fluoride (PMSF), 1 Roche protease inhibitor cocktail tablet) was added. Tissues were homogenized with a Fisher LR400A Lab-Stirrer at approximately 300 rpm, then spun at 20,000 *g* for 20 min at 4 °C. The supernatant was removed as the soluble extract and stored at -20 °C until use.

Recombinant protein expression and purification

Recombinant proteins were expressed in *E. coli* BL21 strains. PCMTD1-Elongins were expressed for 3 h at 37 °C and induced with 0.5 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG). E1, E2, and PCMTD1-trunc enzymes were expressed overnight at 18 °C, with 1 mM IPTG. Cul5-NTR and Cul5-CTR/Rbx2 plasmids were co-expressed 18 °C, with 0.5 mM IPTG.

To maintain consistent conditions, all enzymes were purified in the same buffers with one immobilized metal affinity chromatography step followed by size exclusion chromatography. Following expression, cells were spun down for 15 min at 5,000 *g* at 4 °C and frozen at -80 °C until lysis and purification. Thawed cells were resuspended in 5 mL/g pellet of lysis buffer (50 mM HEPEs, pH 7.6, 300 mM NaCl, 5% glycerol, 1 mM β-mercaptoethanol (βME), 1 mM PMSF, and 1 EDTA-free Pierce protease inhibitor tablet per 50 mL). Lysis was performed by three passes through an Avestin Emulsiflex at 15,000 psi with incubations on ice to minimize temperature of lysate. Lysates were then spun at 13,000 rpm for 50 min at 4 °C. Filtered lysates were then loaded onto three 5 mL HisTrap HP columns equilibrated with wash buffer (50 mM HEPEs, pH 7.6, 150 mM NaCl, 5% glycerol, 20 mM imidazole, 1 mM βME) on a Bio-Rad Biologic FPLC system. Proteins were eluted from the column over two steps, first a linear gradient from 0-100% elution buffer (50 mM HEPEs, pH 7.6, 150 mM NaCl, 5% glycerol, 300 mM imidazole, 1 mM βME) over 60 min, followed by 100% elution buffer wash for 30 min at 1 mL/min. Fractions containing the purified protein were pooled and were loaded onto a HiPrep

16/60 Sephacryl S-200 HR gel filtration column equilibrated with 50 mM HEPES, pH 7.6, 150 mM NaCl, 1 mM βME. Gel filtration as carried out over 200 mL at 0.4 mL/min. Finally, fractions containing purified proteins were pooled, 5% final glycerol was added, and proteins were concentrated using 10 kDa MWCO Amicon centrifugal filters prior to storage at -20 °C. Protein concentrations were determined by a Lowry assay after protein precipitation with 10% trichloroacetic acid (29).

[^{3}H] AdoMet:protein ultraviolet crosslinking experiments

In a final volume of 60 µL 50 mM Tris-HCl, pH 7.5, 3.85 µM protein was mixed with 0.5 µM *S*-adenosyl-L-[methyl-^{3}H] methionine ([^{3}H] AdoMet; PerkinElmer Life Sciences; 75-85 Ci/mml, 0.55 mCi/ml in 10 mM H2SO4:ethanol (9:1, v/v)). Where indicated in the figure legend, 0.5 mM final of either non-radioactive *S*-adenosylhomocysteine or adenosine triphosphate was added. Reactions were placed into NUNC 96 well clear bottom plates and exposed to 254 nm ultraviolet light at 4 °C for 1 h. 15 µL 5x SDS-loading dye was added to quench the reaction, and samples were heated at 100 °C for 3 min. Samples were then separated by SDS-PAGE and visualized by coomassie staining.

Determination of L-isoaspartate-methylation levels by the methanol vapor diffusion assay

PCMT1 was used as a positive control to quantify L-isoaspartate levels in the KASA(isoD)LAKY peptide substrate and the ovalbumin protein substrate (Sigma A5503). In a final volume of 100 µL10 pmol of PCMT1 or 15 pmol PCMTD1-Elongins were incubated with 25 pmol KASA(isoD)LAKY or 500 pmol Ovalbumin (typically ~6% isomerized). Final concentrations in the reactions included 135 mM Bis-Tris-HCl, pH 6.4, and 10 µM S-adenosyl-l-[methyl3H]methionine ([3H]AdoMet) (prepared by a 1600-fold isotopic dilution of a stock of 72 Ci/mmol [3H]AdoMet (PerkinElmer Life Sciences, NET155H00) with nonisotopically labeled AdoMet (p-toluenesulfonate salt; Sigma-Aldrich A2408)). The reaction was stopped by adding 10 µL of 2 M sodium hydroxide, and 100 µL of the 110 µL mixture was transferred to a 9 by 2.5 cm piece of folded thick filter paper (Bio-Rad; catalog number 1650962) wedged in the neck of a

20-mL scintillation vial above 5 mL scintillation reagent (Safety Solve, Research Products International, catalog number 121000), tightly capped, and incubated at room temperature. After 2 h, the folded filter papers were removed, the caps replaced, and the vials were counted thrice for 5 minutes each in a Beckman LS6500 scintillation counter. Background radioactivity in a reaction containing no substrate was determined by incubating the recombinant human PCMT1 or PCMTD1-Elongins, 135 mM Bis-Tris-HCl buffer, and 10 μM [^{3}H] AdoMet as described above. Samples were analyzed in triplicate.

SDS-PAGE fluorography for the analysis of methyltransferase activity

25 μg wild-type and *Pcmt1* knockout extracts were analyzed in a 30 μL reaction volume with final concentrations of 74 mM Bis-Tris-HCl, pH 6.4, 6 μg recombinant human PCMT1 or PCMTD1-Elongins, 0.3 μM S-adenosyl-L-[methyl-^{3}H] methionine (PerkinElmer Life Sciences; 75-85 Ci/mml, 0.55 mCi/ml in 10 mM H_2SO_4:ethanol (9:1, v/v)), and incubated for 2 h at 37 °C. The reaction was stopped by adding 5 μL SDS-PAGE loading buffer (250 mM Tris-HCl, pH 6.8, 10% (w/v) SDS, 50% (v/v) glycerol, 5% (v/v) β-mercaptoethanol, and 0.05% (w/v) bromophenol blue). Samples were heated at 100 °C for 3 min and separated on a 4-20%, 10 well ExpressPlus PAGE gel (Genscript, catalog # M42010) at 140 V for 1 h. Gels were stained with Coomassie (0.1% (w/v) Brilliant Blue R-250, 10% (v/v) glacial acetic acid, and 50% (v/v) methanol) for 1 h and destained with 10% (v/v) acetic acid and 15% (v/v) methanol. For fluorography, gels were subsequently incubated with EN3HANCE (PerkinElmer Life Sciences, catalog number 6NE9701) for 1 h, incubated in water for 30 min, and dried before the gels were exposed to film (Denville Scientific, 8 × 10-inch Hyblot CI) for the length of time designated in the figure legends at −80 °C.

SDS-PAGE fluorography of ubiquitination assays

7.7 μM PCMTD1-Elongins, 70 nM E1, 3 μM E2, 160 nM Cul5/Rbx2, were incubated in 20 mM Tris, pH 7.5, 5 mM KCl, 1.5 mM MgCl2, 1mM DTT with 50 μM ubiquitin and 1.6 mM ATP for 1 h at 37 °C in a final volume of 30 μL. After 1 h, 30 μL of 6 μg PCMT1, 0.3 μM [^{3}H] AdoMet, and

200 mM Bis-Tris-HCl, pH 6.4 were added to the reactions and incubated for an additional 2 h at 37 °C. The reaction was stopped by adding 15 µL SDS-PAGE loading buffer (250 mM Tris-HCl, pH 6.8, 10% (w/v) SDS, 50% (v/v) glycerol, 5% (v/v) β-mercaptoethanol, and 0.05% (w/v) bromophenol blue). Samples were heated at 100 °C for 3 min and separated on a 4-20%, 10 well ExpressPlus PAGE gel (Genscript, catalog # M42010) at 140 V for 1 h. Staining, enhancing, and fluorography proceeded as described in the section above. Signal from fluorographs was quantified by densitometry using ImageJ (30).

Western blot of ubiquitination assays

7.3 pmol of Cul5/Rbx2 was neddylated using the Enzo NEDDylation Kit (catalog # BML-UW0590). 182 nM final concentration neddylated Cul5/Rbx2 was incubated with 2.84 µM PCMTD1-Elongins, 52.5 nM E1, 2.25 µM E2, 118.4 µM cold AdoMet, and 13.4 µM carbonic anhydrase in 20 mM Tris, pH 7.5, 5 mM KCl, 1.5 mM MgCl2, 1mM DTT with 37.5 µM ubiquitin and 1.2 mM ATP for 0 or 30 min at 37 °C in a final volume of 40 µL. The reaction was stopped by adding 10 µL SDS-PAGE loading buffer (250 mM Tris-HCl, pH 6.8, 10% (w/v) SDS, 50% (v/v) glycerol, 5% (v/v) β-mercaptoethanol, and 0.05% (w/v) bromophenol blue). Samples were heated at 100 °C for 3 min and separated on a 4-20%, 10 well ExpressPlus PAGE gel (Genscript, catalog # M42010) at 140 V for 1 h. The gel was then transferred to a PVDF membrane at 30 V for 1 h. Ubiquitination was detected on the blot with a 1:500 dilution of an anti-ubiquitin primary antibody (Mono- and polyubiquitinylated conjugates monoclonal antibody (FK2), Enzo catalog # BML-PW8810-0100), followed by incubation with a 1:10,000 dilution of rabbit anti-mouse secondary antibody. Bound antibody was visualized using the Amersham ECL western blotting detection reagent (GE Healthcare product # RPN2106). Film was exposed to the blot for 10 s.

References:

1. Santos, A.L., Lindner, A.B. Protein Posttranslational Modifications: Roles in Aging and Age-Related Disease. *Oxid. Med. Cell. Longev*. Epub 5716409 (2017).

2. Boschi-Muller, S., Gand, A., Branlant, G. The methionine sulfoxide reductases: Catalysis and substrate specificities. *Arch. Biochem. Biophys*. **474**, 266-273 (2008).
3. Szwergold, B., Manevich, Y., Payne, L., Loomes, K. Fructosamine-3-kinase-related-protein phosphorylates glucitolamines on the C-4 hydroxyl: novel substrate specificity of an enigmatic enzyme. *Biochem. Biophys. Res. Commun*. **361**, 870-875 (2007).
4. McFadden, P.N., Clarke, S. Conversion of isoaspartyl peptides to normal peptides: Implications for the cellular repair of damaged proteins. *Proc. Natl. Acad. Sci. USA*. **84**, 2595-2599 (1987).
5. Chondrogianni, N., Petropoulos, I., Grimm, S., Georgila, K., Catalgol, B., Friguet, B., Grune, T., Gonos, E.S. Protein damage, repair, and proteolysis. *Mol. Aspects Med*. **35**, 1-71 (2014).
6. Berndsen, C.E., Wolberger, C. New insights into ubiquitin E3 ligase mechanism. *Nat. Struct. Mol. Biol*. **21**, 301-307 (2014).
7. Deshaies, R.J., Joazeiro, C.A. RING domain E3 ubiquitin ligases. *Annu. Rev. Biochem*. **78**, 399-434 (2009).
8. Linossi, E.M., Nicholson, S.E. The SOCS box-adapting proteins for ubiquitination and proteasomal degradation. *IUBMB Life*. **64**, 316-323 (2012).
9. Okumura, F., Joo-Okumura, A., Nakatsukasa, K., Kamura, T. The role of cullin 5-containing ubiquitin ligases. *Cell. Div*. **11**, 1 (2016).
10. Lowenson, J.D., Kim, E., Young, S.G., Clarke, S. Limited accumulation of damaged proteins in l-isoaspartyl (D-aspartyl) O-methyltransferase-deficient mice. *J. Biol. Chem*. **276**, 20695-20702 (2001).
11. Kamura, T., Sato, S., Haque, D., Liu, L., Kaelin, W.G., Conaway, R.C., Conaway, J.W. The Elongin BC complex interacts with the conserved SOCS-box motif present in members of the SOCS, ras, WD-40 repeat, and ankyrin repeat families. *Genes Dev*. **12,** 3872-3881 (1998).

12. Babon, J.J., Sabo, J.K., Soetopo, A., Yao, S., Bailey, M.F., Zhang, J.G., Nicola, N.A., Norton, R.S. The SOCS box domain of SOCS3: structure and interaction with the elonginBC-cullin5 ubiquitin ligase*. J. Mol. Biol*. **381**, 928-940 (2008).
13. Stebbins, C.E., Kaelin, W.G., Pavletich, N.P. Structure of the VHL-ElonginC-ElonginB complex: implications for VHL tumor suppressor function. *Science.* **284**, 455-461 (1999).
14. Mahrour, N., Redwine, W.B., Florens, L., Swanson, S.K., Martin-Brown, S., Bradford, W.D., Staehling-Hampton, K., Washburn, M.P., Conaway, R.C., Conaway, J.W. Characterization of Cullin-box sequences that direct recruitment of Cul2-Rbx1 and Cul5-Rbx2 modules to Elongin BC-based ubiquitin ligases. *J. Biol. Chem*. **283,** 8005-8013 (2008).
15. Kershaw, N.J., Laktyushin, A., Nicola, N.A., Babon, J.J. Reconstruction of an active SOCS3-based E3 ubiquitin ligase complex *in vitro*: Identification of the active components and JAK2 and gp130 as substrates*. Growth Factors*. **32,** 1-10 (2014).
16. Zheng, N., Schulman, B.A., Song, L., Miller, J.J., Jeffrey, P.D., Wang, P., Chu, C., Koepp, D.M., Elledge, S.J., Pagano, M., Conaway, R.C., Conaway, J.W., Harper, J.W., Pavletich, N.P. Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. *Nature*. **416**, 703-709 (2002).
17. Ichikawa, J.K., Clarke, S. A highly active protein repair enyme from an extreme thermophile: the L-isoaspartyl methyltransferase from *Thermatoga maritime*. *Arch. Biochem. Biophys.* **358**, 222-231 (1998).
18. Zhu, J.X., Doyle, H.A., Mamula, M.J., Aswad, D.W. Protein Repair in the Brain, Proteomic Analysis of Endogenous Substrates for Protein L-Isoaspartyl Methyltransferase in Mouse Brain. *J. Biol. Chem*. **281**, 33802-33813 (2006).
19. Griffith, S.C., Sawaya, M.R., Boutz, D.R., Thapar, N., Katz, J.E., Clarke, S., Yeates, T.O. Crystal structure of a protein repair methyltransferase from Pyrococcus furiosus with its L-isoaspartyl peptide substrate*. J. Mol. Biol*. **313**, 1103-1116 (2001).

20. Ryttersgaard, C., Griffith, S.C., Sawaya, M.R., MacLaren, D.C., Clarke, S., Yeates, T.O. Crystal structure of human L-isoaspartyl methyltransferase. *J. Biol. Chem*. **277**, 10642-10646 (2002).

21. Katz, J.E., Dlakić, M., Clarke, S. Automated identification of putative methyltransferases from genomic open reading frames. *Mol. Cell. Proteomics*. **2**, 525-540 (2003).

22. Kozbial, P.Z., Mushegian, A.R. Natural history of S-adenosylmethionine-binding proteins. *BMC Struct. Biol*. **5**, 19 (2005).

23. Winston, J.T., Strack, P., Beer-Romero, P., Chu, C.Y., Elledge, S.J., Harper, J.W. The SCFbeta-TRCP-ubiquitin ligase complex associates specifically with phosphorylated destruction motifs in IkappaBalpha and beta-catenin and stimulates IkappaBalpha ubiquitination in vitro. *Genes Dev*. **13**, 270-283 (1999).

24. Maxwell, P.H., Wiesener, M.S., Chang, G.W., Clifford, S.C., Vaux, E.C., Cockman, M.E., Wykoff, C.C., Pugh, C.W., Maher, E.R., Ratcliffe, P.J. The tumor suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. *Nature*. **399**, 271-275 (1999).

25. Carvalho, A.F., Pinto, M.P., Grou, C.P., Vitorino, R., Domingues, P., Yamao, F., Sá-Miranda, C., Azevedo, J.E. High-yield expression in Escherichia coli and purification of mouse ubiquitin-activating enzyme E1. *Mol. Biotechnol*. **51**, 254-261 (2012).

26. Jin, J., Li, X., Gygi, S.P., Harper, J.W. Dual E1 activation systems for ubiquitin differentially regulate E2 enzyme charging. *Nature*. **447**, 1135-1138 (2007).

27. Notredame, C., Higgins, D.G., Heringa, J. T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**, 205-217 (2000).

28. MacKay, K.B., Lowenson, J.D., Clarke, S.G. Wortmannin Reduces Insulin Signaling and Death in Seizure-Prone Pcmt1-/1 Mice. (2012) *PLoS One*. **7**, e46719.

29. Thorne, C.J.R. Techniques in Protein and Enzyme Biochemistry. 2-18 (1978).

30. Schneider, C. A.. Rasband, W. S., Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature methods*. **9**, 671-675 (2012).

FIGURES:

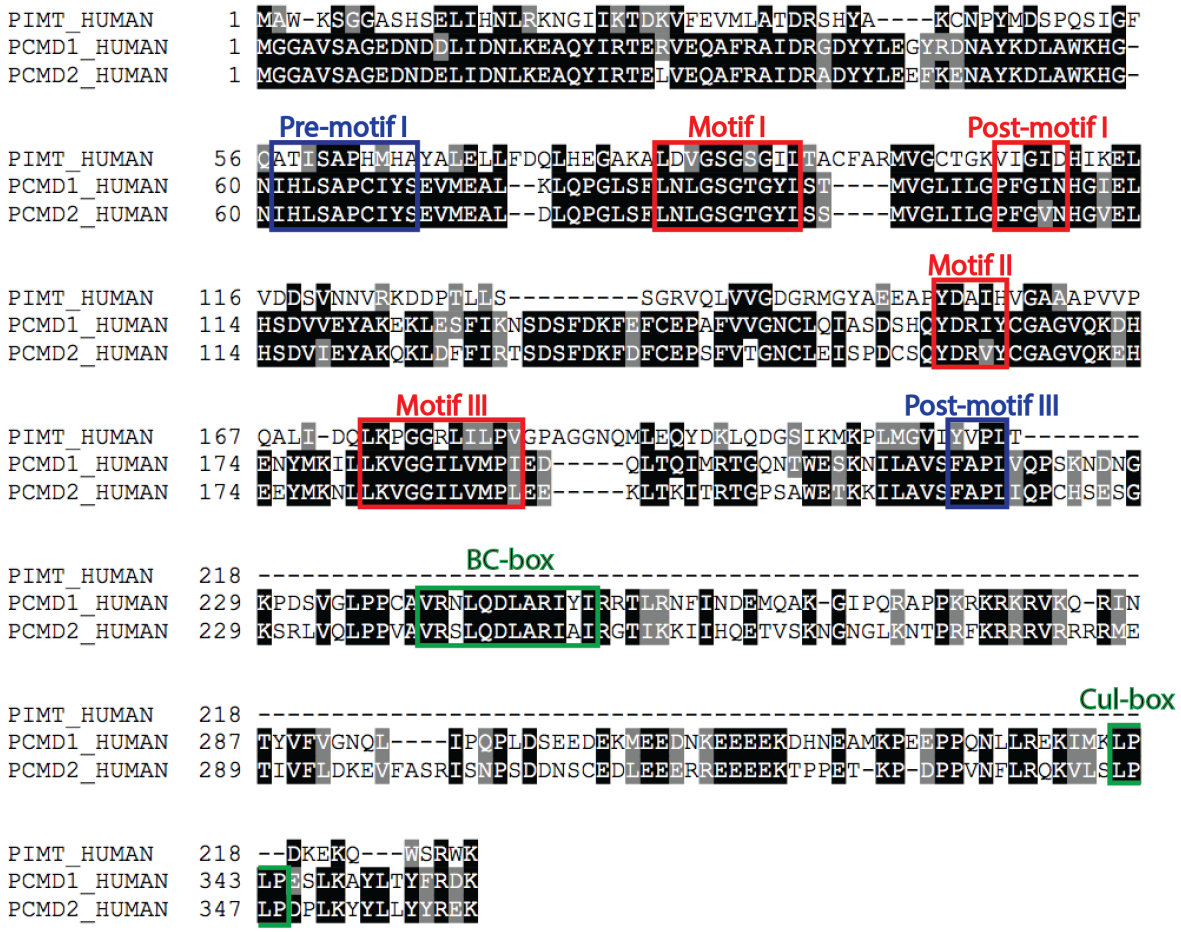


FIG. 1. **Sequence alignment of PCMT1 and PCMTD proteins.** Black and gray colors represent fully conserved and conservative mutations, respectively. Sequence outlined in blue corresponds to PCMT1 isoaspartyl-binding motifs, while sequence outlined in red represent PCMT1 AdoMet-binding motifs (20). Residues boxed in green in the PCMTD proteins comprise the BC-box and Cul-5 box binding motifs of the SOCS box domain.

