MECHANISTIC AND BIOLOGICAL INSIGHTS INTO CARBOHYDRATE AND LIPID MODIFICATIONS OF PROTEINS

by

JIANYE ZHANG

Submitted in partial fulfillment of the requirements
For the degree of Doctor of Philosophy

Dissertation Advisors: Dr. Lawrence M. Sayre
and Dr. Gregory P. Tochtrop
Co-Advisor: Dr. Vincent M. Monnier

Department of Chemistry
CASE WESTERN RESERVE UNIVERSITY

August 2011
We hereby approve the thesis/dissertation of

Jianye Zhang

(candidate for the ________________________________ degree *

(signed) ____________________________________________

Robert G. Salomon

(chair of the committee)

Malcolm E. Kenney

Vincent M. Monnier

Rajesh Viswanathan

Gregory P. Tohtrop

(date) __________________________________________

06/06/2011

*We also certify that written approval has been obtained for any proprietary material contained therein.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title page</td>
<td>i</td>
</tr>
<tr>
<td>Signature sheet</td>
<td>ii</td>
</tr>
<tr>
<td>Table of Contents</td>
<td>iii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>vii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>viii</td>
</tr>
<tr>
<td>List of Schemes</td>
<td>xi</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>xiii</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>xv</td>
</tr>
<tr>
<td>Abstract</td>
<td>xix</td>
</tr>
<tr>
<td><strong>Chapter 1. Introduction</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1. Carbonyl and oxidative post-translational modifications to proteins</td>
<td>2</td>
</tr>
<tr>
<td>1.2. Glycation and advanced glycation end products (AGEs)</td>
<td>2</td>
</tr>
<tr>
<td>1.3. Sugar-derived protein cross-links</td>
<td>8</td>
</tr>
<tr>
<td>1.4. Lipoxidation and advanced lipoxidation end products (ALEs)</td>
<td>11</td>
</tr>
<tr>
<td>1.5. Epoxyketoctadecenoic acid (EKODE)</td>
<td>15</td>
</tr>
</tbody>
</table>
Chapter 2. Identification of sugar-derived modification and cross-linking sites in human alpha A crystallin

2.1. Introduction

2.2. Results

2.2.1. Cloning and expression of recombinant alpha A crystallin and its mutants

2.2.2. Enzymatic digestion of αA crystallin and its mutants

2.2.3. AGE modifications on wild type αA crystallin by glucose

2.2.3.1. Fructosyl lysine (FL) in the peptides

2.2.3.2. Carboxyethyl lysine (CEL) in the peptides

2.2.3.3. Glycation of arginine residues: and formation of 3-deoxyglucosone derived hydroimidazolones (3-DG-H)

2.2.3.4. Methylglyoxal hydroimidazolone (MG-H1)

2.2.3.5. The glucose-derived cross-links

2.3. Discussion

2.3.1 K166 is the hot spot for glycation and glycoxidation

2.3.2. Glutamic acid catalyzing the modification of dicarbonyl compounds on the arginine residues.

2.3.3. The hotspots for glucose-derived cross-links

2.3.4. The protective effect of oxidation against the glucose-derived cross-links

2.4. Conclusions
Chapter 3. Anaerobic vs aerobic pathways of carbonyl and oxidant stress in human lens and skin during aging and in diabetes: a comparative analysis

3.1. Introduction

3.2. Results

3.2.1. The preparation of isotope-labeled AGE standards

3.2.2. From glucose to glucosepane

3.2.3. Glycoxidation vs glyoxal-dependent mechanism of CML formation

3.2.4. Origin and significance of methylglyoxal modifications in lens and skin

3.2.5. Metal-catalyzed and hydrogen peroxide-mediated oxidation

3.2.6. The hierarchy of protein damage during aging

3.3. Discussion

3.3.1. The glycoxidation of fructosyl lysine to CML is favored in lens

3.3.2. Glucosepane is a better marker for diabetes than CML

3.3.3. Multiple sources for dicarbonyl species formation

3.3.4. Lysyl oxidation was more prevalent in skin than in lens

3.4. Conclusions

3.5. Experimental section

3.6. References
4.1. Introduction

4.2. Results and discussion

4.2.1. Preparation of EKODE model compounds

4.2.2. The reactions of EKODE I models with butanethiol

4.2.3. The reactions of an EKODE II model with butanethiol

4.2.4. The reactions of EKODE model compounds with imidazole

4.2.5. The reactions of EKODE model compounds with lysine

4.3. Conclusion

4.4. Experimental section

4.5. References

Appendix. Preparation of the antigen to KODA-derived ketoamide

Bibliography
List of Tables

Chapter 2

Table 1  AGE modifications being investigated  38
Table 2  The mono adducts on the lysine residues of the glycated wild type αA crystallin  41
Table 3  Dicarbonyl compound derived modifications, 3-DG-H and MG-H1  46
Table 4  The cross-linking peptides  49

Chapter 3

Table 1  Effect of diabetes on the relationship between age-related accumulation of protein modifications in lens and skin as reflected by Spearman correlation coefficients ranked according to R value as a function of age and presence of diabetes.  98
# List of Figures

## Chapter 1

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Formation pathways of the major AGEs</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>The major AGE structures: The mono adducts of AGEs</td>
<td>6</td>
</tr>
<tr>
<td>3</td>
<td>The cross-links of AGEs</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>The oxidation of polyunsaturated lipids.</td>
<td>13</td>
</tr>
</tbody>
</table>

## Chapter 2

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SDS-PAGE of crystallin proteins separated using a 4-12% gel.</td>
<td>35</td>
</tr>
<tr>
<td>2</td>
<td>Identification of the glycated α-crystallin peptide VQDDFVEIHGK*HNER in LC-MS.</td>
<td>40</td>
</tr>
<tr>
<td>3</td>
<td>Identification of the peptide modified by FL- H₂O</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Identification of EEKPTSAPSS+160 peptide</td>
<td>42</td>
</tr>
<tr>
<td>5</td>
<td>Tandem mass spectrum of CEL modified peptide AIPVSREEK*PTSAPSS.</td>
<td>43</td>
</tr>
<tr>
<td>6</td>
<td>Tandem mass spectrum of 3-DG-H modified peptide TVLDSGISEVR*SDR.</td>
<td>47</td>
</tr>
<tr>
<td>7</td>
<td>Tandem mass spectrum of MGH1 modified peptide AIPVSR*EEKPTSAPSS.</td>
<td>48</td>
</tr>
<tr>
<td>8</td>
<td>The intra cross-links of glucosepane and DODIC</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Modifications of AIPVSREEKPTSAPSS by 3-DG-H and intra-molecular cross-links glucosepane and DOGDIC</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>51</td>
</tr>
</tbody>
</table>
AGEs levels in wild type crystallin and its mutants K166A (KA), K166A/R163A (KARA), K166/R(KR), R163(RA).

Figure 11 Aspartic acid residues may form an intra-molecular hydrogen bond through a six member ring in proteins, which may decrease the pKa of aspartic acid comparing with glutamic acid.

Figure 12 The crystal structure of truncated bovine αA crystallin (59-163) dimer.

Chapter 3

Figure 1 Age-related changes in levels of fructose-lysine and glucosepane in insoluble lens crystallins and skin collagen-rich fraction

Figure 2 Kinetic studies on collagen glycation

Figure 3 Age-related changes in CML, G-H1, and pentosidine in insoluble lens crystallins and skin collagen-rich fraction.

Figure 4 Age-related changes in levels of MG-H1 and CEL in insoluble lens crystallins and skin collagen-rich fraction.

Figure 5 Formation of AGEs in bovine lens crystallin homogenate incubated with 5 mM ascorbic acid or DHA for 7 days in 100 mM phosphate buffer containing 1 mM DTPA chelating agent.

Figure 6 Age-related accumulation rates of allysine, 2-aminoadipic acid, pentosidine, and ornithine in skin vs lens.
Figure 7  Quantitative comparisons of mean levels ± SD for AGEs and lysine oxidation levels reached at age ≥ 65 years for insoluble lens crystallins and skin collagen-rich fraction.

Figure 8  The glycoxidation efficiency in crystallin vs in collagen.

Chapter 4

Figure 1  The EKODE isomers identified.

Figure 2  UV-monitoring of the reactions of 1 with butanethiol or N-acetyl cysteine methyl ester

Figure 3  The reactions of 1 to N-acetyl cysteine methyl ester, monitored by 600 MHz NMR.

Figure 4  The HHCOSY spectrum of compound 16.

Figure 5  The reactions of 2 with butanethiol or N-acetyl cysteine methyl ester

Figure 6  The reactions of 2 with N-acetyl-cysteine methyl ester (1 mM) in sodium phosphate buffer pH 7.4 containing 5% d6-DMSO

Figure 7  The reaction of 2 to imidazole.

Figure 8  Identification of the two isomers of imidazole Michael adducts to an EKODE II model by HHCOSY
## List of Schemes

### Chapter 1

**Scheme 1**  
The mechanism of the glucosepane and DOGDIC formation 10

**Scheme 2**  
The pathways to form HNE and ONE adducts. 14

**Scheme 3**  
The synthesis of ONE-$d_9$ 15

**Scheme 4**  
The preparation of the antigen to KODA-derived ketoamide 15

**Scheme 5**  
Proposed routes for linoleic acid oxidation 17

### Chapter 2

**Scheme 1**  
Proposed mechanism to account for the preferential glycation of K166 57

**Scheme 2**  
Proposed mechanism for the formation of CEL and CML 57

**Scheme 3**  
The mechanism of modification of arginine by dicarbonyl compounds 60

**Scheme 4**  
Mechanism proposed by Lederer et al. for glucose-derived cross-links 61

**Scheme 5**  
The proposed mechanism for the formation of methionine sulfoxide from the oxidation of fructosyl lysine 65

### Chapter 3

**Scheme 1**  
The synthesis routes for $d_3$-2-amino adipic acid (A), $d_2$-$^{13}$C$_1$-CEL (B), $d_2$-G-H1 (C) and $d_3$-MG-H1 (D). 80
<table>
<thead>
<tr>
<th>Scheme 1</th>
<th>The synthesis of EKODE model compounds</th>
<th>117</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scheme 2</td>
<td>The reactions of butanethiol with EKODE I model compounds</td>
<td>119</td>
</tr>
<tr>
<td>Scheme 3</td>
<td>The proposed mechanism for the reaction of thiol with EKODE I</td>
<td>119</td>
</tr>
<tr>
<td>Scheme 4</td>
<td>The reaction of compound 2 to butanethiol</td>
<td>125</td>
</tr>
<tr>
<td>Scheme 5</td>
<td>The proposed mechanism of the EKODE-II model epoxy ring opening reaction</td>
<td>126</td>
</tr>
<tr>
<td>Scheme 6</td>
<td>The reactions of EKODE II analogues with butanethiol.</td>
<td>126</td>
</tr>
</tbody>
</table>
Acknowledgement

First of all, I would like to express my sincere appreciation to my ex-advisor, Dr. Lawrence Sayre, for his generous support and guidance since 2005 even during a funding tough period. His intelligence and diligence are a precious legacy to my research career.

I would like to express my deepest thanks to the generous support from Dr. Vincent Monnier in last six years. His intelligent instructions gave me the best training in research. He helped to me realize the importance of rigor in science.