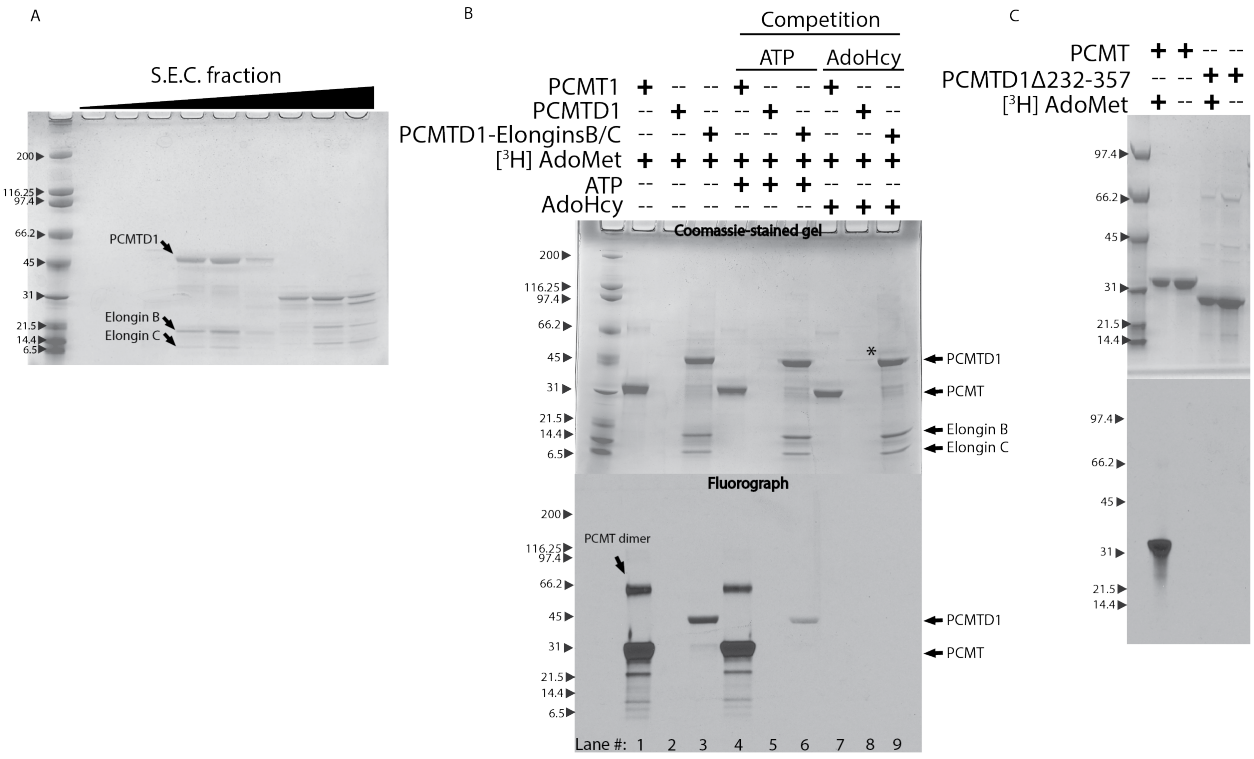


FIG. 2. **PCMTD1 associates with Elongins B and C, which aid in binding of the AdoMet cofactor.** A, PCMTD1 was co-expressed and purified with Elongins B and C as described in the Methods section. The coomassie gel shows fractions from size exclusion chromatography of pooled IMAC fractions. Untagged Elongin B (13 kDa) and Elongin C (10 kDa) co-elute with His-PCMTD1 (42 kDa). B, The coomassie-stained gel and fluorography show separation of [^{3}H] Adomet-crosslinked proteins. The asterisk in the upper panel indicates the His-PCMTD1 alone band. Film was exposed for 5 days. C, PCMT1 and truncated PCMTD1 proteins were UV-crosslinked with [^{3}H] Adomet and separated by SDS-PAGE and AdoMet binding was detected by fluorography as described in the Methods. Film was exposed for 3 days.

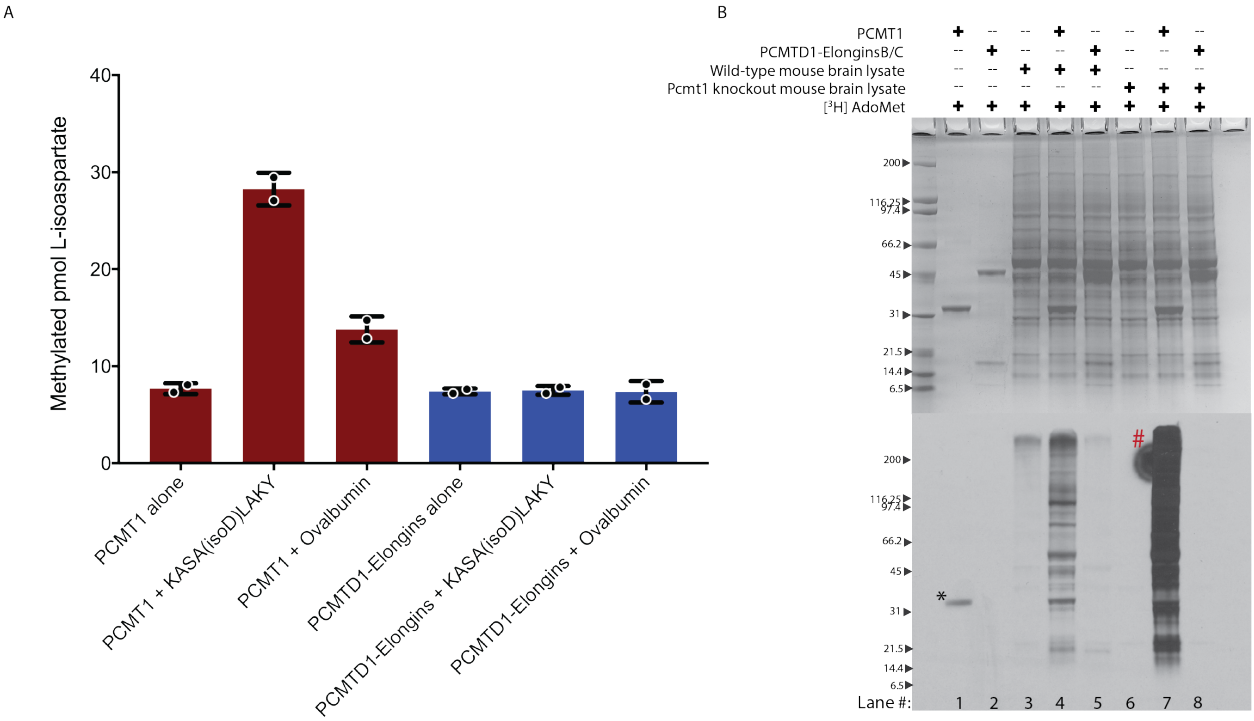


FIG. 3. The PCMTD1-Elongins do not display methyltransferase activity. A, L-isoaspartate methylation of the KASA(isoD)LAKY peptide and ovalbumin were detected by a methanol vapor diffusion assay as described in the Methods section. Assay was performed in duplicate. B, Total methylation was investigated by incubation of proteins with [^{3}H] AdoMet and mouse brain lysate. All [^{3}H]-labeled proteins were separated by SDS-PAGE. Proteins were visualized by coomassie staining and [^{3}H]-methylated proteins were detected by fluorography, by exposure to film for 6 days. The black asterisk (*) indicates PCMT1 automethylation, the red hashtag (#) shows contamination of the film by the unintended shift of the phosphorescent sticker used to mark orientation of the film.

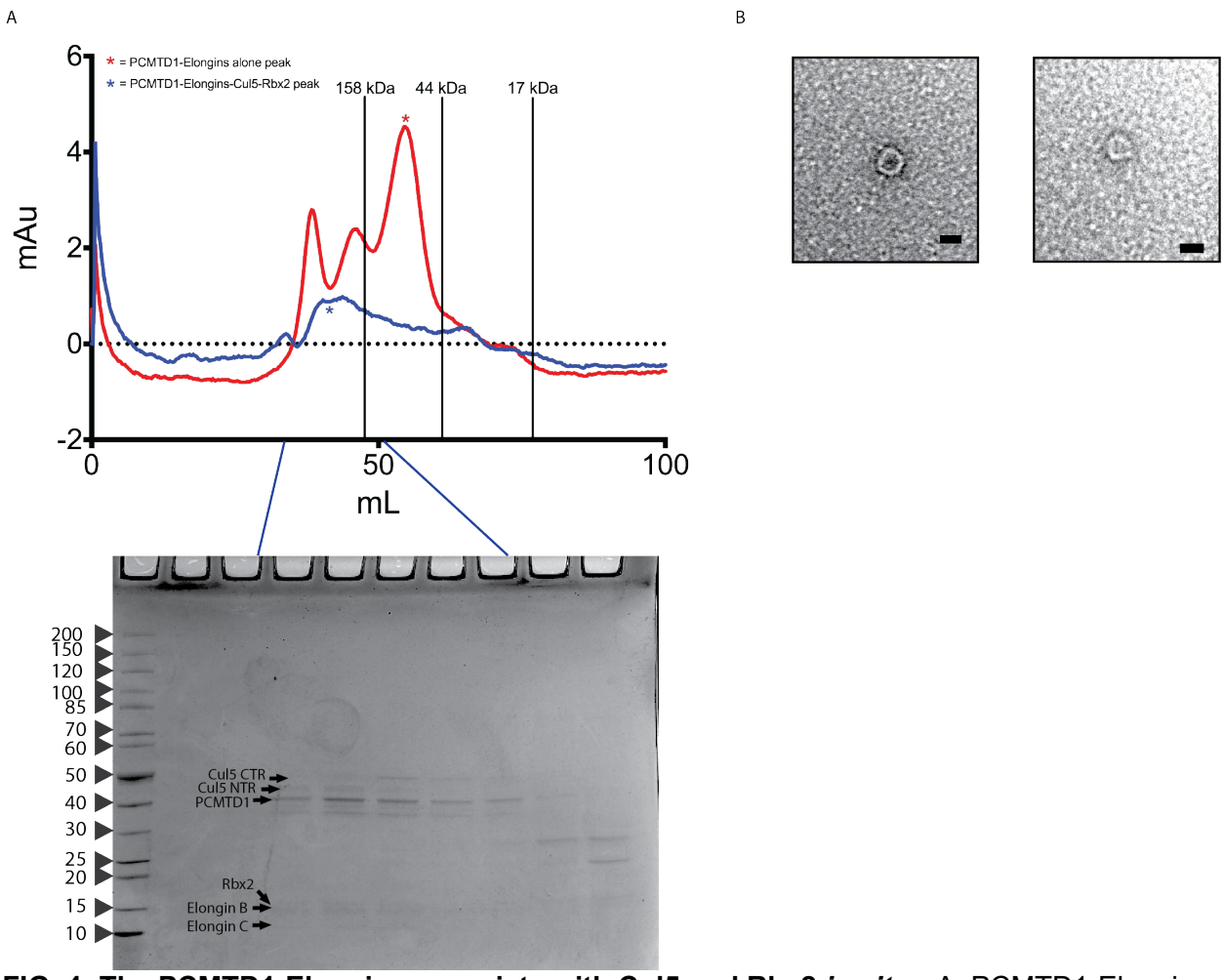


FIG. 4: The PCMTD1-Elongins associate with Cul5 and Rbx2 *in vitro*. A, PCMTD1-Elongins and Cul5-Rbx2 were separately purified as described in the Methods section, then purified proteins were mixed and co-purified by Ni-affinity chromatography, and gel filtration. The size exclusion chromatogram of PCMTD1-Elongins alone (red trace; A_{280}) is compared with the chromatogram of PCMTD1-Elongins-Cul5-Rbx2 (blue trace). Black lines indicate the volume where Bio-Rad gel filtration standards elute. Proteins were detected within these fractions by SDS-PAGE and coomassie staining. B, The purified complex was concentrated to ~0.2 mg/mL and electron micrographs were collected. Black scale bar represents 30 nm.

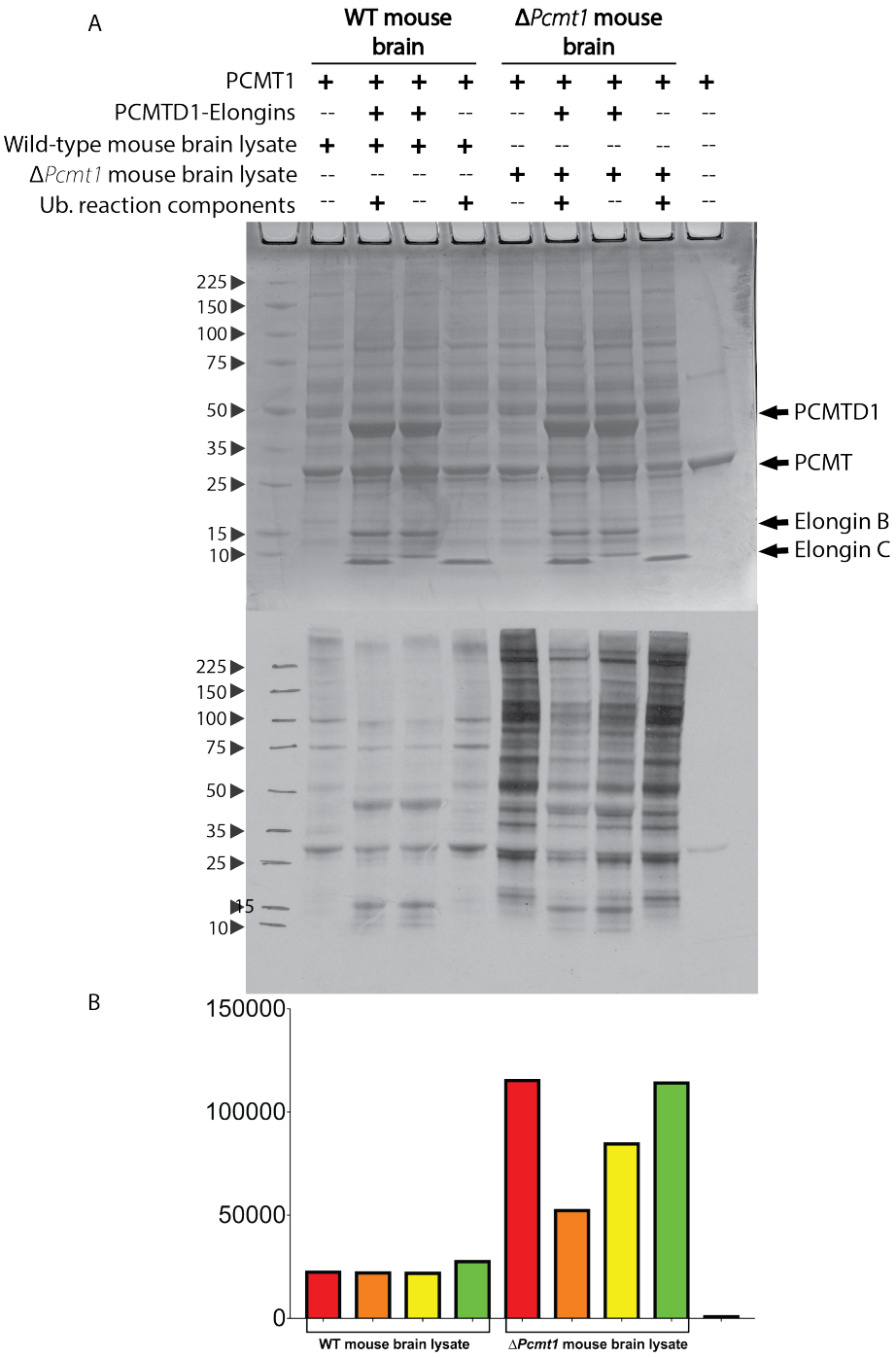


FIG. 5: **PCMTD1-Elongins coupled with cullin-RING ligase components lower detectable levels of L-isoaspartate *in vitro.*** A, PCMTD1-Elongins were incubated with mouse brain lysates and cullin-RING ligase components (Ub. reaction components) prior to L-isoaspartate detection by PCMT1 and [^{3}H] AdoMet. Radiolabeled lysates were separated by SDS-PAGE and

signal was detected by exposure to film for 5 days. B, Radioactive signal from fluorograph lanes was quantified by densitometry.

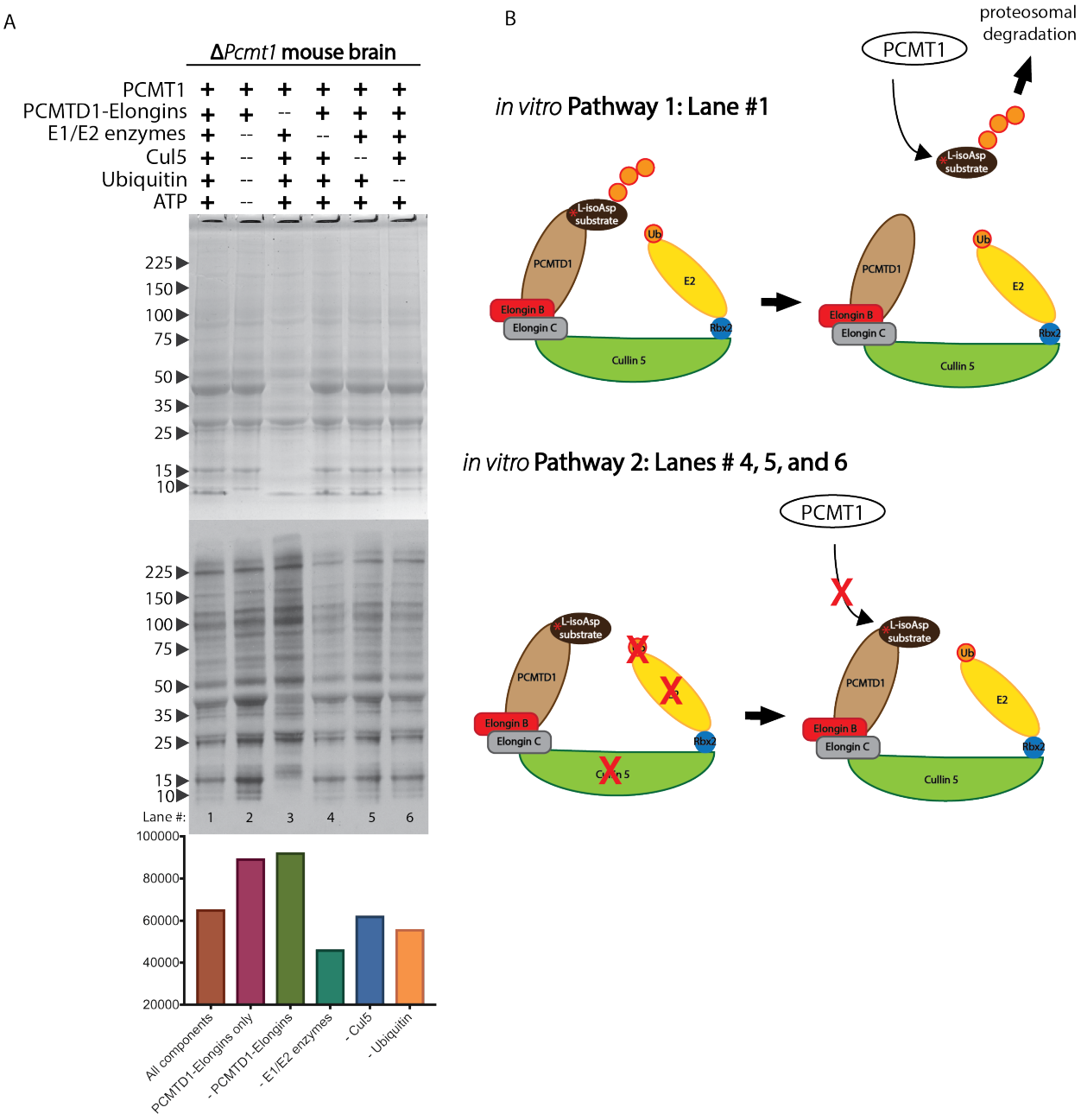


FIG. 6. ***In vitro*** **PCMTD1-Elongins can lower PCMT1-detectable L-isoaspartate without active ubiquitination**. A, PCMTD1-Elongins were incubated with mouse brain lysates and cullin-RING ligase components (Ub. reaction components) prior to L-isoaspartate detection by PCMT1 and [^{3}H] AdoMet. Single cullin-RING ligase components were removed in the designated lanes. Radiolabeled lysates were separated by SDS-PAGE and signal was detected

by exposure to film for 7 days. B, Model for the lowered levels of L-isoaspartate methylated by PCMT1 and [^{3}H] AdoMet *in vitro*.

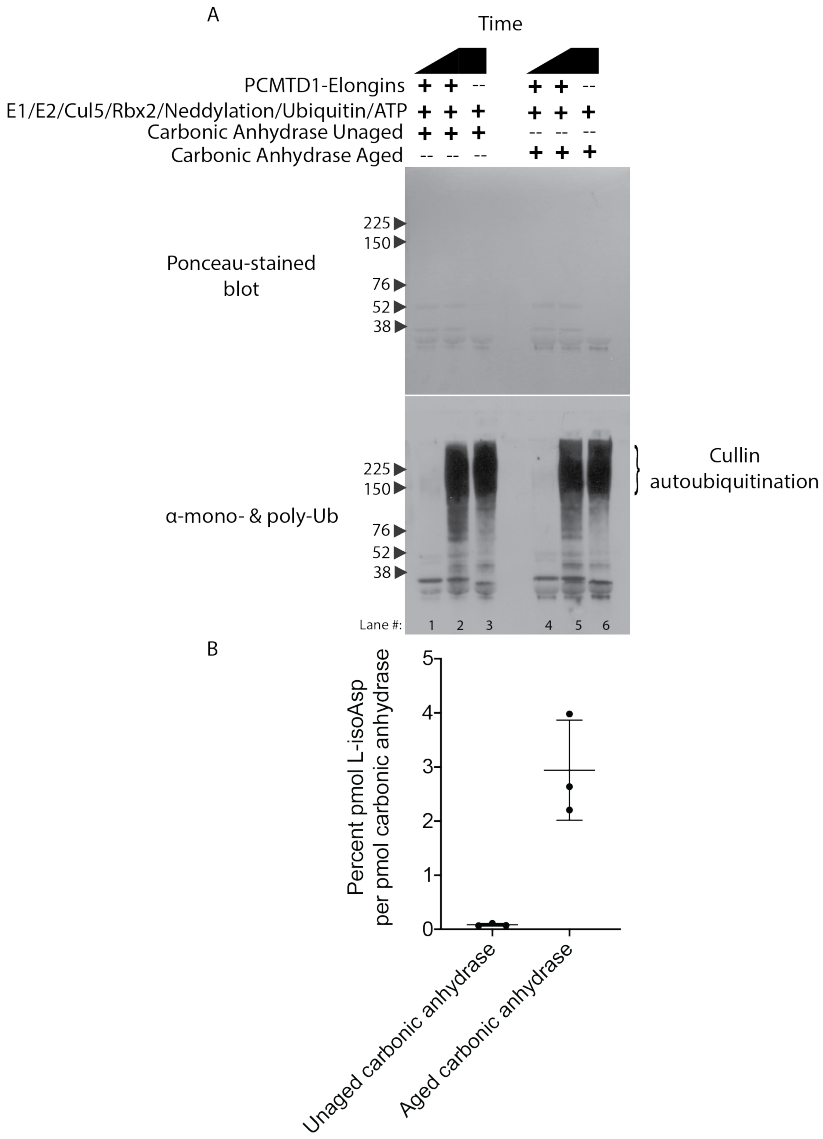


FIG. 7. ***In vitro*** **ubiquitination assay with the carbonic anhydrase substrate reveals PCMTD1-dependent ubiquitination**. A, PCMTD1-Elongins were incubated with the cullin-RING ligase components (E1/E2/Cul5/Rbx2/Neddlyation/Ubiquitin/ATP) and unaged or aged carbonic anhydrase. Reactions in lanes 1 and 4 were stopped immediately after combining all components. Lanes 2, 3, 5, and 6 were stopped after 30 minutes. Loading and transfer was evaluated with Ponceau staining of the blot (top panel), ubiquitination activity was detected by western blot using an anti-ubiquitin antibody (lower panel). B, The stoichiometry of pmol L-isoAsp per pmol carbonic anhydrase in unaged and aged samples was quantified using the methanol vapor diffusion assay.

CHAPTER 6

Perspectives, future research, and concluding remarks

The age-related accumulation of L-isoaspartate within proteins is challenging to characterize due to its isobaric mass and typically sub-stoichiometric levels. Due to its ubiquitous nature, there are a number of approaches that can be taken and a variety of different systems in which to study this modification. This dissertation presents significant advancements in our understanding of L-isoaspartate in physiologically relevant peptides and their aggregation, as well as the characterization of a novel L-isoaspartate maintenance pathway. While the results presented here represent major steps forward in comprehending the complex system of L-isoaspartate damage and repair, they also reveal new questions that need to be addressed by further research. In this chapter I will discuss possible methods for investigating these L-isoaspartate systems further.

What are the exact nature PCMT1's and L-isoaspartate's roles in cataract formation?

In Chapters 2 and 3 of this dissertation, I present work performed in the rodent and human eye lens, respectively. Experiments performed in these chapters definitively show that L-isoaspartate accumulates to high levels in the aged lens, even while endogenous PCMT1 retains the ability to methylate L-isoaspartate *in vitro*. To understand the repercussions of this, we endeavored to localize the damage, and found that while L-isoaspartate does localize to molecular weights of intact crystallins, the strongest signal is seen below these molecular weights (Figure 3.5). We identified the peptides in this region and combined this with previous research on isoaspartyl "hot spots" within crystallins to perform *in vitro* experiments characterizing the properties of these altered peptides. In accordance with the body of literature that exists pertaining to L-isoaspartate, the effects of the modification were variable within the peptides analyzed. In some cases the introduction of the modification abated aggregative properties, while in others aggregation was increased and chaperone-inhibitory properties were increased.

From these results it is clear that this age-related modification plays a complex role in protein aggregation within the aging lens and brings up several key questions. First, is the

formation of the L-isoaspartate modification linked to the formation of these proteolytic fragments? This could be answered by an extensive identification of peptides and L-isoaspartate sites within eye lenses of different ages to see if the fragments initially contain the modification. Next, is there a predictable trend that dictates when the modification will enhance or decrease aggregation? This might be aiding by computational means in which the L-isoaspartate is introduced into sequences and then these are threaded through peptide backbones, which are known to aggregate, similar to current steric zipper prediction softwares (1). However, this would be best aided by exhaustive *in vitro* experimentation testing the effects of the modification within physiologically relevant crystallin peptides. The consequences modification within aggregates should also be addressed, for example, does the modification affect proteolytic attempts to clear aggregates? It would also be informative to determine the rates of L-isoaspartate formation within the significantly damaged peptides, to compare to the loss of transparency in the lens. Lastly, but no less important, is to fully explain why the PCMT1 repair enzyme cannot maintain L-isoaspartate levels efficiently in the eye lens tissue. Understanding this may reveal important steps in the formation of cataract.

Addressing the nature and relevance of L-isoAsp23 amyloid-β (Aβ) in vivo

In vitro we have established that L-isoaspartate at site 23 in Aβ greatly enhances amyloid aggregation. The structural studies performed as a part of this dissertation suggest that this may, at least in part, be due to the formation of a more stable inter-protofilament interface facilitated by the modified residue. This certainly has revealed a novel inter-protofilament interface within the L-isoAsp23 containing 20-34 Aβ segment, which also yields insight into the kinds of amyloid steric zippers that an L-isoaspartate residue can facilitate. This could perhaps even help explain the enhanced amyloid formation of the crystallin $\alpha A^{66\text{-}80,\ isoAsp76}$ peptide discussed in Chapter 3.

However, it is also possible that the observations of faster rates of amyloid formation within the L-isoAsp containing peptide are not due to the formation a more favorable interface.

Instead, it may be the flexibility of the L-isoaspartate's methylene group that allows the peptide to sample a large number of peptide conformations in a shorter amount of time, forming the single protofilament core faster. In effect this would create a nucleus for aggregation faster. If this were the case then the amyloid structure in the initial stages of aggregation would likely lack the second L-isoAsp containing interface. Isolation of aggregates at different time points during aggregation and subsequent fiber diffraction studies might reveal changes in the patterns corresponding to the formation of the second interface. Discerning the root cause in accelerated amyloid formation by L-isoaspartate has important implications for Aβ related diseases in which L-isoAsp23 has been found. A determination of the levels of this modified site from disease brains would also enhance our knowledge of this modification's role in disease.

An undeniable query arises from the discussion in the paragraph above and Chapter 4: Is the novel L-isoAsp23 interface we observed present *in vivo*? Polymorphism-specific antibodies have been generated against Aβ previously (2), antibodies against L-isoAsp23 aggregate containing the novel interface could be raised. These antibodies could then be used to perform immunohistochemical studies of control and Alzheimer's disease brain tissues to check for the presence of this structure. Alternatively, researchers have recently executed the isolation and structure solution of some tau fibril polymorphs from Alzheimer's and Pick's disease brain (3, 4). If Aβ fibrils could be isolated from disease tissue, these fibrils could be probed for the presence of L-isoAsp23 using antibodies that have already been produced against L-isoAsp23 Aβ (5). If the immunolabeled L-isoAsp23 fibers exhibited a distinct morphology, a low-resolution single particle structure might reveal the presence of the second interface.

These studies could admittedly be complicated by cross-seeded aggregates of wild-type or otherwise modified Aβ and the L-isoAsp23 form. Thus, it will also be necessary to undertake the comprehensive characterization of hetero-structures of L-isoAsp23 and wild-type forms. This

could be attempted by co-crystallization of wild-type and isomerized segments or single particle studies of cross-seeded fibers.

What is the substrate specificity of the PCMTD1 ubiquitin ligase and the nature of its ubiquitination activity?

The work presented in Chapter 5 of this dissertation takes steps towards validating the role of PCMTD1 as a novel L-isoaspartate recognizing ubiquitin ligase by showing association with cullin-RING ligase components and lowered levels of PCMT1-methylated L-isoaspartate in the presence of these components. This is an important advancement in our understanding of the already intricate network of protein ubiquitination and degradation. The L-isoaspartate may act as a marker for protein age that is recognized for degradation by the PCMTDs. However, these are just the beginning stages of characterization for these proteins.

If the L-isoaspartate modification does act as a universal “off-switch” which is recognized by the PCMTDs for degradation, then the PCMTDs will have a relatively broad substrate specificity, as does PCMT1. To facilitate further characterization of these E3 ligases, these substrates should be identified. This could be achieved by mammalian cell expression of a tagged-PCMTD construct and pull down experiments. Once specific substrates have been identified, then *in vitro* ubiquitination assays can be performed to determine if the PCMTDs do efficiently catalyze polyubiquitination of the substrate. *In vivo* time course studies could then be done in wild-type and PCMTD knockout cell lines to monitor the degradation of the substrates.

Two significant inquiries, which arise from these studies, are whether or not there is any cross-talk between PCMT1 and the PCMTDs, and how the PCMTDs activity compares to PCMT1. Do they recognize similar substrates? Do the PCMTDs also recognize both L-isoaspartate and D-aspartate, or could they also recognize D-isoaspartate? Why does the PCMTD protein bind AdoMet? These questions can be in part answered by further *in vitro* assays investigating activity against different substrates. Structural studies of the PCMTD

proteins in complex with one or more of the cullin-RING proteins or AdoMet and an isoaspartyl substrate would also provide important insight into the activity of the enzyme.

References:

1. Goldschmidt, L., Teng, P.K., Reik, R., Eisenberg, D. Identifying the amylome, proteins capable of forming amyloid-like fibrils. *Proc Natl Acad Sci USA*. **107,** 3487-3492 (2010).
2. Hatami, A., Albay, R. Monjazeb, S., Milton, S., Glabe, C. Monoclonal antibodies against AB42 fibrils distinguish multiple aggregation state polymorphisms in vitro and in Alzheimer disease brain. *J. Biol. Chem*. **289**, 32131-32142 (2014).
3. Fitzpatrick, A.W.P, Falcon, B., He, S., Murzin, A.G., Murshudov, G., Garringer, H.J., Crowther, R.A., Ghetti, B., Goedert, M., Scheres, S.H.W. Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature.* **547,** 185-190 (2017).
4. Falcon, B., Zhang, W., Murzin, A.G., Murshudov, G., Garringer, H.J., Vidal R., Crowther, R.A., Ghetti, B., Sheres, S.H.W., Goedert, M. Structures of filaments from Pick's disease reveal a novel tau protein fold. *Nature*. **561**, 137-140 (2018).
5. Shimizu, T., Fukuda, H., Murayama, S., Izumiyama, N., Shirasawa, T. Isoaspartate formation at position 23 of amyloid beta peptide enhanced fibril formation and deposited onto senile plaques and vascular amyloids in Alzheimer's disease. *J Neurosci Res.* **70,** 451-461 (2002).

APPENDIX CHAPTER 1

A glutamate/aspartate switch controls product specificity in a protein arginine methyltransferase

The work described in this chapter has been reproduced from:

Erik W. Debler, Kanishk Jain, Rebeccah A. Warmack, You Feng, Steven G. Clarke, Gunter Blobel, and Pete Stravropoulos. “A glutamate/aspartate switch controls product specificity in a protein arginine methyltransferase.” *Proc. Natl. Acad. Sci. U.S.A*. 2016 Feb 113(8): 2068-73.

A glutamate/aspartate switch controls product specificity in a protein arginine methyltransferase

Erik W. Debler[a,b,1], Kanishk Jain[c], Rebeccah A. Warmack[c], You Feng[c], Steven G. Clarke[c], Günter Blobel[a,b,1], and Pete Stavropoulos[a,b,d,1]

[a]Laboratory of Cell Biology, The Rockefeller University, New York, NY 10065; [b]Howard Hughes Medical Institute, The Rockefeller University, New York, NY 10065; [c]Department of Chemistry and Biochemistry and The Molecular Biology Institute, University of California, Los Angeles, CA 90095; and [d]Laboratory of Lymphocyte Biology, The Rockefeller University, New York, NY 10065

Contributed by Günter Blobel, January 10, 2016 (sent for review December 18, 2015; reviewed by Gino Cingolani and Laurie K. Read)

Trypanosoma brucei **PRMT7 (*Tb*PRMT7) is a protein arginine methyltransferase (PRMT) that strictly monomethylates various substrates, thus classifying it as a type III PRMT. However, the molecular basis of its unique product specificity has remained elusive. Here, we present the structure of *Tb*PRMT7 in complex with its cofactor product *S*-adenosyl-L-homocysteine (AdoHcy) at 2.8 Å resolution and identify a glutamate residue critical for its monomethylation behavior. *Tb*PRMT7 comprises the conserved methyltransferase and β-barrel domains, an N-terminal extension, and a dimerization arm. The active site at the interface of the N-terminal extension, methyltransferase, and β-barrel domains is stabilized by the dimerization arm of the neighboring protomer, providing a structural basis for dimerization as a prerequisite for catalytic activity. Mutagenesis of active-site residues highlights the importance of Glu181, the second of the two invariant glutamate residues of the double E loop that coordinate the target arginine in substrate peptides/proteins and that increase its nucleophilicity. Strikingly, mutation of Glu181 to aspartate converts *Tb*PRMT7 into a type I PRMT, producing asymmetric dimethylarginine (ADMA). Isothermal titration calorimetry (ITC) using a histone H4 peptide showed that the Glu181Asp mutant has markedly increased affinity for monomethylated peptide with respect to the WT, suggesting that the enlarged active site can favorably accommodate monomethylated peptide and provide sufficient space for ADMA formation. In conclusion, these findings yield valuable insights into the product specificity and the catalytic mechanism of protein arginine methyltransferases and have important implications for the rational (re)design of PRMTs.**

crystal structure | enzyme catalysis | PRMT | histone methylation | epigenetics

Posttranslational modifications of proteins can affect their structure, catalytic activity, and molecular interactions (1). Methylation of the guanidino group of arginine residues represents a prominent subset of these reactions (2). Histone arginine methylation is associated with gene silencing and activation (3); the modification of arginine residues in a variety of nonhistone proteins, including splicing and transcription factors, can regulate their activity (4, 5).