I also would like to give my deepest thanks to Dr. Gregory Tochtrop for his generous help and support to my research since Dr. Sayre’s death. His intelligence and passion inspired my research greatly. Without his enormous guidance and help, I would not have completed this dissertation. I also would like to thank Dr. tochtrop’s help for my career.

I would like to thank my committee, Dr. Robert Salomon, Dr. Malcolm Kenney, Dr. Rajesh Viswanathan, and Dr. Vincent Monnier for their time and help reviewing this dissertation. And I would like to thank Dr. Ram Nagaraj for generous gift of the plasmids of alpha crystallin.

I would specially thank my labmate, Chuan Shi, for the great help to chapter 4. I would like to thank Dr. Xingjun Fan for designing the primers of αA crystallin mutants and the great help to chapter 3. I would like to thank the great help from Dr. David sell for the preparation of human samples and statistic analysis, and Christopher Strauch for the help in LC-MS and GC/MS analysis.

I would like to thank Dr. Jim Faulk for his help on HPLC-MS technical support and Dr. Dale Ray for help on 600 MHz NMR experiments. I also would like to thank Dr.
Benlian Wang, Kathleen Lundberg, Dr. Janna Kiselar, and Dr. Masaru Miyagi in the Center for Proteomics and Bioinformatics for the help with mass spectrometry, and Dr. Xiongyin Tu and Dr Viven Yee for help with modeling of the crystal structure of αA crystallin.

I would like to thank the ex-lab members in Dr. Monnier’s lab, Dr. Ina Nemet, Xiaoqing Liu, Shuyu Hao, Dr. Francois Collard, Dr. Zhenyu Dai, and Juan Qian, and in Dr. Sayre’s lab, Dr. Yuming Zhang, Dr. Keqing Lin, Dr. De Lin, Dr. Xiaochun Zhu, Dr. Bo Zhong, Dr. Yuming Chen, Detao Gao, Xibo Li for their help to my research. I also would like thank the lab members in Dr. Tochtrop’s lab, Dr. Yong Han, Xiaoxia Tang, Sushabhan Sadhukhan, and Qingjiang Li for their help.

I would like to specially thank my family and friends to support my pursuing the Ph. D degree.

Finally, I will thank Case Western Reserve University for the past seven years life. The research in chapter 2 and 3 are supported by NEI EY07099, and the research in chapter is supported by NHLBI HL053315.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-DG</td>
<td>3-deoxyglucosone</td>
</tr>
<tr>
<td>3-DGH</td>
<td>3-deoxyglucosone-derived hydroimidazolone</td>
</tr>
<tr>
<td>$[^{13}C_6]$/FL</td>
<td>$[^{13}C_6]$-fructosyl lysine</td>
</tr>
<tr>
<td>AAA</td>
<td>amino adipic acid</td>
</tr>
<tr>
<td>AGEs</td>
<td>advanced glycation end products</td>
</tr>
<tr>
<td>ALEs</td>
<td>advanced lipoxidation end products</td>
</tr>
<tr>
<td>BCA</td>
<td>bicinchoninic acid</td>
</tr>
<tr>
<td>BSA</td>
<td>bovine serum albumin</td>
</tr>
<tr>
<td>CEL</td>
<td>carboxyethyl-lysine</td>
</tr>
<tr>
<td>CHCA</td>
<td>α-cyano-4-hydroxycinnamic acid</td>
</tr>
<tr>
<td>CML</td>
<td>carboxymethyl-lysine</td>
</tr>
<tr>
<td>$d_2-[^{13}C_1]/CEL$</td>
<td>$d_2-[^{13}C_1]$-carboxyethyl lysine</td>
</tr>
<tr>
<td>$d_2$/G-H1</td>
<td>$d_2$-glyoxal hydroimidazolone</td>
</tr>
<tr>
<td>$d_3$/MG-H1</td>
<td>$d_3$-methylglyoxal hydroimidazolone</td>
</tr>
<tr>
<td>DHA</td>
<td>dehydroascorbic acid</td>
</tr>
<tr>
<td>DOGDIC</td>
<td>3-deoxyglucosone derived imidazolium cross-link</td>
</tr>
<tr>
<td>DOLD</td>
<td>3-deoxyglucosone derived lysine dimer</td>
</tr>
<tr>
<td>DTPA</td>
<td>diethylene triamine pentaacetic acid</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>EH</td>
<td>5,6-epoxy-2-heptenal</td>
</tr>
<tr>
<td>EKODE</td>
<td>epoxyketo octadecenoic acid</td>
</tr>
<tr>
<td><strong>ELISA</strong></td>
<td>enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td><strong>ESI-CID</strong></td>
<td>electrospray ionization-collision induced dissociation</td>
</tr>
<tr>
<td><strong>ESI-LC-MS</strong></td>
<td>electrospray ionization liquid chromatography-mass spectrometry</td>
</tr>
<tr>
<td><strong>ESI-MRM</strong></td>
<td>electrospray ionization- multiple reaction monitoring</td>
</tr>
<tr>
<td><strong>ESRD</strong></td>
<td>end stage renal diseases</td>
</tr>
<tr>
<td><strong>FABP</strong></td>
<td>fatty acid-binding protein</td>
</tr>
<tr>
<td><strong>FAOX</strong></td>
<td>fructosamine oxidases</td>
</tr>
<tr>
<td><strong>FL</strong></td>
<td>fructosyl-lysine</td>
</tr>
<tr>
<td><strong>FN3K</strong></td>
<td>fructosylamine 3-phosphokinase</td>
</tr>
<tr>
<td><strong>GAPDH</strong></td>
<td>glyceraldehyde-3-phosphate dehydrogenase</td>
</tr>
<tr>
<td><strong>GC/MS</strong></td>
<td>gas chromatography/mass spectrometry</td>
</tr>
<tr>
<td><strong>G-H1</strong></td>
<td>glyoxal-derived hydromimidazolone</td>
</tr>
<tr>
<td><strong>GO</strong></td>
<td>glyoxal</td>
</tr>
<tr>
<td><strong>GODIC</strong></td>
<td>glyoxal derived imidazolium cross-link</td>
</tr>
<tr>
<td><strong>GOLD</strong></td>
<td>gloxal derived lysine dimer</td>
</tr>
<tr>
<td><strong>HbA1c</strong></td>
<td>glycated hemoglobin</td>
</tr>
<tr>
<td><strong>H-D</strong></td>
<td>hydrogen-deuterium</td>
</tr>
<tr>
<td><strong>HHCOSY</strong></td>
<td>$^1$H-$^1$H correlation spectroscopy</td>
</tr>
<tr>
<td><strong>HNE</strong></td>
<td>4-hydroxy-2-nonenal</td>
</tr>
<tr>
<td><strong>HODA</strong></td>
<td>9-hydroxy-12-oxo-10-dodecenoic acid</td>
</tr>
<tr>
<td><strong>HPLC</strong></td>
<td>high performance liquid chromatography</td>
</tr>
<tr>
<td><strong>HSA</strong></td>
<td>human serum albumin</td>
</tr>
<tr>
<td><strong>IPTG</strong></td>
<td>isopropyl β-D-thiogalactopyranoside</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>KODA</td>
<td>9-keto-12-oxo-10-dodecenoic acid</td>
</tr>
<tr>
<td>LC/MS</td>
<td>liquid chromatography/mass spectrometry</td>
</tr>
<tr>
<td>LC-MS/MRM</td>
<td>lc-ms multiple reaction monitoring</td>
</tr>
<tr>
<td>LOOHs</td>
<td>lipid hydroperoxides</td>
</tr>
<tr>
<td>mAb</td>
<td>monoclonal antibody</td>
</tr>
<tr>
<td>MALDI</td>
<td>matrix-assisted laser desorption/ionization</td>
</tr>
<tr>
<td>MDA</td>
<td>malondialdehyde</td>
</tr>
<tr>
<td>MG-H1</td>
<td>methylglyoxal derived hydroimidazolone</td>
</tr>
<tr>
<td>MGO</td>
<td>methylglyoxal</td>
</tr>
<tr>
<td>MODIC</td>
<td>methylglyoxal derived imidazolium cross-link</td>
</tr>
<tr>
<td>MOLD</td>
<td>methylglyoxal derived lysine dimer</td>
</tr>
<tr>
<td>MRM</td>
<td>multiple reaction monitoring</td>
</tr>
<tr>
<td>ONE</td>
<td>4-oxo-2-nonenal</td>
</tr>
<tr>
<td>ONL</td>
<td>N(_\varepsilon)-(4-oxononanoyl) lysine</td>
</tr>
<tr>
<td>OPD</td>
<td>o-phenylenediamine</td>
</tr>
<tr>
<td>OS</td>
<td>oxidative stress</td>
</tr>
<tr>
<td>PCC</td>
<td>pyridinium chlorochromate</td>
</tr>
<tr>
<td>PMSF</td>
<td>phenylmethanesulfonylfluoride</td>
</tr>
<tr>
<td>PTM</td>
<td>post translation modifications</td>
</tr>
<tr>
<td>PUFA</td>
<td>polyunsaturated fatty acids</td>
</tr>
<tr>
<td>RAGE</td>
<td>receptor of AGEs</td>
</tr>
<tr>
<td>RNase A</td>
<td>ribonuclease A</td>
</tr>
<tr>
<td>ROS</td>
<td>reactive oxygen species</td>
</tr>
<tr>
<td><strong>SDS-PAGE</strong></td>
<td>sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td><strong>sHSPs</strong></td>
<td>small heat shock proteins</td>
</tr>
<tr>
<td><strong>SIM</strong></td>
<td>selected ion monitoring</td>
</tr>
<tr>
<td><strong>THP</strong></td>
<td>tetrahydropyrimidine</td>
</tr>
<tr>
<td><strong>TIC</strong></td>
<td>total ion chromatograms</td>
</tr>
</tbody>
</table>
Mechanistic and Biological Insights into Carbohydrate and Lipid Modifications of Proteins

Abstract

by

Jianye Zhang

Glycation and lipoxidation are two typical contributors to irreversible post-translational modification in biological systems, which contribute to age-related diseases. In this dissertation, mechanistic studies are reported that elucidate the chemistry of glycation and lipoxodation, and their biological consequences.

To understand the contribution of glycation to diabetic cataract, glycated recombinant human αA crystallin was analyzed by mass spectrometry. The data revealed that R12, R65, R103, R112, R157 and R163 were hot spots of dicarbonyl modifications. The most prevalent site for glycation was K166, which was confirmed by the fact that the mutations of K166 can drastically decrease the levels of advanced glycation end products (AGEs). The selectivities were hypothetically attributed to the catalytic effect of glutamate adjacent lysine. The major sugar-derived cross-link was observed at K166-R163, K166-R103, and K166-R117. The latter two cross-link sites suggested a transient attraction between the C-terminus and the edge of β6 and β7 in αA crystallin, which may contribute to the inter-molecular attractions between crystallins.

The comparative analysis of carbonyl and oxidant stress in human lens and skin revealed that the major modifications in lens were ornithine, methylglyoxal-derived hydroimidazolone (MG-H1) and carboxymethyllysine (CML), whereas in skin were
ornithine, FL and glucosepane. The potential explanation to these facts is that lens proteins can more efficiently use glycoxidation to prevent sugar-derived cross-links. Most AGEs accumulated with aging both in lens and skin. However, diabetes had no impact on MG-H1, glyoxal hydroimidazolone, CML and carboxylethyllysine. These observations suggested that current diabetes medication can ameliorate dicarbonyl stress, but not glycation and consequent cross-linking. The strikingly high MG-H1 level in human lens originated from the high concentration of ascorbic acid in lens.