Most of the enzymes that catalyze arginine methylation are designated protein arginine methyltransferases (PRMTs) and require the cofactor *S*-adenosyl-L-methionine (AdoMet) as the methyl donor (6). Four types of arginine methylation products have been described: ω-N^G-monomethylarginine (MMA), asymmetric ω-N^G-N^G-dimethylarginine (ADMA), symmetric ω-N^G-$N^{G'}$ dimethylarginine (SDMA), and δ-N^G-monomethylarginine (6, 7). Accordingly, PRMTs can be categorized into four groups: Type I PRMTs catalyze ADMA formation, type II PRMTs catalyze SDMA formation, type III PRMTs catalyze MMA formation, and type IV PRMTs catalyze δ-N^G-monomethylarginine formation. Type I, II, and III PRMTs are widely distributed in nature whereas type IV PRMTs seem to be limited to yeasts and plants (8). Interestingly, whereas type I and II enzymes catalyze MMA production in addition to their dimethyl products, type III enzymes are the only PRMTs that produce MMA alone. To date, only one methyltransferase, PRMT7, has been reported to have type III activity in mice, humans, and trypanosomes (9–13). The function of PRMT7 has been linked to cancer metastasis (4, 14, 15), DNA damage (16), pluripotency (17), and parasite infection (18).

All PRMTs that have been structurally characterized to date share a conserved catalytic core of about 300–350 residues constituting an AdoMet-binding methyltransferase and a β-barrel domain. Another common feature of PRMTs is the dimerization of the catalytic core that is realized in most cases by noncovalent association of two protomers. Covalent linkage of two PRMT modules has also been observed (19–22). Although representative structures of type I, II, and III enzymes have been determined (13, 19, 20, 23–25), our understanding of product specificity in these enzymes remains fragmentary. To unravel the molecular basis of the strict MMA activity of PRMT7 enzymes, we determined the X-ray crystal structure of *Trypanosoma brucei* PRMT7 (*Tb*PRMT7), which was recently reported in two different crystal forms (13). Although this former study established its dimeric state in solution (13), we generated a dimerization-deficient mutant where active-site residues remain intact and demonstrated that dimerization is necessary for catalysis. Importantly, we performed extensive mutational analysis and identified and proved for the first time, to our knowledge, that Glu181 is a key residue for monomethylation by *Tb*PRMT7. We carried out isothermal titration calorimetry (ITC) to characterize peptide binding to WT and mutant proteins and assayed product formation. Collectively, these studies provide new insights into the catalytic mechanism and product specificity of this class of enzymes.

Significance

Posttranslational modifications in proteins profoundly modulate their function, and enzymes that generate these modifications therefore have key regulatory roles in a wide array of biological processes. Protein arginine methyltransferases (PRMTs) attach methyl group(s) to arginines and differ in their product specificity, as they form either monomethyl arginine (MMA), asymmetric dimethylarginine (ADMA), or symmetric dimethylarginine (SDMA), each of which relays specific biological signals. Although the members of the PRMT family are structurally highly homologous, the precise molecular basis of their product specificity has not been determined. Based on our structure of *Tb*PRMT7, which explicitly forms MMA, we identified a glutamate residue as a key determinant of its product specificity, and we were able to engineer a *Tb*PRMT7 mutant capable of ADMA formation.

Author contributions: E.W.D., S.G.C., G.B., and P.S. designed research; E.W.D., K.J., R.A.W., Y.F., and P.S. performed research; E.W.D., K.J., R.A.W., Y.F., S.G.C., G.B., and P.S. analyzed data; and E.W.D., K.J., R.A.W., S.G.C., and P.S. wrote the paper.

Reviewers: G.C., Thomas Jefferson University; and L.K.R., University at Buffalo.

The authors declare no conflict of interest.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 5EKU).

[1]To whom correspondence may be addressed. Email: edebler@rockefeller.edu, blobel@rockefeller.edu, or stavrop@rockefeller.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1525783113/-/DCSupplemental.

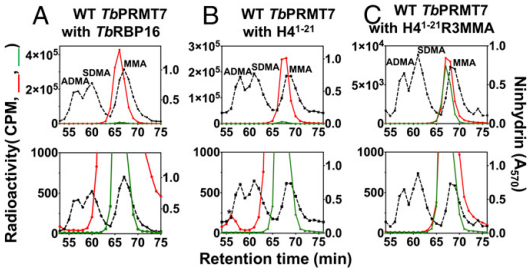


Fig. 1. *Tb*PRMT7 displays Type III PRMT activity. *Tb*PRMT7 was incubated with [*methyl*-^{3}H]-AdoMet and (*A*) *Tb*RBP16; (*B*) an *N*-acetyl peptide corresponding to residues 1–21 of human histone H4 ($H4^{1-21}$); or (*C*) the $H4^{1-21}$ peptide with MMA substituted for arginine at position 3 ($H4^{1-21}$R3MMA). Protein or peptides from the incubation were hydrolyzed and mixed with amino acid standards of ADMA, SDMA, and MMA and then analyzed via high-resolution cation-exchange chromatography. The red lines indicate the radioactivity of ^{3}H-methyl groups for complete reactions. The green lines indicate the radioactivity from a single control reaction where the enzyme alone was incubated with [*methyl*-^{3}H]-AdoMet as an automethylation control and is shown in each panel. The elution of the amino acid standards was determined by ninhydrin reactivity and is shown as a black dashed line. In each case, *Lower* represents an enlargement to show lower levels of methylation. Isotopically labeled ADMA, SDMA, and MMA elute about a minute earlier than the nonisotopically labeled standards due to the effect of tritium on the pK_as of the methylarginine species (11, 32). The asterisked small radioactive peak in *B*, *Lower* migrates about 1 min earlier than expected for [^{3}H]ADMA—the identity of this material is unknown.

Results

***Tb*PRMT7 Is a Type III PRMT That Forms only Monomethylarginine.** *Tb*PRMT7 exclusively produces MMA with protein substrates including histones and trypanosomal RNA binding protein 16 (*Tb*RBP16) (12). *Tb*PRMT7 also effectively methylates a peptide corresponding to the amino acid sequence of the 21 N-terminal residues of human histone H4; however, when the arginine residue at position 3 was replaced by an MMA residue, little or no methylation was observed (13), suggesting that Arg-3 was the prominent site of methylation in this peptide and that the enzyme was unable to catalyze the addition of a second methyl group on this residue to form SDMA or ADMA. However, the sensitivity of the latter assay was low.

To rule out any possible dimethylation, we used *Tb*RBP16 and human histone H4 peptides (residues 1–21) as substrates in an amino acid analysis method that could detect specific dimethylated substrates on a subfemtomole level. We were able to demonstrate the complete inability of *Tb*PRMT7 to produce dimethylarginine species with *Tb*RBP16 under conditions where 0.01% of such methylated species would be detected (Fig. 1*A*); sensitivity is based on the ratio of the background radioactivity to the MMA peak radioactivity. We next demonstrated that a histone $H4^{1-21}$ peptide acetylated at its N terminus is also an excellent substrate for *Tb*PRMT7 and yields only MMA where 0.02% of a dimethylated product would be detected (Fig. 1*B*). Finally, we were able to directly test for dimethylation using the corresponding peptide that had MMA substituted for arginine at position 3 ($H4^{1-21}$R3MMA); the presence of only monomethylated H4 peptide provides a direct substrate for any dimethylation reaction and limits any competition for methylation between unmethylated and monomethylated species. Again, no production of dimethylated arginine species was detected under conditions where 0.5% would be readily detected (Fig. 1*C*). Only a small increase in MMA over the background of automethylation was observed (perhaps due to methylation at arginine-17 and/or arginine-19), indicating that the major site of methylation on this peptide is at arginine-3 (Fig. 1*C*).

Overall Structure of *Tb*PRMT7. To obtain a detailed picture of the active site at atomic resolution, we set out to determine the X-ray crystal structure of *Tb*PRMT7. Limited proteolysis on the full-length protein identified a stable, N-terminally truncated fragment spanning residues 32–390, referred to as *Tb*PRMT7 in the remainder of the text for convenience. Notably, its catalytic activity is comparable with that of the full-length protein. To obtain phase information for the de novo structure determination, the protein was derivatized with seleno-methionine. The 2.8-Å crystal structure was solved using the single anomalous dispersion (SAD) phasing technique and refined to an R_{free}/R_{work} = 26.1/22.1%. For details of the data collection and refinement statistics, see Table S1.

In an independent study, Wang et al. identified a similar core fragment of *Tb*PRMT7 (residues 36–378) and determined its crystal structure in two different crystal forms (13). Despite the alternative crystal packing, *Tb*PRMT7 adopts the same compact, roughly rhombus-shaped homodimer, in which two protomers are arranged in an antiparallel fashion with a twofold symmetry axis perpendicular to the rhombus (Fig. 2*B*) (13). Because the dimeric *Tb*PRMT7–*S*-adenosyl-L-homocysteine (AdoHcy) complexes are essentially identical in the different crystal forms with an rmsd of 0.5 Å for 670 Cα atoms, confirming the validity of the

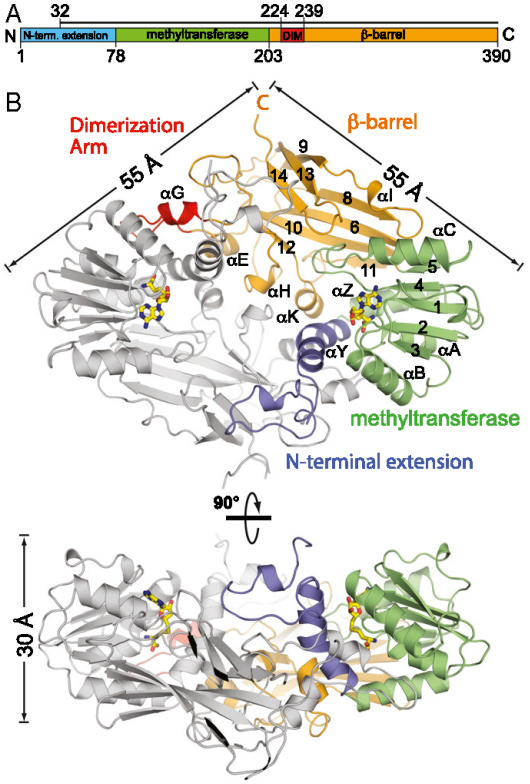


Fig. 2. Overview of *Tb*PRMT7 in complex with AdoHcy. (*A*) Domain organization of *Tb*PRMT7. Domain boundaries are indicated by residue numbers. The bar above the domain structure denotes the crystallized fragment. DIM, dimerization arm. (*B*) Ribbon representation of the *Tb*PRMT7 dimer, using the same color code for the domains as in *A*. The second protomer of the *Tb*PRMT7 dimer is shown in gray. AdoHcy is displayed in yellow stick representation. (*Lower*) A 90° rotated view is shown. Labeling of α-helices and β-strands follows the convention of earlier structures (23, 24).

Table 1. Type III enzymatic activity of WT and mutant *Tb*PRMT7 enzymes

Enzyme	Percentage type III activity
WT	100
Automethylation	1.3
Dimerization mutants	
224–235 mutated to 12 glycines	0.16
1–74 deletion	0.0055
Double E loop mutants	
G180Y	119
I173L F174L	109
I173V	102
G180N	58
E172D	21
E181D	14.2
I173A	8.6
I173G	1.9
E181Q	1.1
E172D E181D	0.32
E172Q	0.17
I173P F174M	0.04
G175D M177E	0.02
THW motif mutants	
Q329H	99.6
Q329F	90.1
Q329A	14.1
W330A	46
Helix αY mutants	
F71A	96.1
M75F	94.5
M75A	5.2

dimeric structure, only key features of the overall structure are recapitulated below.

The structured core of a *Tb*PRMT7 protomer resembles that of other PRMTs and consists of four modules: an N-terminal extension (residues 32–77), the AdoHcy-containing methyltransferase domain (residues 78–202), and the β-barrel domain (residues 203–390) that contains a protruding dimerization arm (residues 224–238) (Fig. 2*B*) (13, 23, 24). The two active sites that would bind the methyl donor AdoMet and the target arginine of the substrate are located ~36 Å apart on the same face of the homodimer. The total buried surface area between the two protomers amounts to 7,560 $Å^2$ as calculated by the PISA server (26) and is primarily formed by the N-terminal extension and the dimerization arm and to a smaller degree by helices αE, αH′, and αK at the center of the dimer (Fig. 2*B*).

Dimerization Is Necessary for *Tb*PRMT7 Catalytic Activity. Using multi-angle light scattering (MALS), we confirmed that *Tb*PRMT7 exists as a dimer in solution (Fig. S1), consistent with the large interface observed in the crystal structure and consistent with previous results from small-angle X-ray scattering (13). To assess the importance of dimerization for catalysis, we created a mutant that is deficient in dimerization while keeping the active site intact. To this end, we replaced residues 224–235 of the dimerization arm with glycines. This dimerization arm mutant was exclusively monomeric in solution, indicating that the mutant is still properly folded whereas its catalytic activity was greatly reduced (Table 1). These data strongly suggest a requirement of dimerization for efficient catalysis. We confirmed this correlation with a second mutant, in which the N-terminal extension (residues 1–74 including helix αY) was truncated. Again, dimerization is abolished with a concomitant loss in catalytic activity (Table 1). However, because helix αY of the N-terminal extension directly contributes several residues to the active site (Fig. 3), the effect of helix αY removal cannot be solely ascribed to the dimerization deficiency.

Active Site of *Tb*PRMT7. The active site of *Tb*PRMT7 is formed by the methyltransferase domain, helix αY in the N-terminal extension, and a short motif of the β-barrel (Fig. 3). The cofactor product AdoHcy is well-defined in the electron density map, adopts its canonical conformation observed in AdoMet-dependent methyltransferases, and engages in highly conserved interactions with the methyltransferase domain (24). AdoHcy sits in the lower part of the binding pocket with the methionine moiety at the bottom whereas the upper part of the cavity is poised for accommodating the incoming arginine of the substrate. The strongly negative electrostatic surface potential of the binding pocket is complementary to a substrate harboring a positively charged arginine residue (Fig. S2). Because the methyl group of AdoMet is transferred from its sulfur atom to a terminal nitrogen atom of the target arginine during catalysis, the residues adjacent to the sulfur atom of AdoHcy are expected to be most relevant for governing product specificity in *Tb*PRMT7. The two conserved eponymous glutamate residues, Glu172 and Glu181, as well as Ile173 of the double E loop in the methyltransferase domain, form one side of the binding pocket (27), opposed by Phe71 and His72 of helix αY in the N-terminal extension on the other side. The bottom of the pocket is formed by Met75 of helix αY whereas Gln329 and Trp330 of helix αK in the β-barrel domain are wedged between the methyltransferase domain and the N-terminal extension. The latter two residues are part of the THW motif termed after the threonine, histidine, and tryptophan residues predominantly found in type I PRMTs. In PRMT7 enzymes, however, methionine and glutamine substitute for the threonine and histidine residues. Comparison with the *Tb*PRMT7–AdoHcy complex of another crystal form (PDB ID code 4M37) reveals that almost all active-site residues adopt identical positions and rotamers, with slight deviations observed for Met75, Glu172, Glu181, and Gln329.

Active-Site Mutations Decrease Type III Methylation Activity. To investigate the impact of individual residues on the methylation activity, mutations were made at key residues in the active site: e.g., the double E loop and THW motif (27). Using *Tb*RBP16 as a substrate, MMA production was assayed by amino acid analysis (Table 1). Several of the mutants displayed similar levels of MMA as WT *Tb*PRMT7, including mutants for three residues in the double E loop (Ile173Val, Ile173Leu/Phe174Leu, and Gly180Tyr), a mutant of the THW motif (Gln329His), and two mutants in helix αY (Phe71Ala and Met75Phe). Although these mutations did not markedly affect activity, the remaining mutants displayed severe losses in activity, highlighting the significance of these residues in catalysis. Notably, no ADMA or

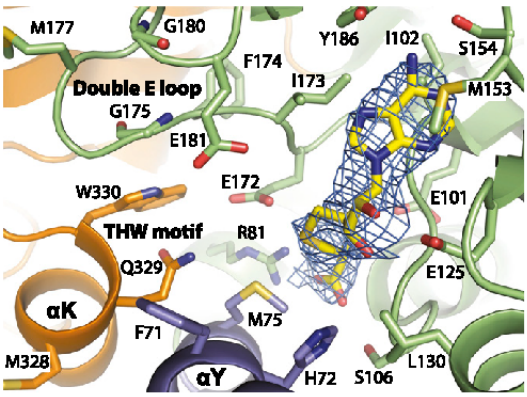


Fig. 3. Active site of *Tb*PRMT7. The electron density of AdoHcy and key residues in the active site are shown.

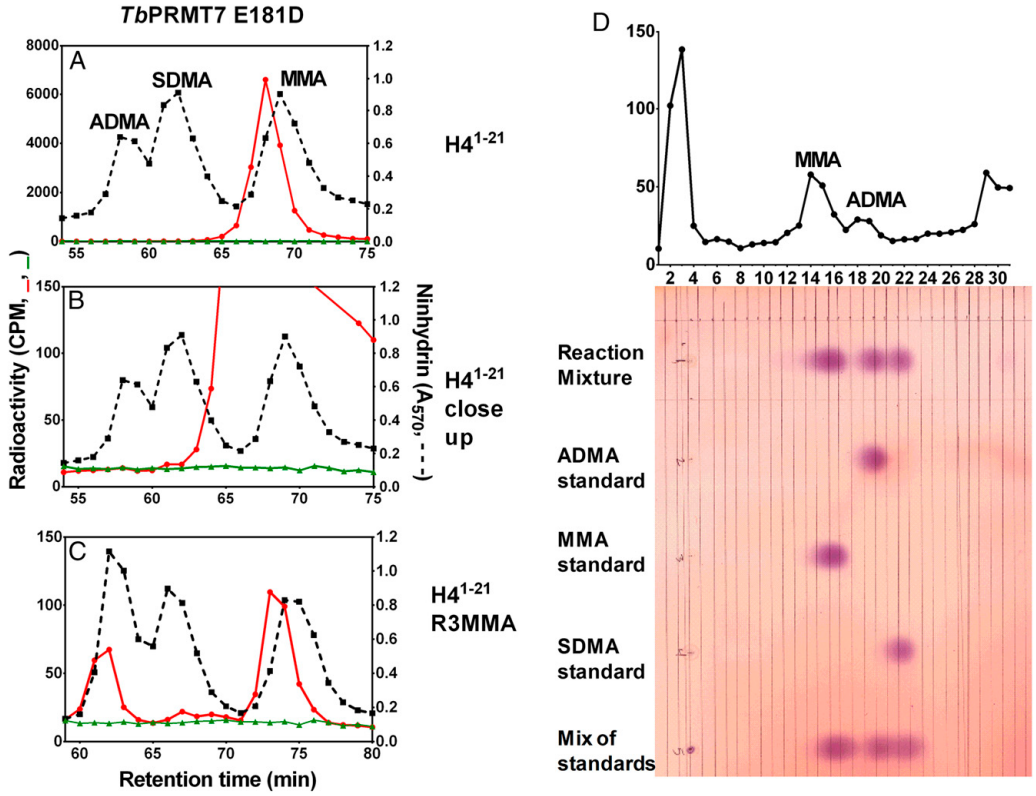


Fig. 4. An active-site mutation in the double E loop of *Tb*PRMT7 produces ADMA. *Tb*PRMT7 Glu181Asp was incubated with [*methyl*-^{3}H]-AdoMet and (*A* and *B*) the H4$^{1-21}$ peptide or (*C* and *D*) the H4$^{1-21}$R3MMA peptide. Hydrolysates of the peptide products were mixed with amino acid standards of ADMA, SDMA, and MMA and were analyzed via high-resolution cation exchange chromatography (*A*–*C*) or TLC (*D*). Red lines in *A*–*C* indicate the radioactivity of ^{3}H-*methyl* groups for complete reactions. The green lines indicate the radioactivity from a single control reaction where the enzyme alone was incubated with [*methyl*-^{3}H]-AdoMet as an automethylation control and is shown in each panel. (*B*) An enlargement of the radioactive data in *A*. The elution of the amino acid standards was determined by a ninhydrin reaction, shown in black dashes. Isotopically labeled ADMA and MMA elute about a minute earlier than the nonisotopically labeled standards. (*D*) TLC for hydrolysates of the reaction mixture and individual and mixed standards of ADMA, MMA, and SDMA. (*Upper*) The radioactivity corresponding to the TLC slices of the reaction mixture lane. (*Lower*) The ninhydrin staining of the TLC plate.