Finally, modifications derived from epoxyketoctadecenoic acids (EKODEs) were investigated. Both the reactions of EKODE I and II with buanethiol are initiated with Michael addition. The resulting adducts can be converted to epoxy ring opening products through an intra-molecular [1, 2] shift or [1, 5] shift of butylthio. The thiolysis reactions were prompted by both heat and light. Similar reactions were observed in the reactions of EKODEs with cysteine, and the modifications of EKODEs to histidine were dominated by Michael adducts.
Chapter 1

Introduction
1.1. Carbonyl and oxidative post-translational modifications (PTM) to proteins

The fine regulation of oxidative stress (OS) is important for maintenance of cell functions in aerobic organisms.\(^1\) A slight surge of reactive oxygen species (ROS) can cause unnecessary oxidative modifications of proteins, in particular to sulphhydryl residues. An array enzymes are involved in the repairing of such damages. If protein damage is too serious for repairing, it will be removed through enzymatic proteolysis by the proteasome system and lysosomes.\(^2\) However, these coping mechanisms are often eroded by the accumulation of irreversible damage.\(^3\) These modifications can be categorized into three kinds of reactions. The first one is the oxidative cleavage of the protein backbone or side chain residues such as proline, arginine, lysine, threonine and tryptophan.\(^4\) This kind of modification often reflects the acute prevalence of ROS. The second kind of damage is sugar-derived modifications, known as glycation, which is not an obligatory consequence of oxidation.\(^5\) The final one is damage caused by lipid peroxidation products.\(^6\) The last two types of modifications originate from relatively stable modifiers such as glucose or 4-hydroxy-2-nonenal (HNE) as compared to ROS, and the process can be chronic. During the last three decades, glycation and lipoxidation have been investigated due to their strong correlations with age-related chronic diseases.\(^7\) However the chemistry underlying glycation and lipoxidation has not been completely elucidated. This dissertation focuses on sugar/lipid-derived modifications of protein side chains and their biological consequences.

1.2. Glycation, and advanced glycation end products (AGEs)
In the last thirty years, broad interest in sugar-related post-translational modification, called glycation, has intensified when solid evidence revealed its strong correlation with age-related diseases such as diabetes, Alzheimer diseases, end stage renal diseases (ESRD), and cardiovascular diseases. Glycation is initiated by the adduction of sugars, such as glucose, ribose, and fructose or sugar derived-oxalaldehydes such as methyl glyoxal (MGO) and glyoxal (GO), to the primary amines. The resulting reversible products, Schiff bases, could be converted to the more stable compounds, called Amadori products, through an enolization process during the Maillard reaction. The Amadori products could further undergo a series of reactions, such as oxidation, cyclization, condensation, or fragmentation (Fig. 1). The resulting irreversible adducts and the Amadori products themselves are all called advanced glycation end products (AGEs). Most AGEs are quite stable under physiological conditions. However, glycation products on arginine residues such as hydroimidazolone or pyrimidine can spontaneously and slowly hydrolyze to form ornithine. Presently, more than 20 AGEs have been clearly identified in vitro and in vivo (Fig. 2 and 3), and the number is still increasing with the intensive investigation of the glycation process.

AGEs are excellent markers and predictors of age-related disease, especially for diabetic complications. The most famous glycation-related marker is glycated hemoglobin (HbA1c), which is clinically used to indicate blood glycemia level. Long-term studies on the development of diabetic complications in the DCCT and UKPDS have conclusively demonstrated a strong correlation between glycemia and the risk of developing diabetic complications. And more comprehensive studies revealed that pentosidine is associated with diabetic nephropathy and neuropathy, while
carboxymethyl-lysine (CML) is associated with the retinopathy and neuropathy.\textsuperscript{16} The Monnier lab recently showed that glucosepane, a glucose derived cross-linking product, more significantly correlated with diabetes and aging in the human skin and lens samples compared to CML and pentosidine.\textsuperscript{17}

\textbf{Figure 1.} Formation pathways of the major AGEs. Used the permission from Dr. Vincent Monnier.

However, the biological consequences of glycation are still incompletely elucidated. Even though it was reported that methylglyoxal derived hydroimidazolone (MG-H1) could benefit the chaperone function of alpha crystallin,\textsuperscript{18} most AGEs have deleterious effects on proteins, especially the long-lived proteins such as the human lens proteins and collagen. In the human lens, accumulation of AGEs can decrease the net charge on the protein surface, change protein interactions, and trigger protein aggregations.\textsuperscript{19} The accumulation of AGEs on the extracellular matrix proteins results in
the decrease of digestibility, solubility and elasticity.\textsuperscript{20} Significant research implied that AGEs bind to the proinflammatory receptor of AGEs (RAGE), triggering a signaling cascade with production of reactive oxygen species (ROS),\textsuperscript{21} and induction of apoptosis and anoikis of cells.\textsuperscript{8c, 22} AGEs also block the active sites of critical enzymes, and result in their dysfunctions.\textsuperscript{23} However, it is still difficult to delineate the role of AGEs in the development of diabetic complications. Are they major contributors to disease, or just markers of disease progress? It is still necessary to intensively investigate the role of glycation in biological systems.

Methodologies for quantitative analysis of AGEs in animal and human samples have been well developed. Liquid chromatography/mass spectrometry (LC/MS),\textsuperscript{24} gas chromatography/mass spectrometry (GC/MS),\textsuperscript{25} enzyme-linked immunosorbent assay (ELISA)\textsuperscript{26} and high performance liquid chromatography (HPLC)\textsuperscript{27} are widely used in the analysis. Of these methods, the LC/MS method in particular is becoming important because of several advantages.\textsuperscript{24, 28} First, LC/MS can analyze multiple AGEs in the same sample simultaneously. Second, the introduction of stable isotope-labeled AGEs as internal standards makes the quantification highly accurate and reproducible. Third, the samples are processed under near physiological conditions, which helps decrease the introduction of artifacts, and make most AGEs suitable for this analytical method. Consequently, LC/MS is still the most sensitive and accurate method for detection and quantitation of AGEs.\textsuperscript{28a}
Figure 2. The major AGEs structures: The mono adducts of AGEs. The configurations of all the chiral centers, except those are specified, are a mixture of R and S configurations.

However, the above methods cannot identify the modification sites in proteins. This makes proteomics methods for peptide mapping important for studies of the biological consequences of AGEs. With the applications of matrix-assisted laser desorption/ionization (MALDI) mass spectrometry and electrospray ionization-collision induced dissociation-LC/MS (ESI-CID-LC/MS), the fructosyl-lysine (FL) modification sites in ribonuclease A (RNase A), crystallins, and hemoglobin have been well identified. The glycation sites by fructose, galactose, lactose, ribose and ascorbic acid
were also characterized.\textsuperscript{33} It was reported that methyl-glyoxal (MGO) selectively modifies the RGD and GFOGER integrin-binding sites in collagen.\textsuperscript{22b} Glyoxal modifications in collagen or RNase were also reported.\textsuperscript{34} Most of these studies focused on only one or two kinds of AGEs. However, under physiological conditions, the damage to the proteins by glycation cannot be narrowed down to one or two kinds of reactions. A comprehensive analysis of the glycation hot spots on proteins would aid our understanding of the biological consequences of glycation.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{The cross-links of AGEs.}
\end{figure}
Specific coping mechanisms against glycation exist in biological systems. It was reported that a deglycating enzyme, fructosylamine 3-phosphokinase (FN3K), exists in the lens. FN3K can selectively phosphorylate fructosyllysine, which can result in the regeneration of lysine residues with a byproduct, 3-deoxyglucosone (3-DG). In vitro studies showed that FN3K can remove FL from glycated hemoglobin. However, the deglycation byproduct 3-DG is very reactive. It can form 3-deoxyglucosone-derived hydroimidazolone (3-DGH) or 3-deoxyglucosone-derived imidazolium cross-link (DOGDIC) on proteins and result in deleterious effects. Fructosamine oxidases (FAOX) are another kind of deglycating enzymes. They can convert fructosamine to a primary amine and glucosone. Recently, our group synthesized the FAOX-II inhibitor, fructosyl-thioacetate, and successfully used it to characterize the enzyme crystal structure. However FAOX has very limited deglycating efficacy for glycated proteins. On the other hand, Baynes et al. found that fructosyl lysine could be oxidized to form carboxymethyl lysine (CML). This reaction was accelerated in high concentration of sodium phosphate buffer, and inhibited by the metal ion chelating reagents such as pyridoxamine, suggesting a catalytic effect of redox metal ions. This oxidation can block the route from FL to glucosepane, though it results in the CML adduct formon.

1.3. Sugar-derived protein cross-links

Among all the AGEs, I am most interested in the sugar-derived cross-links, based on the hypothesis that cross-links have more deleterious effects on proteins compared to mono adducts. Since pentosidine, the first glycation-derived cross-link, was characterized by the Monnier group in 1989, more than ten cross-linking AGEs have been
characterized in vitro or in vivo studies (Fig. 3). In these modifications, glucosepane, directly derived from glucose, was first characterized by Lederer et al. in vitro. They originally thought it was derived from DOGDIC. But incubation of bovine serum albumin (BSA) with 3-deoxyglucosone (3-DG) was subsequently shown to only afford DOGDIC, which proved that DOGDIC and glucosepane are formed in two independent ways. Trapping experiments of the intermediate of glucosepane formation by incubating glucose, lysozyme and o-phenylenediamine (OPD) showed that glucosepane formation starts with the formation of FL, which was then converted to N\(^6\)-(2, 3-dihydroxy-5, 6-dioxohexyl)-L-lysinate by enolizations (Scheme 1). Biemel et al. postulated that an arginine near the fructosyl-lysine could catalyze the enolization to form the intermediate. A similar catalytic effect of arginine was also observed during glycation by Venketraman et al. They also showed that negatively charged residues, such as glutamic acid and aspartic acid, could catalyze the enolization to form fructosyl-lysine. Could negative charged residues also catalyze the formation of glucosepane? On the other hand, Watkins et al. proposed a mechanism whereby the phosphate near lysine residues would catalyze the enolization to form FL. The Monnier group also recently observed the catalytic effect of sodium phosphate on the formation of glucosepane.

Glucosepane is the most abundant sugar-derived cross-link in biological systems currently known. Its level is 3-10 times higher than that of the methylglyoxal-derived imidazolium cross-link (MODIC) in human serum albumin (HSA) samples compared to other sugar-related cross-links, such as the glyoxal-derived imidazolium cross-link (GODIC), methylglyoxal-derived lysine dimer (MOLD), glyoxal-derived lysine dimer (GOLD), and 3-deoxyglucosone-derived lysine dimer (DOLD) that are below detectable
level. Similarly in the human lens, glucosepane is the most abundant cross-link even though the MGO level is elevated due to the relatively high concentration of ascorbic acid in lens,\textsuperscript{50} which also results in increase of MOLD and MODIC. In collagen, glucosepane is also the dominant sugar-derived cross-link.\textsuperscript{49b} Its level is 40-200 times higher than DOGDIC, MODIC and GODIC. And it strongly correlates with age and diabetes. Taken together, these examples show that glucosepane is an excellent marker of aging and diabetic complications.