SDMA product was formed in any of these mutants under conditions where 0.02% would have been detected.

A Mutation in the Double E Loop Converts *Tb*PRMT7 into an ADMA-Producing Enzyme. The ability of PRMTs to dimethylate arginines to form either SDMA or ADMA relies on sufficient space within the active site to accommodate not only the unmodified but also the larger monomethylated arginine residue for subsequent dimethylation (28). Notably, none of the mutants yielded any ADMA or SDMA with *Tb*RBP16 or the H4$^{1-21}$ peptide. Many of these mutants showed significantly decreased activity, making it difficult to observe any level of possible dimethylarginine products. To overcome this limitation, the activity of certain mutants was tested with an already methylated H4 peptide at Arg3 (H4$^{1-21}$R3MMA). When H4$^{1-21}$R3MMA is used as a substrate with mutant Glu181Asp, we demonstrated production of both MMA and ADMA (Fig. 4*C*). The MMA formed here presumably results from the methylation of the unmodified arginine-17 and/or arginine-19 residues. With the unmodified H4$^{1-21}$ peptide, the Glu181Asp mutant shows strictly type III activity (Fig. 4 *A* and *B*).

A similar mutational study of the first glutamate in *Tb*PRMT7's double E loop (Glu172Asp) did not reveal any production of dimethylated arginine species. The side chain carboxyl atoms of the first glutamate residue in the *Tb*PRMT7 double E loop (Glu172) have lower B-values than those of the second glutamate (Glu181), consistent with the deeper burial of Glu172 in the protein interior and its larger number of interactions with neighboring residues than the more solvent-exposed Glu181. A similar situation occurs for the first and second glutamate residues in the double E loop of other PRMTs (Table S2). To further confirm the production of ADMA, TLC was performed with acid hydrolysates of the methylation reaction and methylarginine standards. Radioactivity from TLC slices confirmed that the Glu181Asp mutant does produce both MMA and ADMA with the H4$^{1-21}$R3MMA peptide (Fig. 4*D*).

The *Tb*PRMT7 Glu181Asp Mutant Has a Higher Affinity for the Monomethylated Peptide than WT *Tb*PRMT7. To analyze peptide binding to the WT and mutant Glu181Asp enzymes, we performed isothermal titration calorimetry (ITC) with H4$^{1-21}$ and H4$^{1-21}$ R3MMA peptides (Fig. 5). Consistent with a strong electrostatic interaction, binding of the H4$^{1-21}$ peptide to *Tb*PRMT7 was highly dependent on ionic strength, with relatively strong binding (K_D = 2.7 μM) at low salt concentration (20 mM NaCl) and essentially no binding at 300 mM NaCl (K_D = 0.4 M) (Table S3). The WT

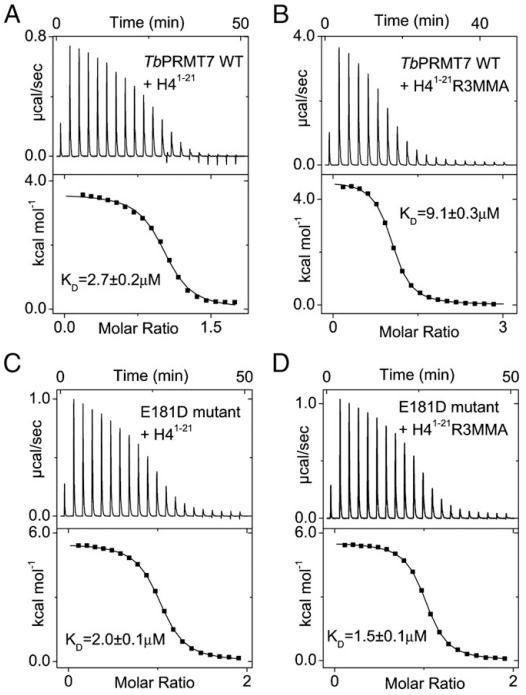


Fig. 5. Isothermal titration calorimetry of *Tb*PRMT7 WT (*A* and *B*) and Glu181Asp mutant (*C* and *D*) with $H4^{1-21}$ and $H4^{1-21}$R3MMA, respectively.

*Tb*PRMT7 enzyme binds its substrate $H4^{1-21}$ with higher affinity (K_D = 2.7 µM) than its monomethylation product $H4^{1-21}$R3MM (K_D = 9.1 µM), thermodynamically favoring product release. Although the affinity for the unmodified $H4^{1-21}$ peptide (K_D = 2.0 µM) is essentially unchanged in the Glu181Asp mutant with respect to the WT enzyme, this mutant strikingly has increased binding affinity for the methylated peptide (K_D = 1.5 vs. 9.1 µM of the WT). Thus, the Glu181Asp mutant favors binding of the bulkier $H4^{1-21}$R3MM peptide, which can be rationalized by providing a more spacious binding pocket, stabilizing the MMA substrate–enzyme interactions and enabling dimethylation.

Discussion

We determined the crystal structure of a protein arginine methyltransferase, PRMT7 from *T. brucei*, in complex with the methyl donor product AdoHcy and performed biochemical and mutational analyses to elucidate its catalytic mechanism and the structural basis of its unique product specificity. Although various suggestions have been made for the role of specific residues in the active site of PRMTs and although a constricted active site has recently been suggested as the basis for MMA activity in different crystal forms of *Tb*PRMT7 (13, 27), little is known experimentally about key residues that direct which methylarginine derivative is produced. This product specificity is crucial to the function of these enzymes because proteins that bind methylated substrates (so-called "methyl readers") differentially recognize ADMA and SDMA (3, 29), and potentially MMA. Thus, defining the exact components of catalysis and product specificity for this class of enzymes has become increasingly important. Although it has not been established that the mammalian and trypanosomal PRMT7 enzymes have the same function, the degree of sequence identity (28% over the full-length *Tb*PRMT7 protein compared with human PRMT7), its strict monomethylating behavior, and its robust enzymatic activity make the protozoan enzyme a facile system to study the biochemistry of type III PRMTs.

Notably, *Tb*PRMT7 and mammalian PRMT7 have a distinct structural organization (9, 10, 13, 20). Although the mammalian enzymes contain two PRMT cores in tandem, the trypanosomal enzyme contains only one PRMT core. However, our light scattering analysis (Fig. S1) as well as previous small-angle X-ray scattering studies (13) showed that *Tb*PRMT7 forms a homodimer in solution. The juxtaposition of the two tandem PRMT cores in mouse and roundworm PRMT7 crystal structures recapitulates the dimeric *Tb*PRMT7 architecture, sharing a similar overall quaternary structure (19, 20). Although only one PRMT module is catalytically active in mouse PRMT7 (*Mm*PRMT7) and roundworm PRMT7 (*Ce*PRMT7) (19, 20), our data (Table 1), as well as those of others, suggest that dimerization of PRMT7 cores—either by a noncovalent assembly or by juxtaposition of two units as part of one polypeptide—is a prerequisite for catalytic activity (19, 20).

Another difference between mammalian and trypanosomal PRMT7 enzymes pertains to their substrate specificity (9, 10, 13, 20). Human PRMT7, for example, has a substrate specificity for RXR motifs surrounded by basic residues, such as Arg17 and Arg19 in $H4^{1-21}$ (9, 10), whereas our experiments with WT *Tb*PRMT7 and the $H4^{1-21}$ peptide demonstrate that arginine methylation does not occur at this motif. By contrast, Arg3 is the major site of methylation in this peptide by *Tb*PRMT7 (Fig. 1), as indicated by the lack of MMA production above automethylation levels with the $H4^{1-21}$ peptide monomethylated at Arg3. Furthermore, the absence of any additional methylated species confirms the inability of *Tb*PRMT7 to catalyze dimethylation reactions.

Site-directed mutagenesis experiments involving residues of the *Tb*PRMT7 active site have given insight into key structural players of this enzyme's function. Several of the mutations are located in known PRMT motifs—including Gln329Ala in the THW motif as well as Glu172Gln and Glu181Gln in the double E loop (13, 27). The importance of the double E loop in determining PRMT activity has been widely reported (27) because the methylated nitrogen atoms of the arginine(s) in substrate proteins are coordinated through these two glutamate residues (13, 24, 25). These results demonstrate the importance of these residues on the activity of the *Tb*PRMT7 enzyme.

By analyzing the effects that certain residues have on *Tb*PRMT7 in a biochemical and structural context, we have identified Glu181 as a crucial molecular component of this enzyme's unique catalysis (Fig. 4). In so doing, we have also expanded the current understanding of how type I and type III enzymes behave and why their methylation behavior differs from one another. A comparison of the distance between the two glutamate residues of the double E loop shows that the glutamates of the type III PRMTs are generally closer together (8.4 Å on average for *Tb*PRMT7, PDB ID codes 4M36, 4M37, 4M38, and 5EKU; 6.8 Å for *Ce*PRMT7, PDB ID codes 3WST and 3XOD; and 8.7 Å for *Mm*PRMT7, PDB ID code 4C4A) than those in type I PRMTs (e.g., 12.5 Å in *Rn*PRMT1, PDB ID code 1ORI; 13.2 Å in *Mm*PRMT6, PDB ID code 4C07; and 8.5 Å in *Hs*PRMT4, PDB ID code 4IKP) (Fig. S3), consistent with the previous finding that *Tb*PRMT7 features a smaller guanidine binding pocket with respect to type I and type II PRMTs (13). The Glu181Asp mutation effectively shortens the side chain at that position by one carbon–carbon bond, increasing the distance between the flanking residues of the double E loop. An enlarged cavity would resemble the larger type I enzyme cavities more closely and allow for the recognition of an already methylated species, resulting in subsequent dimethylation. Indeed, our binding studies show a marked increase in affinity in the Glu181Asp mutant for a methylated substrate ($H4^{1-21}$R3MMA) with respect to the WT enzyme (Fig. 5). The Glu181Asp mutation still maintains the same charge at that position, allowing for similar interactions between the enzyme and the substrate's arginine, albeit at a weaker level. These findings provide a model for predicting the structural determinants of type I methylation, mediated by enzymes with longer distances between glutamates

of the double E loop, versus type III methylation, mediated by enzymes with relatively shorter distances between the glutamates of the double E loop (about 1–2 Å shorter).

Product specificity plays also an important role in histone lysine methyltransferases (HKMTs) because mono-, di-, and trimethylation of lysines relay different biological signals and thus must be strictly controlled. A comparison between the structures of the lysine monomethyltransferase SET7/9 and the trimethyltransferase DIM-5 revealed that the different product specificity can be ascribed to a switch between a tyrosine in SET7/9 for a phenylalanine residue in DIM-5 (30, 31). In detail, DIM-5 contains Phe281 in close proximity to the ε-amino group of the target lysine. In SET7/9, Tyr305 occupies the equivalent position of Phe281 in DIM-5. The addition of the hydroxyl group in Tyr305 narrows the substrate channel and sterically hinders the insertion of a dimethylated ε-amino group at the target lysine, resulting in strict monomethylation activity of SET7/9. Strikingly, the Phe281Tyr mutation in DIM-5 resulted in the alteration of product specificity from a trimethylated lysine to a mono- and dimethylated product whereas the reverse mutation of Tyr305Phe in SET7/9 abolished the strict monomethylation activity and allowed for the formation of di- and trimethylated product. This tyrosine/phenylalanine switch in HKMTs depicts the importance of the absence or presence of a single nonhydrogen atom in the active site for product specificity and recapitulates our finding of an analogous glutamate/aspartate switch that controls product specificity in PRMTs.

In conclusion, our biochemical and structural analyses have defined the components that restrict type III PRMTs to only monomethylate their target arginines and, more generally, have provided insights into the attributes that direct PRMT-mediated catalysis. This knowledge not only will be useful for the prediction of the methylation behavior of novel, thus far uncharacterized members of this family but also will prove valuable in the rational (re)design of other PRMTs.

Methods

The details of molecular cloning, expression, purification, crystallization, X-ray diffraction data collection, structure determination, isothermal titration calorimetry, multiangle light scattering, and amino acid analysis of protein and peptide substrates are described in *SI Methods*. In short, *Tb*PRMT7 was expressed in *E. coli* with an N-terminal His-tag. Recombinant proteins were purified using several chromatographic techniques. The structure was solved by SAD from seleno-L-methionine–labeled crystals.

ACKNOWLEDGMENTS. We thank **H.** Rust (The Scripps Research Institute) and **P.** Thompson (UMass Medical School) for providing the $H4^{1-21}$ and $H4^{1-21}$R3MMA peptides; **L.** Read [State University of New York (SUNY) Buffalo] for providing the expression vector for *Tb*RBP16; M. Dzialo (University of California, Los Angeles) for help with amino acid analysis; **D.** Berman (Rockefeller University) for help with protein expression and purification; **D.** King (University of California, Berkeley) for mass spectrometry analysis; M. Becker and **R.** Fischetti [Advanced Photon Source (APS)] and W. Shi [National Synchrotron Light Source (NSLS)] for support during data collection, and the Structural Biology and the High-Throughput Screening and Spectroscopy Resource Centers at Rockefeller University for support. X-ray data were measured at the GM/CA beamline 23ID-D at APS, supported by funds from the National Cancer Institute (NCI) (Grant ACB-12002) and National Institute of General Medical Sciences (Grant AGM-12006), and at beamline X29 at NSLS, supported by NIH Grants P41GM103473 and P30-EB-00998. E.W.D. was a Frey Fellow of the Damon Runyon Cancer Research Foundation (DRG-1977-08). This work was also supported by grants from NIH, including Grant GM026020 (to S.G.C.) and Ruth L. Kirschstein National Service Award GM007185 (to K.J. and R.A.W.).

1. Walsh G, Jefferis R (2006) Post-translational modifications in the context of therapeutic proteins. *Nat Biotechnol* 24(10):1241–1252.
2. Carr SM, Poppy Roworth A, Chan C, La Thangue NB (2015) Post-translational control of transcription factors: Methylation ranks highly. *FEBS J* 282(23):4450–4465.
3. Dhar SS, et al. (2012) Trans-tail regulation of MLL4-catalyzed H3K4 methylation by H4R3 symmetric dimethylation is mediated by a tandem PHD of MLL4. *Genes Dev* 26(24):2749–2762.
4. Baldwin RM, et al. (2015) Protein arginine methyltransferase 7 promotes breast cancer cell invasion through the induction of MMP9 expression. *Oncotarget* 6(5):3013–3032.
5. Biggar KK, Li SS (2015) Non-histone protein methylation as a regulator of cellular signalling and function. *Nat Rev Mol Cell Biol* 16(1):5–17.
6. Bedford MT, Clarke SG (2009) Protein arginine methylation in mammals: Who, what, and why. *Mol Cell* 33(1):1–13.
7. Herrmann F, Pably P, Eckerich C, Bedford MT, Fackelmayer FO (2009) Human protein arginine methyltransferases in vivo: Distinct properties of eight canonical members of the PRMT family. *J Cell Sci* 122(Pt 5):667–677.
8. Niewmierzycka A, Clarke S (1999) S-Adenosylmethionine-dependent methylation in Saccharomyces cerevisiae: Identification of a novel protein arginine methyltransferase. *J Biol Chem* 274(2):814–824.
9. Feng Y, Hadjikyriacou A, Clarke SG (2014) Substrate specificity of human protein arginine methyltransferase 7 (PRMT7): The importance of acidic residues in the double E loop. *J Biol Chem* 289(47):32604–32616.
10. Feng Y, et al. (2013) Mammalian protein arginine methyltransferase 7 (PRMT7) specifically targets RXR sites in lysine- and arginine-rich regions. *J Biol Chem* 288(52):37010–37025.
11. Zurita-Lopez CI, Sandberg T, Kelly R, Clarke SG (2012) Human protein arginine methyltransferase 7 (PRMT7) is a type III enzyme forming ω-NG-monomethylated arginine residues. *J Biol Chem* 287(11):7859–7870.
12. Fisk JC, et al. (2009) A type III protein arginine methyltransferase from the protozoan parasite Trypanosoma brucei. *J Biol Chem* 284(17):11590–11600.
13. Wang C, et al. (2014) Structural determinants for the strict monomethylation activity by trypanosoma brucei protein arginine methyltransferase 7. *Structure* 22(5):756–768.
14. Yao R, et al. (2014) PRMT7 induces epithelial-to-mesenchymal transition and promotes metastasis in breast cancer. *Cancer Res* 74(19):5656–5667.
15. Yang Y, Bedford MT (2013) Protein arginine methyltransferases and cancer. *Nat Rev Cancer* 13(1):37–50.
16. Karkhanis V, et al. (2012) Protein arginine methyltransferase 7 regulates cellular response to DNA damage by methylating promoter histones H2A and H4 of the polymerase δ catalytic subunit gene, POLD1. *J Biol Chem* 287(35):29801–29814.
17. Wang YC, Peterson SE, Loring JF (2014) Protein post-translational modifications and regulation of pluripotency in human stem cells. *Cell Res* 24(2):143–160.
18. Ferreira TR, et al. (2014) Altered expression of an RBP-associated arginine methyltransferase 7 in Leishmania major affects parasite infection. *Mol Microbiol* 94(5):1085–1102.
19. Hasegawa M, Toma-Fukai S, Kim JD, Fukamizu A, Shimizu T (2014) Protein arginine methyltransferase 7 has a novel homodimer-like structure formed by tandem repeats. *FEBS Lett* 588(10):1942–1948.
20. Cura V, Troffer-Charlier N, Wurtz JM, Bonnefond L, Cavarelli J (2014) Structural insight into arginine methylation by the mouse protein arginine methyltransferase 7: A zinc finger freezes the mimic of the dimeric state into a single active site. *Acta Crystallogr D Biol Crystallogr* 70(Pt 9):2401–2412.
21. Yang Y, et al. (2015) PRMT9 is a type II methyltransferase that methylates the splicing factor SAP145. *Nat Commun* 6:6428.
22. Hadjikyriacou A, Yang Y, Espejo A, Bedford MT, Clarke SG (2015) Unique features of human protein arginine methyltransferase 9 (PRMT9) and its substrate RNA splicing factor SF3B2. *J Biol Chem* 290(27):16723–16743.
23. Weiss VH, et al. (2000) The structure and oligomerization of the yeast arginine methyltransferase, Hmt1. *Nat Struct Biol* 7(12):1165–1171.
24. Zhang X, Zhou L, Cheng X (2000) Crystal structure of the conserved core of protein arginine methyltransferase PRMT3. *EMBO J* 19(14):3509–3519.
25. Sun L, et al. (2011) Structural insights into protein arginine symmetric dimethylation by PRMT5. *Proc Natl Acad Sci USA* 108(51):20538–20543.
26. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372(3):774–797.
27. Fuhrmann J, Clancy KW, Thompson PR (2015) Chemical biology of protein arginine modifications in epigenetic regulation. *Chem Rev* 115(11):5413–5461.
28. Lee HW, Kim S, Paik WK (1977) S-adenosylmethionine: Protein-arginine methyltransferase. Purification and mechanism of the enzyme. *Biochemistry* 16(1):78–85.
29. Migliori V, et al. (2012) Symmetric dimethylation of H3R2 is a newly identified histone mark that supports euchromatin maintenance. *Nat Struct Mol Biol* 19(2):136–144.
30. Xiao B, et al. (2003) Structure and catalytic mechanism of the human histone methyltransferase SET7/9. *Nature* 421(6923):652–656.
31. Zhang X, et al. (2003) Structural basis for the product specificity of histone lysine methyltransferases. *Mol Cell* 12(1):177–185.
32. Gottschling H, Freese E (1962) A tritium isotope effect on ion exchange chromatography. *Nature* 196:829–831.
33. Hayman ML, Read LK (1999) Trypanosoma brucei RBP16 is a mitochondrial Y-box family protein with guide RNA binding activity. *J Biol Chem* 274(17):12067–12074.
34. Otwinowski Z, Minor W (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol* 276:307–326.
35. Adams PD, et al. (2011) The Phenix software for automated determination of macromolecular structures. *Methods* 55(1):94–106.
36. Jones TA, Zou J-Y, Cowan SW, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* 47(Pt 2):110–119.
37. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486–501.
38. Davis IW, et al. (2007) MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35(Web Server issue):W375–W383.
39. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA (2001) Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc Natl Acad Sci USA* 98(18):10037–10041.
40. Wyatt PJ (1997) Multiangle light scattering: The basic tool for macromolecular characterization. *Instrum Sci Technol* 25:1–18.
41. Karplus PA, Diederichs K (2012) Linking crystallographic model and data quality. *Science* 336(6084):1030–1033.