**Scheme 1. The mechanism of the glucosepane and DOGDIC formation.**

To better understand the biological consequences of sugar-derived cross-links, it is necessary to identify cross-linking sites in proteins. In previous model studies, our
group characterized the sugar-derived cross-links in RNase A *in vitro.* However, only DOGDIC at K1 and R39 was characterized as the intermolecular cross-link in RNase. RNase may be too specific, and short-lived to be applied as a general model for the sugar-derived cross-links. It may be more useful to focus on long-lived proteins such as alpha crystallins in human lens or skin collagen.

1.4. Lipoxidation and advanced lipoxidation end products (ALEs)

Like glycation, lipoxidation also correlates with age-related diseases. It is presumed that the polyunsaturated fatty acids (PUFA), the major component of cell membranes and lipoproteins, are susceptible to ROS induced oxidation. The resulting lipid hydroperoxides can further break down to form reactive carbonyl species, such as malondialdehyde (MDA), glyoxal, acrolein, HNE, 4-oxo-2-nonenal (ONE), 9-hydroxy-12-oxo-10-dodecenoic acid (HODA) and 9-keto-12-oxo-10-dodecenoic acid (KODA) ([Fig. 4](#)). The redox metal ions, ascorbic acids, or vitamin E stoichiometrically promote these fragmentations. The lipoxidation-derived carbonyl species are prone to covalent addution of protein and DNA, which in turn results in the formation of ALEs. It should be noted that some ALEs, such as CML and glyoxal-derived hydroimidazolone (G-H1), can also derive from sugar-related modifications. Both ALEs and AGEs are the markers of oxidative stress (OS), which is thought to be a major contributor to the age-related diseases.

The chemistry of ALEs has not been completely elucidated. Some lipoxidation products such as MDA, acrolein, and glyoxal, have been well studied. In the Sayre laboratory, the major interest concerns with protein modifications induced by the linoleic
acid oxidation products such as HNE and ONE.$^6$ Most of HNE and ONE modifications on cysteine and histidine residues of proteins are initiated by Michael addition. The resulting products from HNE can be further stabilized by the formation of cyclic hemiacetals,$^5$ and the ONE derived compounds can condense to form furan-ring or pyrrole-ring (the Paal Knorr condensation) with lysine (Scheme 2).$^{57}$ The major product of HNE reacting with lysine was also postulated to be the Michael addition product, which can be characterized after the sodium brohydride reduction.$^5$ Besides the Michael addition product, HNE and lysine can also form the stable pyrrole-like products. The first example of an HNE-derived ALE is 2-pentylpyrrole $^1$, which was followed by the discovery of the lysine-lysine cross-linking product, 2-hydroxy-2-pentyl-1, 2-dihydropyrrol-3-one iminium $^2$. This iminum compound derives from a mechanistic pathway which is different from the formation of 2-pentylpyrrole.$^6$ It is also thought to be a major contributor to the fluorescence near ex/em 360/430 nm, which is strongly associated with advanced atherosclerotic lesions.$^6$ Compound 2 was also found with much higher yield in the modifications by ONE of primary amine residues than from HNE. Another cross-linking compound 4 can also form from the condensation of two ONE and two lysine molecules.$^6$ However, the major long-lived product of ONE modification of lysine residues is the 4-ketoamide 3, which probably derives from the ONE-lysine Schiff base.$^6$ Recently, the Tochtrop lab also found the 4-ketoamide 3 in the incubation of HNE with lysine. The mechanism points toward direct oxidation of HNE-lysine Schiff base instead of the oxidation of HNE to ONE. Unlike methylglyoxal and glyoxal, ONE and HNE show very low reactivity toward arginine residues.$^6$
Fig. 4. The oxidation of polyunsaturated lipids.

With the mass spectrometry, the biological consequences of HNE and ONE modifications on proteins has been investigated. Both glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and Rat Hsp90 can be modified by HNE resulting in conformational changes and protein dysfunction. Grimsrud et al. found that HNE modification of Cys-117 in adipose fatty acid-binding protein (FABP) can decrease by 10-fold of the ability of the protein to bind fatty acids. β-Lactoglobulin modified by d$_9$-ONE (Scheme 3) also showed a conformational change, which resulted in abnormal His-Lys cross-linking between K141 and H146, even though the distance of these two residues is 15.67 Å in the native protein crystal structure.

Immunology is another important tool to study lipoxidation. The monoclonal antibody (mAb) to ketoamide 9K3 was generated with N$_{ε}$-(4-oxononanoyl) lysine (ONL) as immunogen, and strong immuno-reactivity to ketoamide was observed in atherosclerotic lesions. Recently amide 5 was synthesized (see Scheme 4 and
Appendix) and used to develop the antibody to KODA-derived ketoamide with the collaboration with Salomon Lab.

Scheme 2. The pathways to form HNE and ONE adducts.
Scheme 3. The synthesis of ONE-d₉.

\[
\begin{align*}
\text{C}_4\text{D}_9\text{Br} & \xrightarrow{\text{Mg, Furfural}} \text{Ac}_2\text{O, DMAP} \xrightarrow{100\%} \text{Li} \xrightarrow{75\%} \\
\text{C}_4\text{D}_9\text{O} & \xrightarrow{\text{Br}_2, \text{Na}_2\text{CO}_3} \text{OAc} & \text{OAc} & \xrightarrow{80\%} \\
\text{C}_4\text{D}_9\text{O} & \xrightarrow{\text{Ac}_2\text{O, DMAP}} \text{C}_4\text{D}_9 \xrightarrow{50\%} \\
\text{C}_4\text{D}_9\text{O} & \xrightarrow{\text{H}_2\text{O}_2, \text{TiCl}_3} \text{C}_4\text{D}_9 \xrightarrow{15\% \text{ yield of 3 steps}} \\
\end{align*}
\]

Scheme 4. The preparation of the antigen to KODA-derived ketoamide.

\[
\begin{align*}
\text{HO} & \xrightarrow{\text{H}_2\text{O}_2, \text{TiCl}_3} \text{HO} & \text{Ph} & \xrightarrow{\text{CbzCl, Et}_3\text{N, DMAP}} \text{Ph} & \xrightarrow{70\% \text{ yield}} \\
\text{HO} & \xrightarrow{\text{Br}_2, \text{BzOH, NaOAc, DMF/AcOH}} \text{Ph} & \xrightarrow{\text{Br}_2, \text{BzOH, NaOAc, DMF/AcOH}} \text{Ph} & \xrightarrow{\text{DMF/H}_2\text{O 9:1 Reflux}} \\
\text{HO} & \xrightarrow{1. \text{NHS, DCC}, 2. \text{Pd/C, H}_2} \text{5} & \xrightarrow{\text{DMF/PBS 1:4 + KLH antigen}} \text{antigen} & \xrightarrow{55\% \text{ yield}} \\
\end{align*}
\]

1.5. Epoxyketoctadecenoic acid (EKODE)

The fragmentation of the polyunsaturated fatty acid hydroperoxides is not the only route of lipoxidation. It is well known that prostanoid endoperoxides can rearrange to give levulinaldehydes without the fragmentation of the acyl chain. Recently, The
Sayre lab observed a series of compounds, named epoxyketoocadecenoic acid (EKODE), in mild linoleic acid oxidations. These compounds account for the major modifications on the histine-rich protein apomyoglobin when linoleic acid is non-enzymatically oxidized. HPLC analysis using authentic samples generated by chemical synthesis as internal standards helped us identify at least six EKODE isomers (Scheme 5). In fact, the EKODE-Ia has been reported as an *in vivo* lipoxidation product by Goodfriend et al. It was also reported that EKODE can stimulate corticosterone production, aldosterone secretion, and adrenal steroidogenesis in rat adrenal cells, and activate the antioxidant response elements.

The discovery of the ability of EKODEs to modify protein side chains imperatively leads to the question of the nature of EKODE derived ALEs. It was already mentioned above that EKODEs were the major lipoxidation products that modify apomyoglobin. However, when the substrate protein was changed to β-lactoglobulin, the modification patterns became much more complex. At least, six distinct types of modifications were observed: 1) HNE Michael adducts on H146 and H161; 2) HODA Michael adducts on H146 and H161; 3) ONE-derived-ketoamide on K60, K77, K91, K101, K135, and K138; 4) KODA-derived-ketoamide on K47, K60, K91, and K135; 5) EKODE adduct on H146; 6) ONE-derived H146-K pyrrole cross-links. It is not surprising to observe similar modification patterns between HODA and HNE since they are quasi “mirror image”. The similar phenomenon happens between the KODA and ONE, the latter displaying the most abundant modifications. The take-home message of these results is that the regioselectivity of lipoxidation-derived modifications is impacted by chemical environment of protein side chains.
1.6. Research Strategies

**Chapter 2** focuses on understanding the chemistry of sugar-derived modifications of human lens proteins. Human αAcrystallin was used as model protein, which was incubated with a mixture of $^{13}$C-labeled glucose and unlabeled glucose (1:1). The resulting glycated proteins were enzymatically digested by trypsin and characterized by MALDI and ESI-LC/MS/MS. The mass peaks of each modified peptide appeared as a
pair of two peaks separated by a certain mass difference, and were easily identified. In **Chapter 3**, a comparative analysis of carbonyl and oxidant stress in human lens and skin was carried out with quantification of all the major AGEs by LC-MS multiple reaction monitoring (LC-MS/MRM), using isotope-labeled internal standards prepared by unambiguous chemical synthesis. This helped us understand the relative importance of one AGE to the others. In **Chapter 4**, model compounds of EKODEs were synthesized. Their reactions with amino acid surrogates were characterized by NMR spectrometry and spectrophotometry, revealing the mechanisms underlying modification of protein side chains by EKODEs. These studies are expected to contribute to knowledge of protein carbonylation and oxidation in the age-related diseases.

1.7. References


36. Biemel, K. M.; Reihl, O.; Conrad, J.; Lederer, M. O., Formation pathways for lysine-arginine cross-links derived from hexoses and pentoses by Maillard processes:


49. (a) Biemel, K. M.; Friedl, D. A.; Lederer, M. O., Identification and quantification of major maillard cross-links in human serum albumin and lens protein. Evidence for glucosepane as the dominant compound. *J Biol Chem* **2002**, *277* (28), 24907-15; (b) Sell, D. R.; Biemel, K. M.; Reihl, O.; Lederer, M. O.; Strauch, C. M.; Monnier, V. M.,


56. Schauenstein, E.; Esterbauer, H., Formation and properties of reactive aldehydes. 


70. Salomon, R. G.; Kaur, K.; Podrez, E.; Hoff, H. F.; Krushinsky, A. V.; Sayre, L. M., HNE-derived 2-pentylpyrroles are generated during oxidation of LDL, are more


Chapter 2

Identification of sugar-derived modification and cross-linking sites in human alpha A crystallin
2.1. Introduction

Cataract, one of the major causes of impaired vision and blindness in the world, shows a strong correlation with diabetes and aging.\(^1\) The incidence of lens opacity is elevated with increasing levels of glycated hemoglobin,\(^2\) and 80% of the human population over age 70 may develop cataracts. Non-hereditary cataracts of the senile type are strongly associated with the acceleration of lens protein modifications. In these modifications, two kinds of reactions are sugar-related, glycation and glycoxidation, which result in the formation of advanced glycation end (AGE) products. In the past 30 years, the modifications and the biological consequences of AGEs on lens proteins, i.e. crystallins, have been investigated in depth.\(^3\) Even though there are some controversies on the role of certain AGEs such as methylglyoxal hydroimidazolone (MG-H1) on the function of alpha crystallins,\(^4\) most AGEs can alter the net charges on the protein surface, change the interaction strength among lens proteins, reduce the stability of lens proteins, and trigger protein aggregation thought to promote the cataract formation.\(^3a\)

To understand the correlations between AGE modifications and cataract formation, qualitative and quantitative mapping of the AGEs modification sites in human lens, and the mechanistic studies of crystallin cross-linking are needed. With the development of mass spectrometric technology, the identities of most mono-modifications of AGEs in lens proteins have been ascertained. However, information on the role of AGEs in crystallin cross-linking is still limited. There are several limitations concerning the acquisition of information on lens protein cross-links. First, the chemistry of sugar-derived cross-links is much more complex than that of mono modifications, and authentic samples are difficult to obtain. Second, most popular mass spectrometric search
engines such as Mascot or Bioworks cannot search for cross-links, and thus manual analysis of every cross-linking peptide is a massive and challenging task. Third, the moiety of the sugar part of the cross-links has its own unique fragmentation pattern, which increases the complexity of the tandem mass spectrum.