APPENDIX CHAPTER 2

Protein Arginine Methyltransferase Product Specificity Is Mediated by Distinct Active-site Architectures

The work described in this chapter has been reproduced from:

Kanishk Jain, Rebeccah A. Warmack, Erik W. Debler, Andrea Hadjikyriacou, Pete Stavropoulos, and Steven G. Clarke. “Protein arginine methyltransferase product specificity is mediated by distinct active-site architectures.” J. Biol. Chem. 2016 Aug 291(35): 18299-308

Protein Arginine Methyltransferase Product Specificity Is Mediated by Distinct Active-site Architectures*

Received for publication, June 1, 2016, and in revised form, July 5, 2016 Published, JBC Papers in Press, July 7, 2016, DOI 10.1074/jbc.M116.740399

Kanishk Jain[‡], Rebeccah A. Warmack[‡], Erik W. Debler[§1], Andrea Hadjikyriacou[‡], Peter Stavropoulos[§¶], and Steven G. Clarke[‡2]

From the [‡]Department of Chemistry and Biochemistry and the Molecular Biology Institute, UCLA, Los Angeles, California 90095 and the [§]Laboratory of Cell Biology and [¶]Laboratory of Lymphocyte Biology, The Rockefeller University, New York, New York 10065

In the family of protein arginine methyltransferases (PRMTs) that predominantly generate either asymmetric or symmetric dimethylarginine (SDMA), PRMT7 is unique in producing solely monomethylarginine (MMA) products. The type of methylation on histones and other proteins dictates changes in gene expression, and numerous studies have linked altered profiles of methyl marks with disease phenotypes. Given the importance of specific inhibitor development, it is crucial to understand the mechanisms by which PRMT product specificity is conferred. We have focused our attention on active-site residues of PRMT7 from the protozoan *Trypanosoma brucei*. We have designed 26 single and double mutations in the active site, including residues in the Glu-Xaa$_8$-Glu (double E) loop and the Met-Gln-Trp sequence of the canonical Thr-His-Trp (THW) loop known to interact with the methyl-accepting substrate arginine. Analysis of the reaction products by high resolution cation exchange chromatography combined with the knowledge of PRMT crystal structures suggests a model where the size of two distinct subregions in the active site determines PRMT7 product specificity. A dual mutation of Glu-181 to Asp in the double E loop and Gln-329 to Ala in the canonical THW loop enables the enzyme to produce SDMA. Consistent with our model, the mutation of Cys-431 to His in the THW loop of human PRMT9 shifts its product specificity from SDMA toward MMA. Together with previous results, these findings provide a structural basis and a general model for product specificity in PRMTs, which will be useful for the rational design of specific PRMT inhibitors.

Methylation of proteins is a major type of post-translational modification involved in the regulation of a variety of cellular processes mediated by protein-protein interactions, including splicing, transcription, translation, and signaling (1–3). Recent studies have implicated arginine methylation in altering the metabolic landscape of the cell, linking it to cancer metastasis (4–6), DNA damage (7), pluripotency (8), and parasite infection (9, 10). Catalysis of arginine methylation on the terminal nitrogen atoms of the guanidine group is mediated by a family of enzymes designated as protein arginine methyltransferases (PRMTs).[3] Most of these enzymes harbor a conserved ~310-residue core that comprises the methyltransferase domain conserved in *S*-adenosylmethionine (AdoMet)-dependent methyltransferases and a β-barrel domain unique to the PRMT family. These enzymes can be further categorized based on which methylarginine product they catalyze as follows: type I PRMTs catalyze the production of ω-N^G-monomethylarginine (MMA) and asymmetric ω-N^G,N^G-dimethylarginine (ADMA); type II PRMTs catalyze the production of MMA and symmetric ω-N^G,N'^G-dimethylarginine (SDMA); type III PRMTs catalyze the production of only MMA; and type IV PRMTs catalyze δ-N^G-monomethylarginine production (11). Notably, most PRMTs fall under the first three types of PRMTs. Type IV enzymes have only been reported in yeast and plants, although the presence of free δ-N^G-monomethylarginine has been reported in human plasma in a recent proteomic study (12).

ADMA and SDMA methyl marks on histones are recognized by different "reader" proteins and can lead to distinct downstream outcomes. For example, whether a particular arginine residue on histone tails is asymmetrically or symmetrically dimethylated can lead to gene repression or activation (13–17). However, few studies have been conducted to determine the role of MMA marks (18). It has been proposed that MMA marks are used mainly as precursors for dimethylation by the various type I and II PRMTs (17, 19).

Given the biological significance of the type of methylated arginine derivative formed, it is important to understand how product specificity is determined in PRMTs. It has been suggested that small variations in the structure of the active site of these enzymes govern the methylation activity type (2, 3, 20–23). Although previous studies utilizing site-directed mutagenesis have given some support for this hypothesis, efforts to efficiently change the activity type of PRMTs have not yet been fruitful. Two such studies using moderately sensitive analytical techniques have been reported for PRMT1 (14) and

* This work was supported by National Institutes of Health Grant GM026020 (to S. G. C.) and Ruth L. Kirschstein National Service Award GM007185 (to K. J., R .A. W., and A. H.). The authors declare that they have no conflicts of interest with the contents of this article. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

[1] Frey Fellow of the Damon Runyon Cancer Research Foundation supported by Grant DRG-1977-08.

[2] To whom correspondence should be addressed: Dept. of Chemistry and Biochemistry and Molecular Biology Institute, UCLA, 607 Charles E. Young Dr. East, Los Angeles, CA 90095. Tel.: 310-825-8754; E-mail: clarke@mbi.ucla.edu.

[3] The abbreviations used are: PRMT, protein arginine methyltransferase; MMA, ω-N^G-monomethylarginine; ADMA, ω-N^G,N^G-asymmetric dimethylarginine; SDMA, ω-N^G,N'^G-symmetric dimethylarginine; AdoHcy, *S*-adenosyl-L-homocysteine; AdoMet, *S*-adenosyl-L-methionine; [*methyl*-^{3}H]-AdoMet, *S*-adenosyl-[*methyl*-^{3}H]-L-methionine; ITC, isothermal titration calorimetry.

TABLE 1

Product analyses of wild-type and mutant *Tb*PRMT7 enzymes with the H4(1–21) R3MMA peptide

The number of experiments is indicated in parenthesis. As shown under the "Experimental Procedures," 86 cpm correspond to 1 fmol of methyl groups.

	[^{3}H]Methyl group radioactivity in MMA (average cpm)	[^{3}H]Methyl group radioactivity in ADMA (average cpm)	[^{3}H]Methyl group radioactivity in SDMA (average cpm)
*Tb*PRMT7 enzyme			
Wild type (n = 4)	11,093	0	0
Automethylation (n = 4)	122	0	0
Double E loop mutants			
G180N (n = 1)	6,648	0	0
G180Y (n = 1)	5,365	0	0
E172Q (n = 1)	0	0	0
E181D (n = 3)	294	161	0
E181Q (n = 1)	62	0	0
I173G (n = 1)	0	0	0
I173A (n = 1)	121	0	0
I173V (n = 1)	7,231	0	0
I173P F174 M (n = 1)	0	0	0
I173L F174L (n = 1)	29,487	0	0
G175D M177E (n = 1)	0	0	0
E172D E181D (n = 1)	0	0	0
E181D I173G (n = 2)	64	0	0
THW loop mutants			
Q329A (n = 1)	420	0	0
Q329F (n = 1)	95	0	0
Q329H (n = 1)	23,218	0	0
W330A (n = 1)	437	0	0
Q329N (n = 1)	3,322	0	0
Helix αY mutants			
F71A (n = 1)	123	0	0
M75A (n = 1)	205	0	0
M75F (n = 1)	644	0	0
Double E loop and THW loop double mutants			
E181D W330A (n = 1)	52	0	0
E181D Q329A (n = 2)	89	0	500
E181D Q329N (n = 1)	0	0	0
Double E loop and helix αY double mutants			
E181D M75A (n = 1)	65	0	0
E181D F71A (n = 1)	143	0	0

PRMT5 (24), but they have not put forth a general model for the factors that guide product specificity for the three main types of PRMTs.

Using an approach where MMA, ADMA, and SDMA can be detected with sub-femtomole sensitivity, we have been able to demonstrate the transformation of PRMT7 from *Trypanosoma brucei* (*Tb*PRMT7) from an enzyme that strictly produces MMA to one also forming ADMA by replacing a glutamate residue in the double E loop (Glu-181) with an aspartate residue (Fig. 1) (25). The double E loop is a conserved feature of PRMTs that has been shown to directly interact with the methyl-accepting arginine residue (2, 11). *Tb*PRMT7 had been initially characterized for a possible role in the transcriptional control of gene expression in this organism (26). Here, we have focused on *Tb*PRMT7 because it displays robust type III activity and has been amenable to structural analysis (25–27). Within this work, we further examine the effects of key active-site residues on the enzymatic activity of *Tb*PRMT7 through mutagenesis and highly sensitive amino acid analysis techniques to demonstrate the importance of the THW loop (MQW for *Tb*PRMT7) (Fig. 1) (2). Complementary studies with PRMT9 from *Homo sapiens*, previously characterized as a type II enzyme (28, 29), corroborate our PRMT7 results. Based on this evidence, we propose a structural model for how PRMTs can limit their activities to type I, type II, or type III methylation.

Results

TbPRMT7 Active-site Double Mutation, E181D/Q329A, Converts the Enzyme to an SDMA-producing PRMT—Given the ability of the double E loop E181D mutation of *Tb*PRMT7 to alter the methylation type (25), seven *Tb*PRMT7 double mutants were generated with the E181D background to probe the effects of further increasing the size of the active site. Notably, the double mutant E172D/E181D was previously tested and found inactive (Table 1) (25). The additional substitutions in the six new double mutants included M75A, Q329A, Q329N, W330A, F71A, and I173G, each with the E181D mutation, based on their immediate vicinity to the sulfur atom of AdoHcy from which the methyl group of AdoMet is transferred to the arginine residue in protein and peptide substrates (Fig. 1). Using [*methyl*-^{3}H]AdoMet as a cofactor, we analyzed the hydrolyzed products of arginine methylation with high resolution cation exchange chromatography. Two of these double mutants showed little or no activity toward either the H4(1–21) peptide, comprised of the acetylated 21 N-terminal residues of the human histone H4 protein (data not shown), or the acetylated H4(1–21) R3MMA peptide, ω-monomethylated at the third arginine (Table 1). The H4(1–21) R3MMA peptide was used to enhance the detection of dimethylarginine derivatives by providing a substrate where a single methyla-

tion reaction at the primary site of modification could result in dimethylation of the peptide. Strikingly, one of the double mutants, E181D/Q329A, produced SDMA when incubated with the H4(1–21) R3MMA peptide (Fig. 2, *A* and *B*). The small amount of MMA produced in the reaction of the E181D/Q329A mutant with the H4(1–21) R3MMA peptide is most likely due to methylation of the secondary Arg-17 and Arg-19 sites on the histone peptide because the level of radioactivity here is higher than the level recorded for the enzyme alone (Fig. 2*B*). Importantly, although this enzyme contains the ADMA-producing mutation E181D (25), as well as a Q329A mutation in the THW loop, no ADMA formation was detected. SDMA production catalyzed by the E181D/Q329A mutant was confirmed by TLC analysis where the radioactive product co-migrated with the non-radioactive SDMA standard (Fig. 2*C*). The wild-type *Tb*PRMT7 does not produce any dimethylarginine products with either H4(1–21) or H4(1–21) R3MMA peptide (Fig. 3). The single Q329A mutant shows no evidence of dimethylarginine formation (Table 1).

TbPRMT7 E181D/Q329A Shows Higher Binding Affinity for the Monomethylated Histone H4(1–21) Peptide Than for the Unmethylated Peptide—Using isothermal titration calorimetry (ITC) with H4(1–21) and H4(1–21) R3MMA peptides, we previously demonstrated that the wild-type *Tb*PRMT7 enzyme binds its substrate H4(1–21) with higher affinity than its monomethylated product, H4(1–21) R3MMA, whereas the ADMA-producing *Tb*PRMT7 E181D mutant has markedly increased affinity for H4(1–21) R3MMA that even surpasses that for H4(1–21) (25). Similarly, we measured the affinity of the SDMA-producing *Tb*PRMT7 E181D/Q329A enzyme and

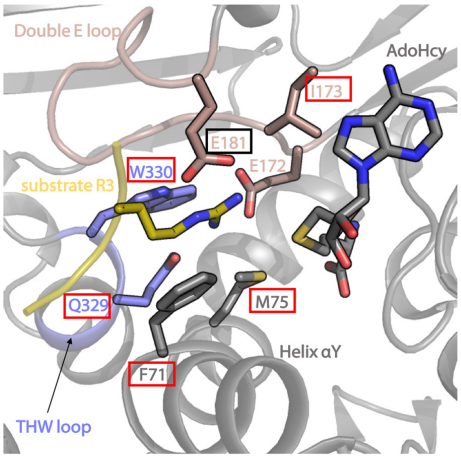


FIGURE 1. **Active site of *T. brucei* PRMT7.** Residue Glu-181, highlighted in the *black box*, is the site of mutation (E181D) shared by the six double mutants in this study with their second mutated residue highlighted in a *red box* (W330A, Q329A, Q329N, F71A, M75A, and I173G). The double mutant E172D/E181D was previously analyzed and Glu-172 is therefore not highlighted here (25). The double E loop is shown in *dark salmon*, the THW loop in *slate*, the substrate arginine residue in *yellow*, and the AdoHcy cofactor, helix αY, and adjacent residues of *Tb*PRMT7 (Protein Data Bank code 4M38) in *gray*.

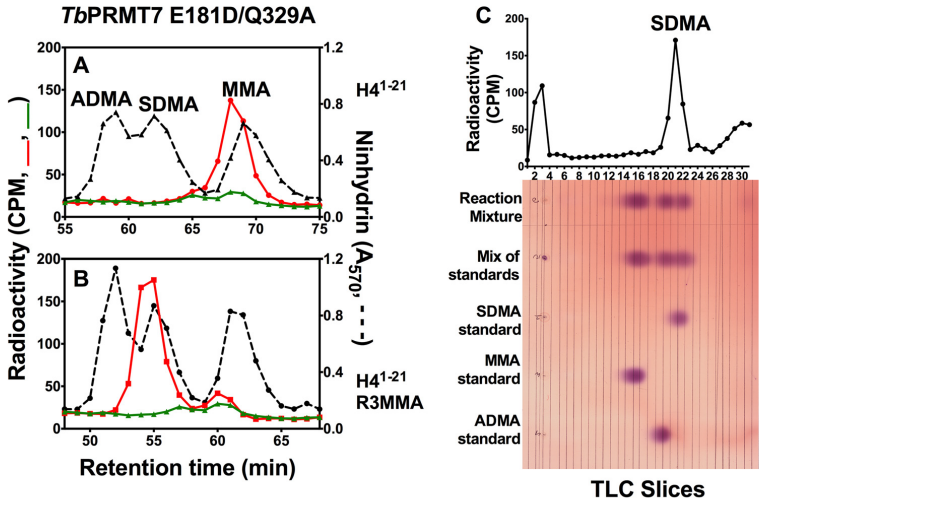


FIGURE 2. ***Tb*PRMT7 E181D/Q329A double mutant produces SDMA with the H4(1–21) R3MMA peptide.** The specificity of this mutant was determined using cation exchange chromatography and TLC as described under "Experimental Procedures." *Tb*PRMT7 E181D/Q329A (4.8 μg of protein) was incubated with the H4(1–21) or H4(1–21) R3MMA peptide (10 μM) and [*methyl*-^{3}H]AdoMet in a final volume of 60 μl. *A*, *Tb*PRMT7 E181D/Q329A double mutant with the H4(1–21) peptide. *B*, *Tb*PRMT7 E181D/Q329A with the H4(1–21) R3MMA peptide. The *red lines* in *A* and *B* represent radioactivity of the E181D/Q329A mutant with the different substrates, and the *green lines* indicate radioactivity of the methylation reaction with no substrate. As noted previously (25), radioactive methylarginine derivatives elute 1 min earlier than their non-radioactive counterparts due to the isotope effect (39, 40). As given under "Experimental Procedures," 86 cpm correspond to 1 fmol of methyl groups. For the number of biological replicates, see Table 1. *C*, representative TLC for hydrolysates of the reaction mixture and individual and mixed standards of ADMA, MMA, and SDMA. The *lower portion* shows the ninhydrin staining of the TLC plate; the *upper portion* shows the radioactivity corresponding to the TLC slices of the reaction mixture lane. Note: the ninhydrin standards on the TLC plate are the same as those shown in Fig. 4*D* of Ref. 25 where a different reaction mixture was chromatographed adjacent to the ADMA standard lane. The experiment is one of two biological replicates.

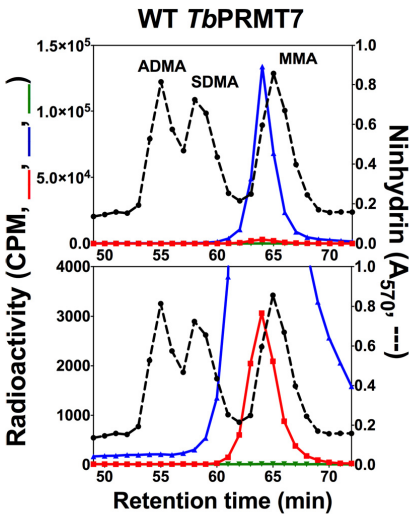


FIGURE 3. **Wild-type *Tb*PRMT7 displays no dimethylarginine production with H4(1–21) and H4(1–21) R3MMA peptides.** *In vitro* methylation and cation exchange chromatography were used as described under "Experimental Procedures" to assess wild-type *Tb*PRMT7 activity and product specificity with H4(1–21) (*blue*), H4(1–21) R3MMA peptides (*red*), or with the enzyme alone (*green*). *Dashed black lines* indicate elution profile of non-radioactive methylarginine species as measured by a ninhydrin assay (see "Experimental Procedures"). The *lower panel* represents enlargement of the radioactivity in the *upper panel* to show low levels of methylation. As given under the "Experimental Procedures," 86 cpm correspond to 1 fmol of methyl groups. For the number of biological replicates, see Table 1.

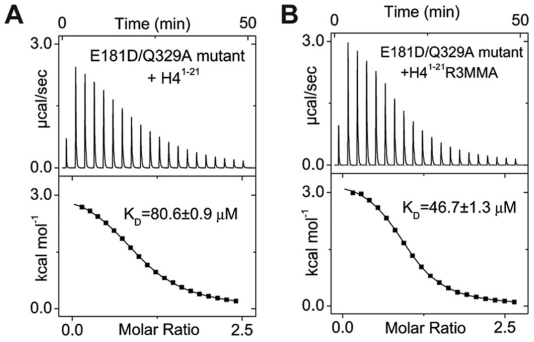


FIGURE 4. **Isothermal titration calorimetry of the *Tb*PRMT7 E181D/Q329A mutant with H4(1–21) (*A*) and H4(1–21) R3MMA (*B*), respectively.** Each titration was performed twice.

found that this mutant displays higher affinity for H4(1–21) R3MMA (K_D = 46.7 μM) *versus* its unmethylated counterpart H4(1–21) (K_D = 80.6 μM) (Fig. 4). Thus, the two mutant enzymes capable of dimethylation consistently favor binding of the bulkier H4(1–21) R3MMA peptide, which can be rationalized by providing a more spacious binding pocket, stabilizing the MMA substrate-enzyme interactions and enabling dimethylation.

Active-site Mutations Lead to Decreases in Type III PRMT7 Activity and Shifts in Recognition Site Specificity—Overall, we have reacted 26 single and double mutants of *Tb*PRMT7 with the H4(1–21) R3MMA peptide to test whether active-site mutations could display changes in the methylation type when presented with a primed monomethylarginine (Table 1). These mutations were generated based on their location in the active site of *Tb*PRMT7, including residues in the double E loop, the AdoMet-binding motif, the THW loop, and an N-terminal extension (helix αY). The majority of the active-site mutations result in decreases in enzyme activity. However, monomethylation is still observed, indicating that the modification of Arg-17 and Arg-19 on the substrate peptide is occurring, as Arg-3 is already methylated in this peptide. This finding suggests that there may be a change in recognition site specificity from glycine-arginine-rich regions to arginine residues in basic regions (26, 30). Notably, the THW (MQW) loop mutant Q329H showed significant increases in MMA production. Most remarkably, the double mutant E181D/Q329A produced both MMA and SDMA, as described above.

Mutation in the THW Loop of Human PRMT9, a Type II PRMT, Shifts Product Specificity from SDMA toward MMA—The human PRMT9 has recently been characterized as a type II PRMT, joining PRMT5 as an enzyme that catalyzes SDMA production (28, 29). This methyltransferase contains a Thr-Cys-Trp (TCW) sequence in place of the canonical Thr-His-Trp (THW) residues (28). To further investigate the role of spatial restrictions conferred by key active-site residues, the cysteine residue was mutated to a bulkier histidine residue to mimic type I and type III PRMTs. These mutant and wild-type enzymes were reacted with a GST fusion of the splicing factor SF3B2, a known substrate of PRMT9 (28). Comparison of wild-type and mutant activities reveals an impressive 8-fold increase in MMA production and almost complete elimination of SDMA production (<0.037%) (Fig. 5).

Rattus norvegicus PRMT1 M48F Mutant Enzyme Does Not Produce SDMA with Histone H4 Peptides—A previous study (24) reported a mutation in rat PRMT1 at Met-48, a residue conserved in the αY helix of many PRMTs, to Phe. This change led to the apparent production of SDMA along with ADMA and MMA, the wild-type products of a type I PRMT, as determined by *o*-phthalaldehyde-derivatized reverse-phase liquid chromatography and LC-MS analysis. However, it appeared that the degree of dimethylarginine formation was quite different when analyzed by these two methods. In our studies with *Tb*PRMT7, the homologous mutation, M75F, showed no change in PRMT7's type III activity with substrates, including RBP16 (25) and the H4(1–21) R3MMA peptide (Table 1). To validate the PRMT1 mutant activity (24), we compared the product specificity of the wild-type human PRMT1 (Fig. 6*A*) with the H4(1–21) and H4(1–21) R3MMA peptides to that of the rat PRMT1 M48F enzyme (Fig. 6*B*). We chose these peptides because H4(1–21) has been shown to be a robust PRMT1 substrate (31, 32). However, in contrast to the earlier work (24), we were unable to distinguish any difference in the product specificity of the wild-type human PRMT1 and the rat PRMT1 M48F mutant with the H4 peptide substrates using high resolution cation exchange chromatography (Fig. 6). With both enzymes, only MMA and ADMA were formed under conditions where we could detect SDMA at a level of less than 0.4% of

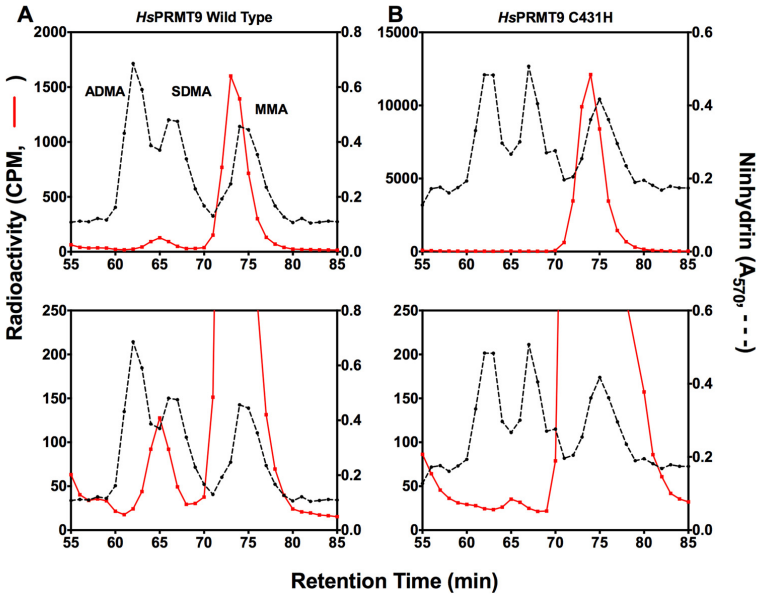


FIGURE 5. ***Hs*PRMT9 C431H mutant displays diminished SDMA and greatly increased MMA production with GST-SF3B2.** *A,* amino acid analysis of methylated arginine derivatives produced by the wild-type human GST-PRMT9 (*Hs*PRMT9) with substrate GST-SF3B2 as described under "Experimental Procedures." *B,* amino acid analysis of methylated arginine derivatives produced by the C431H mutant human GST-PRMT9 with substrate GST-SF3B2. In each case, the *lower panels* represent enlargement of the radioactivity in the *upper panels* to show low levels of methylation. As given under the "Experimental Procedures," 86 cpm correspond to 1 fmol of methyl groups. This experiment is one of two biological replicates.

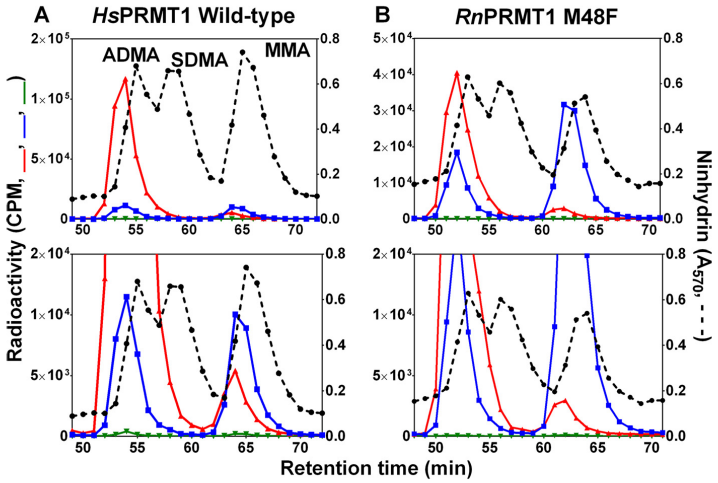


FIGURE 6. ***Rn*PRMT1 M48F mutant enzyme does not produce SDMA with histone H4 peptides.** *In vitro* methylation and cation exchange chromatography were used as described under the "Experimental Procedures" to assess PRMT1 activity and product specificity with H4(1–21) (*blue*), H4(1–21) R3MMA (*red*), and the enzyme alone (*green*). *Dashed black lines* indicate elution profile of non-radioactive methylarginine species as measured by a ninhydrin assay (see "Experimental Procedures"). *A,* amino acid analysis of methylated arginine derivatives produced by human PRMT1 (*Hs*PRMT1). *B,* amino acid analysis of methylated arginine derivatives produced by rat PRMT1 (*Rn*PRMT1) M48F mutant. In each case, the *lower panels* represent enlargement of the radioactivity in the *upper panels* to show low levels of methylation. As given under the "Experimental Procedures," 86 cpm correspond to 1 fmol of methyl groups. This experiment represents one of two biological replicates.

the radioactivity in ADMA. Significantly, in the presence of an already methylated substrate such as H4(1–21) R3MMA, the rat PRMT1 M48F was still unable to produce any SDMA (Fig. 6*B*). Additionally, there is MMA production above automethylation levels for wild-type human PRMT1 and rat PRMT1 M48F when given H4(1–21) R3MMA as a substrate. The MMA

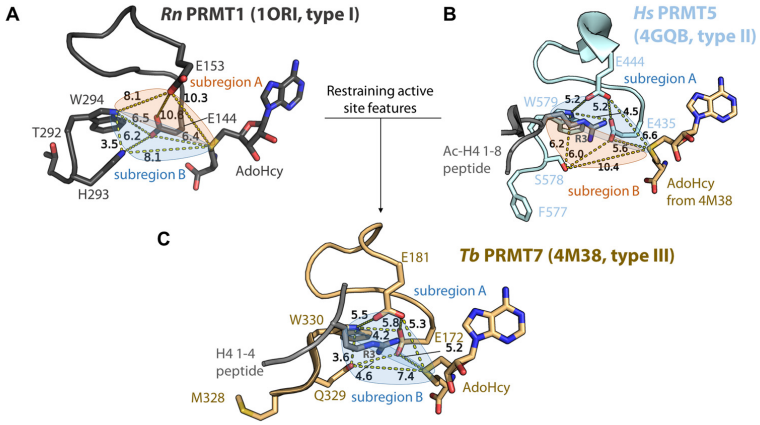


FIGURE 7. **PRMT active sites display distinct spatial architectures.** The active site (double E loop, THW loop, and AdoHcy) from *Rn*PRMT1 (1ORI, chain A; *dark gray*) (*A*), *Hs*PRMT5 (4GQB, chain A; *cyan*) (*B*), and *Tb*PRMT7 (4M38, chain A; *wheat*) (*C*) are shown. Crowded subregions of the active sites are highlighted in *light blue* and open subregions are highlighted in *orange*. *B* and *C*, substrate peptides co-crystallized with the enzyme are also shown. Distances between atoms are given in Ångströms and indicated by *yellow dashed lines*. Images were made using PyMOL (Schrödinger, LLC).

being produced with this peptide would be expected to occur at positions Arg-17 and Arg-19. These results indicate that the residue Met-48 may not be involved in mediating product specificity in mammalian PRMT1.

Discussion

Different methylarginine marks can be recognized by distinct reader proteins (17) and often behave as epigenetic switches, affecting the activation or silencing of certain genes (15, 16). Given the significance of ADMA and SDMA marks, it has become increasingly important to understand how product specificity arises to generate these residues. Having previously demonstrated the conversion of *Tb*PRMT7, a strictly MMA-producing type III enzyme, into a type I enzyme forming ADMA by mutation (25), we now present another *Tb*PRMT7 mutant that is capable of producing SDMA, exhibiting the product specificity of type II PRMTs (Fig. 2). Biochemical and mutational analyses of the enzyme's catalytic activity reveal that SDMA production occurs when it is presented with an already monomethylated substrate, demonstrating that this mutant of PRMT7, in contrast to the wild type, is able to recognize a monomethylated molecule as a substrate and carry out further methylation. In fact, the E181D/Q329A mutant enzyme binds H4(1–21) R3MMA with a higher affinity than the corresponding unmethylated peptide (Fig. 4). This observation illustrates that although the activity of the E181D/Q329A mutant is low, it still behaves, on the catalytic level, as a type II PRMT.

We also examined a mammalian PRMT1 mutant enzyme that was previously reported to produce SDMA along with its wild-type products, ADMA and MMA (24). We were unable to observe any symmetric dimethylation on histone H4 peptide substrates from this rat PRMT1 mutant enzyme (M48F) (Fig. 6). Coupled with our results from amino acid analysis of a homologous mutation in *Tb*PRMT7 (M75F; Table 1) and its mutation to alanine (M75A; Table 1) (25), our work did not confirm any role of Met-48 in affecting PRMT1 product specificity in the mammalian enzyme. It should be noted that the H4 peptide substrates used in our study differ from the GGRGGFGGRGGFGGRGGFG peptide used previously (24). Additionally, immunoblot analysis revealed that the reverse mutation in the PRMT5 enzymes of humans and *Caenorhabditis elegans*, where the corresponding wild-type residue is a phenylalanine (F327M and F379M, respectively) caused asymmetric dimethylation of human histone H4 (14). It would be interesting to examine these mutants with our more sensitive amino acid analysis techniques to determine any changes in product specificity more precisely.

Our previous mutagenesis results (25), coupled with those discussed here, highlight the major features of the PRMT active site, which may control mono- and dimethylation specificity. Conceptually, the active site of PRMTs, defined by the double E loop, the THW loop, and the AdoMet/AdoHcy cofactor, can be divided into two subregions, one of which is located between the two glutamate residues of the double E loop and above the substrate arginine (subregion A), while subregion B is adjacent to the THW loop and the region underneath the substrate arginine as displayed in Fig. 7. Our analysis reveals that the nature of these two subregions correlates well with, and therefore seems predictive of, product specificity in PRMTs. Specifically, type I PRMTs contain an open subregion A and a spatially restricted subregion B (Fig. 7*A*). The nature of these subregions in type II active sites is reversed with respect to type I PRMTs, with an open subregion B and a restricted subregion A (Fig. 7*B*). PRMT7's active site by contrast contains two restricted subregions, combining the restraining features of subregions A (type I) and B (type II) of the other two types of PRMTs (Fig. 7*C*). These spatial restrictions may be the key for PRMT7 to only monomethylate its substrates, thus classifying it as a type III PRMT enzyme.

The E181D mutation of *Tb*PRMT7 increases the space within subregion A of the active site by a single carbon-carbon

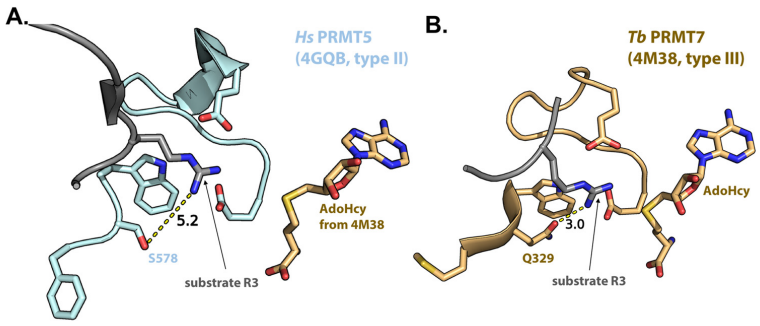


FIGURE 8. **THW loop of PRMT5 is further away from the substrate arginine than the THW loop of PRMT7.** *A*, active site of human PRMT5 (4GQB) is shown. *B*, active site of *Tb*PRMT7 (4M38) is shown. Distances between atoms are given in Ångströms and indicated by *yellow dashed lines*.

bond where the substrate arginine is stabilized. The distance between the glutamates of the double E loop, however, is not the essential factor in SDMA production because the glutamates of human PRMT5, the major SDMA producer in the cell, are actually closer together than even those in *Tb*PRMT7 (Fig. 7, *B* and *C*). The methylation type alteration can be largely attributed to the Q329A mutation, which, in combination with E181D, may result in the opening up of subregion B in the active site underneath the substrate arginine. It is important to note that the Q329A single mutant did not produce SDMA, suggesting that the THW loop may not be the sole contributor in determining type II methylation. Human PRMT5 has a serine residue in place of the corresponding glutamine residue in *Tb*PRMT7 that is located at a greater distance (5.2 Å *versus* 3.0 Å; Fig. 8) from the substrate arginine than the glutamine of *Tb*PRMT7 and is also pointed away from the active site (Fig. 7). In our *Tb*PRMT7 E181D/Q329A construct, the glutamine to alanine substitution removes an acetamide moiety in subregion B. This active-site alteration now allows for methylated arginines to bind more favorably and is better suited to accommodate a methylated nitrogen atom near the THW loop, allowing the other terminal nitrogen atom (positioned near the double E loop) to become methylated. A specific role of the THW loop in determining PRMT product specificity was first suggested in two recent reviews from Thompson and co-workers (2, 20).

In support of the importance of the THW loop in determining type II PRMT product specificity, the mutation of C431H in human PRMT9 shows a significant decrease in SDMA production relative to the wild-type enzyme (Fig. 5). Although no structure has been determined for this enzyme, the cysteine to histidine mutation introduces a bulkier moiety into the THW loop potentially contributing to further crowding in the active site, which in turn may prevent SDMA production. The concomitant marked increase in MMA production of the PRMT9 is consistent with a partially processive methylation mechanism, a characteristic of type I PRMTs (33).

Structural alignments of known type I, II, and III PRMTs show that the geometries of the active sites are highly conserved within each PRMT type (Fig. 9 and Table 2). Although our proposed model will benefit from further validation through structural studies of novel PRMTs and additional mutant enzymes, our results illustrate how small changes in the active site of PRMTs can markedly alter their catalytic specificity and thus aid in creating a spectrum of methylarginine species that may differentially mediate various biological pathways.

The emerging role of PRMTs in cancer (4, 5, 34, 35) has profoundly spurred the research into PRMT inhibitors (36). One of the major issues in this field, however, has been the promiscuity of many PRMT inhibitors derived from small molecule library screening (37). Approaches based on finding bisubstrate analogs that mimic the cofactor and the substrate arginine have the disadvantages of promiscuity and additionally, due to their highly charged nature, limited bioavailability precluding their administration as oral drugs (37). In light of such obstacles in the development of small molecule inhibitors of PRMTs involved in various diseases, it is our hope that our model will facilitate the rational design of specific and potent PRMT inhibitors by providing detailed insight into the distinct active-site architectures of the three types of PRMTs.