Ten years ago, Biemel et al. characterized a series of sugar-derived cross-links.\textsuperscript{5} Among these, glucosepane, the most abundant cross-link known, was shown to be strongly correlated with diabetic complications.\textsuperscript{6} In our laboratory glucosepane in biological specimens is quantitated by mass spectrometry following exhaustive proteolytic digestion.\textsuperscript{7} This method, however, is not suitable for analysis of modification sites in the protein. Meanwhile, Aebersold et al. have developed a sophisticated search engine xQuest for protein cross-link analysis.\textsuperscript{8} The application of xQuest is especially powerful when the cross-links are isotopically-labeled.

Alpha-A crystallin is the most abundant chaperone protein in human lens. It accounts for more than 30% lens protein by weight. The modifications of $\alpha$A crystallin decrease its chaperone function, accelerate lens protein destabilization and are associated with cataract formation.\textsuperscript{4b} The lack of a complete crystal structure of alpha-crystallins, due to their polydispersity in solution, makes the characterization of the cross-links more challenging. In this work, we undertook the task of characterizing the sites of sugar-derived modifications and cross-links of $\alpha$A crystallin modified by $^{13}$C labeled glucose.

2.2. Results

The cloning and expression of recombinant αA crystallin and several of its mutants were performed in *E. coli* BL21-(DE3)pLysS strains following the procedures described by Nagaraj et al. The resulting mutant proteins were soluble in Tris-HCl buffer (pH = 8.0) when the bacterial cell pellets were sonicated. This made it possible to extract and purify recombinant αA crystallin proteins without denaturating agents, and the resulting protein structures are close to αA wild type crystallin structure. The purity and sequence of αA crystallin and its mutants were confirmed by SDS-PAGE (Fig. 1A and B) and mass spectroscopy. To avoid denaturing of the crystallin, all the crystallin proteins were lyophilized to powder and stored at -80 °C.

Analysis of protein composition by SDS-PAGE showed that recombinant αA crystallin (40 μg/well) eluted from a nickel column migrated as a major monomeric band at ~20kDa (Fig. 1 Panel A). Minor bands were also present with molecular weight about 40 kDa, suggesting the presence dimers of crystallin. These dimer bands are commonly observed in recombinant proteins and tend to be resistant to reducing (DTT, 2-mecaptoethanol) and denaturing agents (8 M urea) and detergents (Fig. 1). Upon incubation for 4 weeks in Chelex-100 treated phosphate buffer with and without 500 mM glucose added, protein dimerization was intensified (Panel C), whereby treatment with reducing agent such as β-mercaptoethanol(lane 2) or dithiothreitol under two different conditions (lane 3 and 4) had only minor effect on the presence of protein multimers. This suggests that the dimers result either from non-disulfide covalent crosslinks such as glycation-derived crosslinks or oxidation related crosslinks such as dityrosine, or non-cleavable disulfide crosslinks.
Figure 1. SDS-PAGE of crystallin proteins separated using a 4-12% gel. Panel A: lanes 3-8, the fractions of column eluent when the recombinant αA crystallin was extracted and purified by nickel column.\textsuperscript{9} Fractions 6 and 7 were used for glucose incubation. Panel B: mutants of αA crystallin: lanes 1. K163R, 2. K163A, 3. R166A, 4. K163A/R166/A; Panel C lane 1. αA crystallin was incubated in 500 mM glucose for 2 weeks; lane 2. αA crystallin was incubated in 500 mM glucose for 2 weeks, then reduced by 2-mecapto ethanol at 37 °C for 2 days; lane 3. αA crystallin was incubated in 500 mM glucose for 2 weeks, then reduced by DTT (200mM) at 37 °C for 2 days; lane 4. αA crystallin was incubated in sodium phosphate buffer for 2 weeks, then reduced by DTT (200mM) at 37 °C for 2 days; lane 5. Freshly prepared WT αA crystallin.

2.2. Enzymatic digestion of αA crystallin and its mutants.
For purification purpose, the recombinant αA crystallin variants we used were all his-tag labeled with a six histidine (HHHHHH) peptide connected through the enterokinase site (DDDDK-M) at the N terminal of the proteins. Attempts to remove his-tag by enterokinase were unsuccessful. Short term treatment of the protein with enterokinase only partially removed the his-tag while longer term incubation resulted in degradation of the protein. Since Nagaraj confirmed that the presence of the his-tag has very little effect on structure, molar mass and chaperone function of wild type αA crystallin, all experiments were carried out with his-tag modified proteins.

For the purpose of analysis by mass spectrometry all samples were reduced with DTT, alkylated with iodoacetamide in urea, and digested with trypsin. To achieve complete digestion, fresh trypsin was added at least twice. Digestion efficiency was tested based on the recovery of the native protein sequences in the mass spectrum. In the matrix-assisted laser desorption ionization (MALDI) mass spectrum, sequence recovery of native wild type αA crystallin was 89%. In the electrospray ionization liquid chromatography-mass spectrometry (ESI-LC-MS) spectrum, sequence recovery reached 94%. This difference was attributed to the absence of the peak with m/z of 1032.48 (^{164}\text{EEKPTSAPSS}^{173}) in the MALDI spectrum while this peptide had signals at 1032.48 (1+), 516.75 (2+), 371.1 (3+) with the unambiguous MS^2 spectrum in the ESI-LC-MS.

The peptide EEKPTSAPSS had one net negative charge, which suggests that this peptide was difficult to protonate. In MALDI, all the tryptic peptides were ionized simultaneously, possibly resulting in suppressed ionization of the peptide EEKPTSAPSS by other peptides. In the LC-MS, this peptide could be easily separated from other peptides by high performance liquid chromatography (HPLC). In addition, the
ESI-MS instrument has a higher sensitivity than the MALDI-TOF instrument. Combining with the modified peptides, the sequence recovery reached 100%. In the LC-MS analysis, most unmodified peptides appeared as zero missed cleavage, while the major modified peptides had 0 or 1 missed cleavage events. This further confirmed that the digestion efficiency was excellent.

2.2.3. AGE modifications of wild type αA crystallin by glucose.

Considering the potential competing relationship between glycation and glycoxidation, glycoxidation was minimized by carrying out all incubations in chelex pre-treated 100 mM sodium phosphate buffer in the presence of diethylene triamine pentaacetic acid (DTPA). This procedure helped eliminate redox active transition metal ions from the incubation buffers.

The strategy of labeling the modifiers with isotope atoms has been widely used in proteomic studies. For investigation of protein glycation, Dai et al. successfully used \(^{13}\)C\(_6\)-labeled glucose mixed with unlabeled glucose in 1:1 ratio whereby the reactivities of labeled and unlabeled glucose were expected to be identical.\(^{11}\) As expected, mass peaks of each modified peptide appeared as a pair of two peaks separated by a certain mass difference (Table 1). These two peaks have the same intensity and retention time in the mass spectrum and similar MS\(^2\) patterns (Fig. 2).

When the molecular weight of the modified peptide was above 2000, the isotope effect became stronger, and the peptide also became multiply charged. The high resolution of the LC-MS system made separation of each isotopic peak possible, even with only 1 Da difference, and the charge number of each peptide was easily calculated.
In Fig. 2 D, each neighboring isotopic peak only had a 0.33 m/z difference, which means the peptide charge number was 3 (1/0.33 = 3).

In this study, assignments for fructose-lysine (FL), glucosepane, 3-deoxyglucosone imidazole dimer crosslink (DOGDIC), 3-deoxyglucosone hydroimidazolone (3-DG-H), MG-H1 and carboxyethyl-lysine (CEL) modifications of αA-crystallins incubated with glucose were made with high confidence (Table 2-4).

**Table 1 The AGE modifications being investigated.**

<table>
<thead>
<tr>
<th>Modification</th>
<th>Increased M.W.</th>
<th>Modified AA</th>
<th>Mass difference of double peak</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fructosyl-lysine</td>
<td>162.053</td>
<td>K</td>
<td>6</td>
</tr>
<tr>
<td>3-DG-H</td>
<td>144.042</td>
<td>R</td>
<td>6</td>
</tr>
<tr>
<td>CEL</td>
<td>72.021</td>
<td>K</td>
<td>3</td>
</tr>
<tr>
<td>MG-H1</td>
<td>54.011</td>
<td>R</td>
<td>3</td>
</tr>
<tr>
<td>CML</td>
<td>58.005</td>
<td>K</td>
<td>2</td>
</tr>
<tr>
<td>G-H1</td>
<td>39.995</td>
<td>R</td>
<td>2</td>
</tr>
<tr>
<td>GOLD</td>
<td>34.985</td>
<td>K,K(crosslink)</td>
<td>2</td>
</tr>
<tr>
<td>MOLD</td>
<td>49.001</td>
<td>K,K(crosslink)</td>
<td>3</td>
</tr>
<tr>
<td>DOLD</td>
<td>139.039</td>
<td>K,K(crosslink)</td>
<td>6</td>
</tr>
<tr>
<td>pyrraline</td>
<td>108.021</td>
<td>K</td>
<td>6</td>
</tr>
<tr>
<td>argpyrimidine</td>
<td>80.026</td>
<td>R</td>
<td>5</td>
</tr>
<tr>
<td>THP</td>
<td>144.042</td>
<td>R</td>
<td>6</td>
</tr>
<tr>
<td>GOLA</td>
<td>39.994</td>
<td>K,K(crosslink)</td>
<td>2</td>
</tr>
<tr>
<td>pentosidine</td>
<td>58.99</td>
<td>K,R(crosslink)</td>
<td>5</td>
</tr>
<tr>
<td>DODIC</td>
<td>126.032</td>
<td>K,R(crosslink)</td>
<td>6</td>
</tr>
<tr>
<td>Glucosepane</td>
<td>108.021</td>
<td>K,R(crosslink)</td>
<td>6</td>
</tr>
<tr>
<td>ox-Glucoseapne</td>
<td>124.016</td>
<td>K,R(crosslink)</td>
<td>6</td>
</tr>
</tbody>
</table>
Figure 2. Identification of the glycated α-crystallin peptide VQDDFVEIHGK*HNER in LC-MS. A. total ion chromatograms (TIC); B. the selected ion monitoring (SIM) of m/z = 662.32; C. SIM of m/z = 664.32; D. MS1 of the glycated peptide VQDDFVEIHGK*HNER with triple charges. E. MS² of the light peaks of the glycated peptide, parent m/z = 662.65 (triple charged); F. MS² of the heavy peaks of the glycated peptide, parent m/z = 664.66 (triple charged), the # represented the C13 labeling; G. the expansion of E.

2.3.1. Fructosyl-lysine (FL) in the peptides.

Not surprisingly, the most prevalent protein modification is the Amadori product, fructosyl-lysine, i.e. the product from the first step of the Maillard reaction (Table 2). All lysine-glycated peptides were easily identified by their own signature in the tandem mass spectrum using the method of Zhang et al.¹² and Gadgil et al.¹³. In the MS² spectrum, several dominant peaks were observed, such as M-36 (minus 2H₂O), M-54 (minus 3H₂O to form the pyrylium ion), M-84 (minus 3H₂O and 1formaldehyde to form the furylium ion) and M-162 (minus the sugar moiety). Even though peptide bond fragmentation was strongly inhibited by the sugar moiety fragmentations in the MS² spectrum, the remaining
peaks still provide enough information on the sequence and modification sites in the peptide (Fig. 2 G).

As shown in Table 2, all the lysine residues in αA crystallin can be modified by glucose. Considering the high concentration of glucose (500 mM) and sodium phosphate buffer (100 mM) together with the relatively long incubation period (4 weeks), this result was not surprising. The interesting finding was that K166 was not only modified by fructosyl lysine, but also by FL-H2O (Fig. 3). In the studies on glycated ribonuclease A (RNase), Monnier et al. found this FL-H2O modification on the hot spot of a cross-link involving K41.11 Biemel et al. postulated that this modification represents N6-(2,3-dihydroxy-5,6-dioxohexyl)-L-lysinate, the precursor of glucosepane.14

Table 2. The mono adducts on the lysine residues of the glycated wild type αA crystalline.

<table>
<thead>
<tr>
<th>Modification sites</th>
<th>M.W.</th>
<th>Modification type</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDVTIQHPWFKR</td>
<td>K11</td>
<td>1718.88 FL</td>
</tr>
<tr>
<td>VQDDFVEIHGKHNER</td>
<td>K99</td>
<td>1983.94 FL</td>
</tr>
<tr>
<td>HFSPEDLTVK</td>
<td>K88</td>
<td>2601.28 FL</td>
</tr>
<tr>
<td>DKFVIFDVK</td>
<td>K70</td>
<td>1384.76 FL</td>
</tr>
<tr>
<td>LPSNVDQSALSCLSADGMLTFCGPKIQTGLDATHAER</td>
<td>K145</td>
<td>4209 FL</td>
</tr>
<tr>
<td>FVIFLDVKHFSPEDLTVK</td>
<td>K78</td>
<td>2295.2 FL</td>
</tr>
<tr>
<td>AIPVSREEKPTSAPSS</td>
<td>K166</td>
<td>1816.9 FL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1726.88 CEL</td>
</tr>
<tr>
<td>EEEKPTSAPSS</td>
<td>K166</td>
<td>1193.54 FL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1191.54 ox-FL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1175.54 FL-H2O</td>
</tr>
</tbody>
</table>
Figure 3. Identification of the peptide modified by FL-H2O. A. SIM of m/z = 588.77 (upper) and m/z = 597.77 (lower), the peak with the retention time 12.77 min represents FL-H2O on $^{164}_{164}$EEK*PTSAPSS$^{173}$, the peak with the retention time 10.65 min represents FL on $^{164}_{164}$EEK*PTSAPSS$^{173}$; B. MS$^1$ at 12.77-12.85 min; C. MS$^1$ at 10.65-10.87 min; D. the MS$^2$ of peak 588.77 in B.
Besides the signals with $\Delta m/z +144$, the EEEKPTSAPSS peptide also was the only peptide found with molecular weight increase of $m/z = 160$ (Fig. 4). The peak intensities of EEEKPTSAPSS +160 are much stronger than those of EEEKPTSAPSS + FL.

**Figure 4.** Identification of the EEEKPTSAPSS +160 peptide. A. SIM of $m/z = 596.77$, the peak at the retention time 14.65 min represents EEEKPTSAPSS + 160 with the intensity of 1.21E5; B. SIM of $m/z = 597.77$, the peak at the retention time 10.65 min represents EEEKPTSAPSS + 162 with the intensity of 4.24E4; C. The tandem mass of peak 596.72, which was the double-charged peak of EEEKPTSAPSS + 160.
2.2.3.2 Carboxyethyl lysine (CEL) in the peptides.

Fructosyl lysine was not the only modification found on K166. CEL also was found at this position, and no other lysine residue was found to be modified by CEL (Fig.5). CEL is usually attributed to methylglyoxal (MGO). The modifier introduced in the incubation was glucose, and the major modification of glucose on the protein was the Amadori product. Both glucose and the Amadori product can degrade to MGO under oxidative conditions, especially in presence of catalytic trace amounts of transition metal ions in sodium phosphate buffer. All incubation buffers were pre-treated with Chelex-100 resin, and DTPA was added to the solution in order to minimize metal catalyzed oxidative degradation of glucose and fructosyl-lysine. However, trace amounts of MGO were apparently still produced independently from the transition metal ions explaining why CEL was detected on K166. This fact supports the hypothesis that K166 is the most reactive lysine residue in αA crystallin.

Figure 5. Tandem mass spectrum of CEL modified peptide AIPVSREEK*PTSAPSS. The parent mass was 864.94, representing double charge of normal CEL modification.
2.2.3.3. Glycation of arginine residues: and formation of 3-deoxyglucosone derived hydroimidazolones (3-DG-H).

Besides fructosyl-lysine, there were several other glucose-derived modifications which appeared as the double peaks with Δm/z = +6 Da in the mass spectrum. The second most abundant double peaks represented the 3-DG-H peptides. 3-DG-H originates from the modification of arginine by 3-deoxyglucosone, and results in an increase in m/z of 144 Da. One challenge in the identification of the 3-DG-H-modified peptides was to distinguish them from the MGO-derived tetrahydropyrimidine (THP). Two MGO molecules are needed to form one THP molecule. So the THP peptides also could have a molecular weight increase of 144 Da and appear as a double peak pair with Δm/z = 6 Da. However, MGO levels are considerably suppressed due to chelation of transition metal ions in the incubation buffer by DTPA while considerable 3-DG is known to be present in glucose incubation mixtures. Therefore it seems more reasonable to assign the double peaks with a molecular weight increase of 144 Da to 3-DG-H peptides. These assignments were confirmed by tandem mass spectrum. Brock et al. elucidated the MS² and MS³ spectra of THP peptides. For the THP peptide, the dominant peaks in MS² are THP - H₂O (-18 Da), THP - [H₂O + CO₂] (-62 Da) when the collision energy is 40 eV. With increasing of collision energy, the dominant peak becomes the THP – [2 H₂O + CO₂] (-80 Da). In our study, the collision energy for sequencing is 35 eV, and the dominant tandem mass peaks of Δm/z = 144 Da were the peptide bond fragmentation peaks and the M – [2 H₂O] (-36 Da) (Fig. 6). No obvious M - H₂O (-18 Da), M - [H₂O + CO₂] (-62 Da), or M – [2 H₂O + CO₂] (-80 Da) peaks were observed, confirming the validity of the assignment. All the candidate 3-DG-H peptides are listed in Table 3.
Table 3. Dicarbonyl compound derived modifications, 3-DG-H and MG-H1.

<table>
<thead>
<tr>
<th>Modification site</th>
<th>M.W.</th>
<th>Modification</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTLGPFYPSR</td>
<td>R12</td>
<td>1246.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1336.7</td>
</tr>
<tr>
<td>TVLDGISEVRSDR</td>
<td>R65</td>
<td>1676.84</td>
</tr>
<tr>
<td>HNERQDDHYISR</td>
<td>R103</td>
<td>1679.76</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1769.78</td>
</tr>
<tr>
<td>QDDHGYISREFHR</td>
<td>R112</td>
<td>1201.88</td>
</tr>
<tr>
<td>QTGLDATHAERAIPVSR</td>
<td>R157</td>
<td>1988.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2078.08</td>
</tr>
<tr>
<td>AIPVSRREEKPTSAASS</td>
<td>R163</td>
<td>1708.88</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1798.92</td>
</tr>
</tbody>
</table>

Further analysis of αA crystalline incubated with glucose showed that out of a total of 13 arginine residues only six, i.e. R12, R65, R 103, R112, R157 and R163 residues were glycated in contrast to the fact that all the lysine residues could be glycated. Interestingly, except for R12, which is close to the N terminus of the protein and possibly sticks out from the protein core, all other glycated arginine residues have at least one neighboring glutamic acid. On the other hand, the unmodified R21, R49, R54, R68, R116, R117 and R119 residues are far away from glutamic acid residues. These data suggest that there is a catalytic effect of glutamate residues on the formation of 3-DG-H. Interestingly in spite of aspartic acid having a similar structure to glutamic acid, it does not show a catalytic effect to promote crystallin glycation. R68, lying between D67 and D69, was free from the modification of 3-DG-H.
Figure 6. Tandem mass spectrum of 3-DG-H modified peptide TVLDGISEVR*SDR.

The parent mass was 839.42, representing double charge of normal 3-DGH modification. The peak of 821.56 represented the parent mass losing 2 H$_2$O.

2.2.3.4. Methylglyoxal hydroimidazolone (MG-H1).

Even though transition metal ions were removed from the incubation buffers by chelex resin, the production of methylglyoxal (MGO) apparently was not completely inhibited. The existence of CEL peptide suggests the production of methylglyoxal during the incubation. In fact, the major product of MGO modification was MG-H1. MGO is very reactive, and its selectivity for arginine residues was observed only at very low concentration of MGO (below 50 µM) and very short incubation (less than 2 days).\textsuperscript{4a} In this study, the relatively longer incubation period (4 weeks) with high concentration of glucose resulted in that only 4 arginine residues (R 12, R103, R157 and R163) could be modified by MGO to form MG-H1 (Table 3), indicating the strong inhibition of the production of MGO.
2.2.3.5 The glucose-derived cross-links.

To search for putative cross-linking peptides, all the double peak pairs with Δm = 6 Da were manually identified and entered into the web-based search engine, xQuest (http://prottools.ethz.ch/orinners/public/htdocs/xquest/), to analyse whether the modifications were mono-modifications, intra-peptides cross-links, inter-peptides cross-links or artifacts. All the cross-linking peptides confirmed are listed in Table 4.

Peptide $^{158}$AIPVSR*EEKPTSAPSS$^{173}$ was the only peptide identified as intra-peptide cross-link. The major challenge with identification of the intra-peptide cross-links is to distinguish them from the mono-adducts. The peaks of m/z = + 126 can be explained as 3-DG-H - H₂O, and the peaks of m/z = + 108 can be attributed to pyralline instead of glucosepane. Fig.8 shows that Both the MS² spectra of $^{158}$AIPVSR*EEKPTSAPSS$^{173}$ + 126 and + 108 lacked y₈, y₉, y₈₀, b₆, b₇, or b₈ signals, which reflects the fragmentation peptide bonds between R163 and K166. Whereas, AIPVSR*EEKPTSAPSS’s MG-H1 (Fig. 7) and 3-DG-H modifications showed y₈, y₉ and
b₈ peaks clearly. And the corresponding CEL (Fig. 5) and FL modifications showed y₈ and b₈ unambiguously. These facts supported the assignment of 1780.9 as the monoisotopic modification of ²⁵⁸AIPVSREEKPTSAPSS³⁶³ by DOGDIC, and the peptide 1762.9 as glucosepane.