Experimental Procedures

Peptide Substrates—Histone H4(1–21) (Ac-SGRGKGGKGLGKGGAKRHRKV) and histone H4(1–21) R3MMA (Ac-SGR(me)GKGGKGLGKGGAKRHRKV) peptides were kind gifts from Heather Rust (The Scripps Research Institute, Jupiter, FL) and Paul Thompson (University of Massachusetts Medical School, Worcester, MA). Peptides used for ITC analysis were purchased from AnaSpec.

Protein Expression and Purification—*Tb*PRMT7 wild-type and mutant enzymes were cloned, expressed, and purified as described previously (25). GST-PRMT9 wild-type, GST-PRMT9 C431H mutant, and GST-SF3B2(401–550) fragment were expressed and purified as described previously (28).

Human PRMT1 (*Hs*PRMT1) was expressed from a pET28b(+) vector with a short N-terminal His tag obtained from Dr. Paul Thompson (University of Massachusetts Medical School, Worcester, MA) (38). Rat PRMT1 (*Rn*PRMT1) M48F was expressed from a pET28b(+) vector obtained from Dr. Joan Hevel (Utah State University, Logan, UT) (24). Both constructs were expressed in *Escherichia coli* BL21(DE3) cells (Invitrogen) and grown in LB media con-

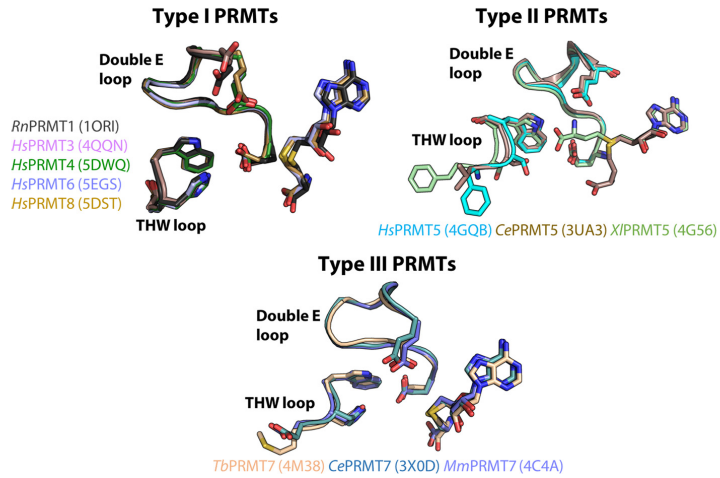


FIGURE 9. **Structural alignment of PRMT active sites.** Active sites of all three types of PRMTs are shown.

TABLE 2

Root mean square deviation (r.m.s.d.) values for structural alignments of the active-site double E loop, the THW loop, and AdoHcy made in PyMOL for type I, II, and III PRMTs from the indicated crystal structures

		r.m.s.d. (Å)	
	PRMTs	Cα	All atoms
Type I			
	*Rn*PRMT1 (1ORI)	0	0
	*Hs*PRMT3 (4QQN)	0.5	1.2
	*Hs*PRMT4 (5DWQ)	0.6	1.3
	*Hs*PRMT6 (5EGS)	0.7	1.2
	*Hs*PRMT8 (5DST)	0.6	1.1
Type II			
	*Hs*PRMT5 (4GQB)	0	0
	*Ce*PRMT5 (3UA3)	0.3	0.6
	*Xl*PRMT5 (4G56)	0.4	1.6
Type III			
	*Tb*PRMT7 (4M38)	0	0
	*Ce*PRMT7 (3X0D)	0.7	1.0
	*Mm*PRMT7 (4C4A)	0.7	1.1

taining kanamycin at 37 °C to an OD_{600} of ~0.6. Expression was induced with 1 mM isopropyl β-D-thiogalactoside (GoldBio) at 18 °C for 16 h. The cells were then harvested by centrifugation at 5,000 × *g* and 4 °C. The harvested cells were lysed using an EmulsiFlex cell homogenizer (Avestin) in 50 mM HEPES (pH 8.0), 300 mM NaCl, 0.5 mM phenylmethylsulfonyl fluoride (Sigma), and complete EDTA-free protease inhibitor mixture (Pierce). Lysed cells were centrifuged at 15,000 rpm for 50 min at 4 °C. The clarified lysate was loaded onto a 5-ml HisTrap HP Ni^{2+} column (GE Healthcare). The column was washed with 10 column volumes of the lysis buffer, including 50 mM imidazole-HCl (pH 8.0), and the protein was eluted with a 50–500 mM imidazole-HCl (pH 8.0) gradient. The eluted protein's purity was verified through SDS-PAGE analysis to be >95% (~40.6 kDa). The protein was then dialyzed against a storage buffer containing 50 mM HEPES (pH 8.0), 1 mM DTT, and 15% glycerol (v/v).

Isothermal Titration Calorimetry—ITC measurements were performed at 15 °C using a MicroCal auto-iTC200 calorimeter (MicroCal, LLC). Protein was incubated with 2-fold molar excess of AdoHcy for 1 h at room temperature. Protein and peptide samples were then extensively dialyzed against a buffer containing 20 mM HEPES (pH 7.5), 20 mM NaCl, and 0.5 mM tris(2-carboxyethyl)phosphine. 2 μl of 1–4 mM peptide was injected into 0.2 ml of 0.1–0.4 mM protein in the chamber every 150 s. Baseline-corrected data were analyzed with ORIGIN software.

Amino Acid Analysis of Protein and Peptide Substrates—*In vitro* methylation assays and amino acid analysis using the *Tb*PRMT7 wild-type and mutant enzymes were performed as described previously (25) in a buffer of 50 mM HEPES (pH 8.0), 10 mM NaCl, 1 mM DTT, and 5% glycerol in a final volume of 60 μl. Assays and amino acid analysis using human PRMT9 were also carried out as described previously (28, 29). For methylation assays with PRMT1, human and rat enzymes were used. The wild-type control was done with human PRMT1, and the mutant reactions were done with rat PRMT1 M48F. In both cases, 2.5 μg of PRMT1 and either 50 μM H4(1–21) or H4(1–21) R3MMA peptide were incubated at 37 °C for 3 h in a mixture containing 0.7 μM of *S*-adenosyl-L-[*methyl*-^{3}H]methionine ([*methyl*-^{3}H]AdoMet) (PerkinElmer Life Sciences; stock solution of 7 μM (78.2 Ci/mmol) in 10 mM H_2SO_4/EtOH (9:1, v/v)), 50 mM HEPES (pH 8.0), 10 mM NaCl, 1 mM DTT, and 5% glycerol in a final volume of 60 μl. Reactions were stopped, acid-hydrolyzed, and analyzed with cation exchange chromatography as described previously (25). Given the specific radioactivity of the [*methyl*-^{3}H]-AdoMet of 78.2 Ci/mmol and a counting efficiency of 50%, 1 fmol of methyl groups corresponds to 86 cpm.

Author Contributions—K. J., R. A. W., E. W. D., A. H., P. S., and S. G. C. planned the experiments. Experiments were performed by K. J., R. A. W., E. W. D., and A. H. K. J. and R. A. W. drafted the manuscript. K. J., R. A. W., E. W. D., A. H., P. S., and S. G. C. participated in the data analysis and interpretation, and K. J., R. A. W., E. W. D., A. H., P. S., and S. G. C. participated in manuscript revisions.

Acknowledgments—We thank H. Rust (The Scripps Research Institute) and P. Thompson (University of Massachusetts Medical School) for providing the H4(1–21) and H4(1–21) R3MMA peptides and the human His-PRMT1 plasmid; J. Hevel (Utah State University, Logan, UT) for providing the rat His-PRMT1 M48F plasmid; M. Dzialo (University of Leuven, Belgium) for help with amino acid analysis; G. Blobel (Rockefeller University) for support; D. Berman (Rockefeller University) for help with protein expression and purification; D. King (University of California at Berkeley) for mass spectrometry analysis; and the High Throughput Screening and Spectroscopy Resource Center at Rockefeller University.

References

1. Walsh, G., and Jefferis, R. (2006) Post-translational modifications in the context of therapeutic proteins. *Nat. Biotechnol.* **24,** 1241–1252
2. Fuhrmann, J., Clancy, K. W., and Thompson, P. R. (2015) Chemical biology of protein arginine modifications in epigenetic regulation. *Chem. Rev.* **115,** 5413–5461
3. Morales, Y., Cáceres, T., May, K., and Hevel, J. M. (2016) Biochemistry and regulation of the protein arginine methyltransferases (PRMTs). *Arch. Biochem. Biophys.* **590,** 138–152
4. Baldwin, R. M., Haghandish, N., Daneshmand, M., Amin, S., Paris, G., Falls, T. J., Bell, J. C., Islam, S., and Côté, J. (2015) Protein arginine methyltransferase 7 promotes breast cancer cell invasion through the induction of MMP9 expression. *Oncotarget.* **6,** 3013–3032
5. Yao, R., Jiang, H., Ma, Y., Wang, L., Wang, L., Du, J., Hou, P., Gao, Y., Zhao, L., Wang, G., Zhang, Y., Liu, D.-X., Huang, B., and Lu, J. (2014) PRMT7 induces epithelial-to-mesenchymal transition and promotes metastasis in breast cancer. *Cancer Res.* **74,** 5656–5667
6. Yang, Y., and Bedford, M. T. (2013) Protein arginine methyltransferases and cancer. *Nat. Rev. Cancer* **13,** 37–50
7. Karkhanis, V., Wang, L., Tae, S., Hu, Y. J., Imbalzano, A. N., and Sif, S. (2012) Protein arginine methyltransferase 7 regulates cellular response to DNA damage by methylating promoter histones H2A and H4 of the polymerase δ catalytic subunit gene, POLD1. *J. Biol. Chem.* **287,** 29801–29814
8. Wang, Y.-C., Peterson, S. E., and Loring, J. F. (2014) Protein post-translational modifications and regulation of pluripotency in human stem cells. *Cell Res.* **24,** 143–160
9. Lott, K., Zhu, L., Fisk, J. C., Tomasello, D. L., and Read, L. K. (2014) Functional interplay between protein arginine methyltransferases in *Trypanosoma brucei*. *Microbiologyopen* **3,** 595–609
10. Ferreira, T. R., Alves-Ferreira, E. V., Defina, T. P., Walrad, P., Papadopoulou, B., and Cruz, A. K. (2014) Altered expression of an RBP-associated arginine methyltransferase 7 in *Leishmania major* affects parasite infection. *Mol. Microbiol.* **94,** 1085–1102
11. Bedford, M. T., and Clarke, S. G. (2009) Protein arginine methylation in mammals: who, what, and why. *Mol. Cell* **33,** 1–13
12. Martens-Lobenhoffer, J., Bode-Böger, S. M., and Clement, B. (2016) First detection and quantification of *N*(δ)-monomethylarginine, a structural isomer of *N* (G)-monomethylarginine, in humans using MS(3). *Anal. Biochem.* **493,** 14–20
13. Wang, H., Huang, Z. Q., Xia, L., Feng, Q., Erdjument-Bromage, H., Strahl, B. D., Briggs, S. D., Allis, C. D., Wong, J., Tempst, P., and Zhang, Y. (2001) Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor. *Science* **293,** 853–857
14. Sun, L., Wang, M., Lv, Z., Yang, N., Liu, Y., Bao, S., Gong, W., and Xu, R.-M. (2011) Structural insights into protein arginine symmetric dimethylation by PRMT5. *Proc. Natl. Acad. Sci.* **108,** 20538–20543
15. Dhar, S. S., Lee, S.-H., Kan, P.-Y., Voigt, P., Ma, L., Shi, X., Reinberg, D., and Lee, M. G. (2012) Trans-tail regulation of MLL4-catalyzed H3K4 methylation by H4R3 symmetric dimethylation is mediated by a tandem PHD of MLL4. *Genes Dev.* **26,** 2749–2762
16. Migliori, V., Müller, J., Phalke, S., Low, D., Bezzi, M., Mok, W. C., Sahu, S. K., Gunaratne, J., Capasso, P., Bassi, C., Cecatiello, V., De Marco, A., Blackstock, W., Kuznetsov, V., Amati, B., *et al.* (2012) Symmetric dimethylation of H3R2 is a newly identified histone mark that supports euchromatin maintenance. *Nat. Struct. Mol. Biol.* **19,** 136–144
17. Gayatri, S., and Bedford, M. T. (2014) Readers of histone methylarginine marks. *Biochim. Biophys. Acta* **1839,** 702–710
18. Suárez-Calvet, M., Neumann, M., Arzberger, T., Abou-Ajram, C., Funk, E., Hartmann, H., Edbauer, D., Kremmer, E., Göbl, C., Resch, M., Bourgeois, B., Madl, T., Reber, S., Jutzi, D., Ruepp, M.-D., *et al.* (2016) Monomethylated and unmethylated FUS exhibit increased binding to Transportin and distinguish FTLD-FUS from ALS-FUS. *Acta Neuropathol.* **131,** 587–604
19. Dhar, S., Vemulapalli, V., Patananan, A. N., Huang, G. L., Di Lorenzo, A., Richard, S., Comb, M. J., Guo, A., Clarke, S. G., and Bedford, M. T. (2013) Loss of the major type I arginine methyltransferase PRMT1 causes substrate scavenging by other PRMTs. *Sci. Rep.* **3,** 1311
20. Fuhrmann, J., and Thompson, P. R. (2016) Protein arginine methylation and citrullination in epigenetic regulation. *ACS Chem. Biol.* **11,** 654–668
21. Schapira, M., and Ferreira de Freitas, R. (2014) Structural biology and chemistry of protein arginine methyltransferases. *Medchemcomm* **5,** 1779–1788
22. Cura, V., Troffer-Charlier, N., Wurtz, J. M., Bonnefond, L., and Cavarelli, J. (2014) Structural insight into arginine methylation by the mouse protein arginine methyltransferase 7: a zinc finger freezes the mimic of the dimeric state into a single active site. *Acta Crystallogr. D Biol. Crystallogr* **70,** 2401–2412
23. Hasegawa, M., Toma-Fukai, S., Kim, J. D., Fukamizu, A., and Shimizu, T. (2014) Protein arginine methyltransferase 7 has a novel homodimer-like structure formed by tandem repeats. *FEBS Lett.* **588,** 1942–1948
24. Gui, S., Gathiaka, S., Li, J., Qu, J., Acevedo, O., and Hevel, J. M. (2014) A remodeled protein arginine methyltransferase 1 (PRMT1) generates symmetric dimethylarginine. *J. Biol. Chem.* **289,** 9320–9327
25. Debler, E. W., Jain, K., Warmack, R. A., Feng, Y., Clarke, S. G., Blobel, G., and Stavropoulos, P. (2016) A glutamate/aspartate switch controls product specificity in a protein arginine methyltransferase. *Proc. Natl. Acad. Sci. U.S.A.* **113,** 2068–2073
26. Fisk, J. C., Sayegh, J., Zurita-Lopez, C., Menon, S., Presnyak, V., Clarke, S. G., and Read, L. K. (2009) A type III protein arginine methyltransferase from the protozoan parasite *Trypanosoma brucei*. *J. Biol. Chem.* **284,** 11590–11600
27. Wang, C., Zhu, Y., Caceres, T. B., Liu, L., Peng, J., Wang, J., Chen, J., Chen, X., Zhang, Z., Zuo, X., Gong, Q., Teng, M., Hevel, J. M., Wu, J., and Shi, Y. (2014) Structural determinants for the strict monomethylation activity by *Trypanosoma brucei* protein arginine methyltransferase 7. *Structure* **22,** 756–768
28. Hadjikyriacou, A., Yang, Y., Espejo, A., Bedford, M. T., and Clarke, S. G. (2015) Unique features of human protein arginine methyltransferase 9 (PRMT9) and its substrate RNA splicing factor SF3B2. *J. Biol. Chem.* **290,** 16723–16743
29. Yang, Y., Hadjikyriacou, A., Xia, Z., Gayatri, S., Kim, D., Zurita-Lopez, C., Kelly, R., Guo, A., Li, W., Clarke, S. G., and Bedford, M. T. (2015) PRMT9 is a type II methyltransferase that methylates the splicing factor SAP145. *Nat. Commun.* **6,** 6428
30. Feng, Y., Maity, R., Whitelegge, J. P., Hadjikyriacou, A., Li, Z., Zurita-Lopez, C., Al-Hadid, Q., Clark, A. T., Bedford, M. T., Masson, J. Y., and Clarke, S. G. (2013) Mammalian protein arginine methyltransferase 7 (PRMT7) specifically targets RXR sites in lysine- and arginine-rich regions. *J. Biol. Chem.* **288,** 37010–37025
31. Feng, Y., Xie, N., Jin, M., Stahley, M. R., Stivers, J. T., and Zheng, Y. G. (2011) A transient kinetic analysis of PRMT1 catalysis. *Biochemistry* **50,** 7033–7044
32. Huang, S., Litt, M., and Felsenfeld, G. (2005) Methylation of histone H4 by arginine methyltransferase PRMT1 is essential in vivo for many subsequent histone modifications. *Genes Dev.* **19,** 1885–1893
33. Wang, M., Xu, R.-M., and Thompson, P. R. (2013) Substrate specificity, processivity, and kinetic mechanism of protein arginine methyltransferase 5. *Biochemistry* **52,** 5430–5440
34. Bao, X., Zhao, S., Liu, T., Liu, Y., Liu, Y., and Yang, X. (2013) Overexpression of PRMT5 promotes tumor cell growth and is associated with poor

disease prognosis in epithelial ovarian cancer. *J. Histochem. Cytochem.* **61**, 206–217
35. Tarighat, S. S., Santhanam, R., Frankhouser, D., Radomska, H. S., Lai, H., Anghelina, M., Wang, H., Huang, X., Alinari, L., Walker, A., Caligiuri, M. A., Croce, C. M., Li, L., Garzon, R., Li, C., *et al.* (2016) The dual epigenetic role of PRMT5 in acute myeloid leukemia: gene activation and repression via histone arginine methylation. *Leukemia* **30**, 789–799
36. Schapira, M., and Arrowsmith, C. H. (2016) Methyltransferase inhibitors for modulation of the epigenome and beyond. *Curr. Opin. Chem. Biol.* **33**, 81–87
37. Hu, H., Qian, K., Ho, M.-C., and Zheng, Y. G. (2016) Small molecule inhibitors of protein arginine methyltransferases. *Expert Opin. Investig. Drugs.* **25**, 335–358
38. Zhang, X., and Cheng, X. (2003) Structure of the predominant protein arginine methyltransferase PRMT1 and analysis of its binding to substrate peptides. *Structure* **11**, 509–520
39. Zurita-Lopez, C. I., Sandberg, T., Kelly, R., and Clarke, S. G. (2012) Human protein arginine methyltransferase 7 (PRMT7) is a type III enzyme forming ω-N^G-monomethylated arginine residues. *J. Biol. Chem.* **287**, 7859–7870
40. Gottschling, H., and Freese, E. (1962) A tritium isotope effect on ion exchange chromatography. *Nature* **196**, 829–831