Table 4. The cross-linking peptides.

<table>
<thead>
<tr>
<th>Modification site</th>
<th>Modification</th>
<th>Cross-link type</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIPVSREEKPTSAPSS R163, K166</td>
<td>1780.9 DOGDIC</td>
<td>intra</td>
</tr>
<tr>
<td>EEKPTSAPSS-HNERQDDHGYISR K166, R103</td>
<td>2783.25 DOGDIC</td>
<td>inter</td>
</tr>
<tr>
<td>EEKPTSAPSS-RYR K166, R117</td>
<td>1632.78 GLUCOSEPANE</td>
<td>inter</td>
</tr>
</tbody>
</table>

Compared with intra-peptide cross-links, the inter-molecular tryptic-peptide cross-links are more interesting. They may come from two different αA crystallin molecules. Three inter-molecular tryptic-peptide cross-links were identified (Fig. 9). The K166 can be cross-linked to R103 by glucosepane or DOGDIC, and to R117 by glucosepane only. As mentioned above, R103 and R163 are neighbored by a glutamic acid residue, and R117 is not. It seems that glutamic acid near to arginine can catalyze DOGDIC formation, but has less impact on glucosepane formation.
Figure 8. The intra cross-links of glucosepane and DODIC. A. SIM of 3-DGH on AIPVSREEKPTSAPSS, the retention time was 15.9 min; B. SIM of DOGDIC, the retention time was 16.9 min; C. SIM of glucosepane, the retention time was 19.9 min; D the tandem mass of DOGDIC B; E. the tandem mass of glucosepane C.
Figure 9. Modifications of AIPVSREEKPTSAPSS by 3-DG-H and intra-molecular cross-links glucosepane and DOGDIC. A. SIM of 3-DG-H on AIPVSREEKPTSAPSS, the retention time was 15.9 min; B. SIM of DOGDIC, the retention time was 16.9 min; C. SIM of glucosepane, the retention time was 19.9 min; D the tandem mass of DOGDIC B; E. the tandem mass spectrum of glucosepane C.
It is interesting that K166 was the only lysine residue involved in cross-link formation. K166 was also the only lysine residue that was found to be modified by CEL. These observations support the hypothesis that K166 is the most reactive lysine residue in αA crystallin, and is the hot spot for glycation. To confirm this, two mutants of αA crystallin were made, K166/A, and K166/R. Furthermore, R163, close to K166, was assumed to catalyze glucosepane formation, according to Biemel et al.5,11 To probe this catalytic effect, R163 was also mutated to alanine. The mutants of K166/A and R163/A were also prepared. All mutant sequences were confirmed by DNA sequencing of their plasmids. The purities of the resulting mutant proteins were checked by SDS-PAGE and LC-MS.

For quantitative analysis and verification of modifications, the four mutants of αA crystallin and the wild type protein were incubated with normal glucose using the same conditions as those used for the mass spectrometric study. After the incubation, the wild type crystallin and its mutants were exhaustively digested with proteolytic enzymes, and analyzed for AGE quantitation by LC-MS with multiple reaction monitoring (MRM) as described.7 All data were reported as mmol AGEs/mol amino acids (lysine, arginine, or methionine) because wild type crystallin had different number of lysine or arginine residues compared to the mutants. The resulting AGE levels are shown in Fig. 10.

The most abundant product is fructosyl-lysine, which reached 300 mmol/mol lysine in wild type crystallin. Other AGEs were below 35 mmol/mol amino acids. The results show that all the mutants in which K166 was removed had a 40-50 % decrease in the total amount of fructosyl lysine. On the other hand, K166 is also the only lysine residue that can be modified by ox-FL, which cannot be accurately quantified by LC-MS.
because of a lack of authentic standard. These data suggest that even though all lysine residues could be modified by fructosyl lysine, K166 probably was the major hot spot for FL. Glucosepane had a higher selectivity for K166 than FL did. Its level decreased 50-60% in the mutants of K166. This confirms the finding that K166 was the hot spot of fructose-lysine derived cross-links in the mass spectrometric study. Interestingly, R163/A also had a 19% decrease in glucosepane compared to wild type. This suggests that R163 is the favored but not exclusive arginine target for glucosepane formation.

The relatively low levels of CEL, CML, G-H1, MG-H1 confirmed the inhibition of glycoxidation via chelation of metal ions in the incubation. The mutants had no effects on the levels of G-H1 and MG-H1. And the MG-H1 level was much higher than that of GH-1, suggesting that glycoxidation generates more methylglyoxal than glyoxal. This is also supported by the above mass data showing that MG-H1 peptides formed more readily than G-H1 peptides. Unlike MG-H1 and GH-1, CEL and CML levels decreased dramatically (up to 78%) in all mutants. The decreases of CEL and CML in all K166 mutants suggest that K166 is also the hot spot of wild type crystallin for these modifications. The decrease of CEL and CML in the R163 mutant suggests that R163 had a catalytic effect on CML and CEL modifications on the K166. The fact that the CML level is almost 10 fold higher than CEL was contradictory with the fact that the MG-H1 level was much higher than G-H1, and that the CEL peptide was easier to detect by peptide mapping than was the CML peptide. This can be explained by the hypothesis that after incubation, the change of buffer for exhaustive digestion introduced transition metal ions, and the fructosyl lysine and ox-FL were oxidatively degraded to CML by metal ion-catalyzed processes.  

16
Figure 10. AGEs levels in wild type crystallin and its mutants K166A (KA), K166A/R163A (KARA), K166/R(KR), R163(RA). Glycated crystallins were obtained from the incubation of wild type and mutated crystallins in 500 mM glucose solution with metals-free sodium phosphate buffer; the control crystallins came from the incubation of crystallins with only transition-metal-deleted sodium phosphate buffer.
Besides the AGEs, a change of methionine sulfoxide level was also observed. For the wild type crystallin and its R163/A mutant, an elevated oxidation level of methionine was found in glycated crystallins than in control crystallins. However this difference was eliminated when the K166 was mutated to other amino acids. It appears that glycation of K166 has a catalytic effect on the production of methionine sulfoxide.

2.3. Discussion

2.3.1 K166 is the hot spot for glycation and glycoxidation.

In this study, K 166 was identified as the hot spot for fructosyl-lysine generation, which probably accounted for more than 50% of all fructosyl-lysine in glycated αA crystallin. This result is different from the earlier studies. Abraham et al. has identified that the glycation sites of αA crystallin as K 11 and K 78.17 This difference probably is due to the fact that the glycated EEKPTSAPSS peptide had a very weak response in the mass spectrometer. Abraham et al. incubated αA crystallin with 1 M glucose for only 5 days, and transition metal ions were apparently not removed from the incubation buffer. This may have resulted in strong glycoxidation, which can degrade the Amadori product. Indeed, our data suggest glycated K 166 seems to be preferentially oxidized to CML by transition metals.18 All these factors resulted in the decrease of the glycated EEKPTSAPSS peptide to an undetectable level in the mass spectrum. Ortwerth and coworkers used ascorbic acid instead of glucose for glycation.19 They found that K166 was the major ascorbylation sites besides K 11 and K 78 in αA crystallin. Ascorbylation was much faster than glycation, and it was easier to get enough modification for detection by mass spectrometry. In a recent study wild type αA crystallin was mutated to K166/T,20
and the mutant was incubated with fructose. The resulting protein was only 27% less glycated, in contrast with our K166 mutants that had a 40-50% decrease in glycation. This suggests that glucose-derived glycation of αA crystallin is more specific than fructose-derived glycation.

We attributed the selectivity of glycation at K166 to the two glutamic acid residues (E164, E165) near K166 (Scheme 1). The epsilon amino group of the lysine residue first reacts with the aldehyde group of glucose to form a Schiff base. Then the Schiff base is deprotonated to form an enol intermediate, which is rapidly converted to the Amadori product, fructosyl lysine. Obviously, the key step was the enolization, which is promoted by a base. Either glutamic acid residue E165 or E164 played the catalytic role in glycation of K166. Venketraman et al. predicted that negatively charged residues, such as glutamic acid and aspartic acid, could play a role as a base favoring glycation in a helical peptide model.21 K166 is the only lysine residue of αA crystallin having neighboring glutamic acid in the amino acid sequence. With limited knowledge achieved in the crystal study of truncated αA crystallin22 or its NMR structure,23 no data is available to determine other lysine residues, except for K166, are spatially close to any glutamic acid residues in the native folded protein.

Thanks to our approach based on crystallin mutants, we found that K166 was not only the hot spot for fructosyl-lysine generation, but also the preferred site of CEL and CML formation. The selectivity of CEL was confirmed in the peptide map of glycated wild type crystallin. Interestingly, both CEL and CML had a higher selectivity than FL did. The modifications on K166 probably accounted for 60-78% CEL and CML, but only for 40-50% FL. CML probably came from the degradation of fructosyl-lysine by
transition metal ions introduced in the exhaustive digestion. CEL probably came from the modification by MGO on lysine, and MGO probably came from the oxidation of glucose during the incubation. The glutamic acid residues should have similar catalytic effects as for fructosyl lysine formation (Scheme 2). MGO and glyoxal are highly electrophilic, and the existence of R163 may lower the pKa of the Schiff base to help the deprotonation step. This may partially explain why the mutant of R163/A also showed a decrease of CEL and CML.

Scheme 1. Proposed mechanism to accounts for the preferential glycation of K166.

Scheme 2. Proposed mechanism for the formation of CEL and CML.

R = CH₃ or H
K166 is not the only lysine residue of αA crystallin close to negatively charged residues. K70 and K78 are close to aspartic acid. K70 and K78 also were found to be strongly modified by fructosyl lysine, but not by CML and CEL. This implies that, in spite of having similar structures, the aspartic acid and the glutamic acid residues have different electron-donating abilities. We hypothesized that because of the shorter side chain of aspartic acid, it would form the intra-molecular hydrogen bond through the six member ring, which would weaken the ability of aspartic acid to function as a base.

![Aspartic acid residues](image)

**Figure 11.** Aspartic acid residues may form a intra-molecular hydrogen bond through a six member ring in proteins, which may decrease the pKa of aspartic acid compared to glutamic acid.

### 2.3.2 Glutamic acid catalyzes the modification of dicarbonyl compounds on the arginine residues.

As described above, the specific modification of arginine residues by 3-deoxyglucosone and MGO are more difficult to explain than those occurring at lysine residues in this work. R12, R65, R103, R112, R157 and R163 were identified in 3-DG-H peptides. R12, R103, R157 and R163 were also identified in MG-H1 peptides. The arginine residues R103, R112, R157 and R163 are next to glutamic acid in the amino acid sequence of αA crystallin. R65 is located between E63 and D67 with the interruptions by V64 and S66. No data support that R12 is close to any negatively charged residues. But it
is presumed to be located at the outer layer of αA crystallin, exposed to attack by dicarbonyl species. R21, R49, R54, R116, R117 and R119 do not have any negatively charged residues in close proximity and were free from MGO and 3-deoxyglucosone modifications. Even though R68 just sits between D67 and D69, no peptide containing R68 was found in the MG-H1 or 3-DG-H peptides. Obviously, glutamic acid exerts a better catalytic effect for generating hydroimidazolone modifications than does aspartic acid in our incubations.

Nagaraj et al. found when αA crystallin was incubated with 20 μM MGO, the MG-H1 sites were R12, R68, R157 and R163.4a With an increase in MGO concentration, more arginine residues became modified by MG-H1. When the concentration was increased to 500 μM, all the arginine residues could be modified by MGO. In this study, R12, R157 and R163 were found to be modified by both MGO and 3-deoxyglucosone. In the incubation, glucose-derived MGO and 3-deoxyglucosone concentrations probably were below 20 μM. Dobler et al. found that when type IV collagen was incubated with MGO, the hotspots of MG-H1 were at the RGD and GER motifs.24 Collagen in contrast has a characteristic triple-helical structure. On the triple-helix, the arrangement of amino acids is two-dimensional with arginine residues being exposed to MGO at almost the same level. The distribution of MG-H1 is influenced by the negative charge close to the arginine residues.
2.3.3. The hotspots for glucose-derived crosslinks.

There are several factors impacting cross-linking selectivity. The first one is the accessibility of lysine residues to glucose. Since all the lysine residues in αA crystallin can be glycated, it may be more important to find the lysine residue favoring the formation of cross-link intermediates. Lederer et al. first proposed a mechanism of formation of glucosepane and DOGDIC.\textsuperscript{5, 25} It starts with the glycation of lysine residues to form the fructosyl-lysine. The Amadori product is then converted to a dicarbonyl compound, N\textsuperscript{6}-(2, 3-dihydroxy-5, 6-dioxohexyl)-L-lysinate or diketo-like compound, 3-deoxyglucosone Schiff base. These intermediates were trapped by arginine residues to form glucosepane or DOGDIC (Scheme 4). The enolization of fructosyl-lysine to the dicarbonyl intermediates are the key steps for the formation of cross-links. K166 is the lysine residue close to glutamic acid residues (E164 and E165), which should catalyze the enolization of FL on K166. In the mass spectrum, it was found that the modification of FL-H\textsubscript{2}O was exclusively on K166, and we assumed it to be N\textsuperscript{6}-(2, 3-dihydroxy-5, 6-dioxohexyl)-L-lysinate. This may explain why cross-linking selectively happened at
K166. On the other hand, three arginine residues were found in cross-linked peptides. R103 and R163, close to glutamic acid residues, were involved both in glucosepane and DOGDIC cross-links. R117 modification was only found in glucosepane peptide. Interestingly, R103 and R163 were also modified by 3-deoxyglucosone to form 3-DG-H, and R117, without any negative charge nearby, was free from 3-DG-H and MG-H1. It seems that DOGDIC formation was similarly impacted by the environment surrounding arginine residues involved in 3-DG-H and MG-H1 formation.

Scheme 4. Mechanism proposed by Lederer et al. for glucose-derived cross-links.
The second factor influencing the formation of cross-links is the distance between lysine and arginine residues. Even though, the crystal structure of native alpha crystallin was unavailable due to its polydispersity in solution, the understanding of the structure of this protein was partially achieved in the last ten years through three approaches. The first one was based on the crystal structure from other small heat shock proteins (sHSPs). The second proposition was based on the crystal structure of truncated α crystallin since crystallin’s polydispersidy was supposedly related to the flexibility of the protein N- and C- terminal domains. The third approach, based on NMR, provides information about the residue interactions. According to these studies, we know that: (1) two α crystallin monomers interact with each other to form a dimer through their alpha-crystallin domain core, with R103 and R117 probably located at the β6 + β7 loop edge of the immunoglobulin-like folded alpha-crystallin domain core (Fig. 12); (2) the C-terminus of each monomer, binds via its 159IPV161 motif to the cleft between the β4 and β8 domains of another crystallin dimer to form an octamer; (3) the C-terminal part, beginning with R163 is highly flexible.

If the octamer structure was fixed, the distance between R 163 to R103 or R117 would be more than 20 Å. There would be no way for the cross-links of K166 to R103 or R117 to form because both glucosepane and DOGDIC require a distance between arginine and lysine below 5 Å. There are two potential explanations for the cross-links between K166 and R103 or R117. The first one is that one monomer dissociates from the crystalline oligomer, and binds randomly to another monomer. The second explanation is interactions between octamers or oligomers result in cross-linking. Both hypotheses require a transient attraction of 163REEK166 to the region of R103 and R117. It was found
that the truncated αA crystallin1-162 dramatically decreased the molar mass and chaperone function, but αA crystallin1-168, which kept the REEK motif, did not appear to undergo any obvious change. An interesting observation is that several cataract-related mutations, such as N101/D, R116/C and Y118/D, are located at the edge of the β6 + β7 region. All these mutations increase the negative charge of αA crystallin and interrupt the molecule interaction. These results suggest a transient attraction between the C-terminus of αA crystallin and the edge of the β6 + β7 region. If true, it may partially explain the polydispersity and oligomerization of αA crystallin.

**Figure 12** The crystal structure of truncated bovine αA crystallin (59-163) dimer. The origin pdb file was 3L1E. The sequence of bovine αA crystallin is 93% similar to that of human αA crystallin. The C-terminal sequence was arbitrarily extended through introduction of 10 residues using coot software. Critical residues for glycation, ie. R163, E164, E165 and R166, are displayed.

### 2.3.4. The protective effect of oxidation against the glucose-derived cross-links.
In the work above we observed that glucose promoted the oxidation of methionine. Glycation of wild type crystallin increased 3-fold the levels of methionine sulfoxide compared to control crystallin. No evidence was found that glucose or fructosyl-lysine could directly oxidize methionine. One reasonable explanation for these observations is that glycation of K166 results in an increase in hydrogen peroxide production. It is known that fructosyl-lysine can be oxidized to CML by Fe$^{3+}$ ion.\textsuperscript{18} The regeneration of Fe$^{3+}$ ion by oxidation of Fe$^{2+}$ can produce H$_2$O$_2$, which can oxidize methionine to methionine sulfoxide. Fe$^{3+}$ ion in our incubation buffers was assumed to be sequestered by chelex resin and DTPA. However, it is unknown if chelation can completely block metal-catalyzed oxidation. The hypothesis that glutamic acid could promote the binding of metal ions to proteins has been postulated by several groups.\textsuperscript{29} If this were true, the inhibition of fructosyl-lysine oxidation at K166 might not be inhibited as completely as usual since two glutamate residues (E164 and E165) are close to K166 in $\alpha$A crystallin.

It should be noted that, in the foregoing proteomics studies, the peak of m/z +160 was found on K166 while CML peptide cannot be detected. Dai et al. postulated that this modification of m/z +160 is oxidized fructosyl-lysine.\textsuperscript{11} Taking all these facts together, it is reasonable to hypothesize that this modification of m/z +160 was an incomplete oxidation of fructosyl-lysine reflecting an insufficient access to catalytic metal ions (Scheme 5). And I attributed the incomplete oxidation of fructosyl-lysine at K166 to the limited metal ion binding ability of E164 and E165. I assumed that the catalytic metal ion concentration surrounding K166 was higher compared to other lysine residues because of the REEK motif. But the binding of metal ions to the REEK motif was not strong enough to achieve complete glycoxidation.
It seems reasonable to hypothesize that the existence of the REEK motif in the C-terminal extension of αA crystallin could be an evolutionary result. The role of the complicated charge system in the C-terminal extension could be to finely adjust the interaction of αA crystallin with other molecules. However the charged residues also could increase the chance of αA crystallin cross-linking to other proteins. The REEK motif, instead of RDDK, could be a highly efficient adaption that avoids potential cross-linking that is expected to be favored by RDDK. This oxidative mechanism may increase the oxidation, glycation and glycoxidation level, but it could substantially decrease the inter-molecular cross-linking risk. Compared with mono-modifications, the cross-links may have much more malicious biological consequences to αA crystallin than simple
adducts. So avoiding cross-links would appear to provide a selective evolutionary advantage.

2.4. Conclusions

Through the study of glycation of αA crystallin, a catalytic effect of negatively-charged residues, especially glutamic acid, was observed. The negatively charged residues could facilitate the enolization of glucose-derived protein adducts, which result in the increase of glycation, glycoxidation, oxidation and cross-links in the glycated αA crystallin. The negatively-charged residues on αA crystallin may play an important role in diabetic cataract formation. In addition, the study of the cross-linking sites provides further understanding of the quaternary structure of αA crystallin. The C-terminal extension was transiently attracted to the β6 + β7 region by an electrostatic effect through the REEK motif, which may contribute to the assembly of αA crystallin oligomers.

2.5 Experimental section

**General.** D-(6C13) glucose was purchased from Cambridge Isotope Laboratories (Andover, MA). Chelex 100 resin was bought from Bio-Rad (Hercules, CA). Deionized water (18.2 MΩ cm⁻¹) was used for all experiments. Sequencing grade trypsin was purchased from Promega (Madison, WI). All other reagents were obtained in the highest quality available from Sigma (St. Louis, MO), unless specifically indicated. The concentration of proteins in solution was determined by bicinchoninic acid (BCA) assay (Pierce, Rockford, IL).
Overexpressing and purification of recombinant human lens αA crystallin.

The plasmid (1 μl) containing the cDNA of αA crystallin with N-terminal His Tag and an enterokinase site in a pET23d vector (a generous gift from Dr. Nagaraj, Case Western Reserve University) was transformed into the *Escherichia coli* cell line BL21-Gold (DE3) pLysS (Stratagene). The resulting suspension was kept on ice for 30 mins, and then heated to 42 °C for 45 sec. After cooling over ice for 2 min, 50 μl of the cell suspension was plated on an agar plate containing 100 μg/mL ampicillin and 34 μg/mL chloramphenicol, and the cells were grown overnight at 37 °C. A colony of BL21-Gold (DE3) pLysS was transferred into a 2× YT LB medium with 100 μg/mL ampicillin and 34 μg/mL chloramphenicol (50 mL), and agitated with a gyroratory shaker with 275 rpm at 37 °C for 16 hours. Then 5 mL of the resulting bacterial suspension was transferred into a fresh 2× YT LB medium with 50 μg/mL ampicillin and 34 μg/mL chloramphenicol (500 mL), and agitated with a gyroratory shaker with 300 rpm at 37 °C until optical density at 600 nm reached 0.6. At this point, isopropyl β-D-1-thiogalactopyranoside (IPTG) was added into the culture solution so that the final concentration was 1 mM. Then the suspension was further shaken at 300 rpm at 37 °C for another 4 hours, and centrifuged at 5000 rpm for 15 min to harvest the cells. The resulting cell pellets were kept at –80 °C. The cell pellets from 500 mL 2 × YT LB medium culture were resuspended in 25 mL Tris-HCl solution (pH = 8.0, 50 mM) with 100 μM phenylmethanesulfonylfluoride (PMSF) and 1 mM dithiothreitol (DTT), and sonicated at 4 °C for 15 min. The resulting suspension was centrifuged at 15,000 rpm for 30 min. The supernatant was loaded onto a short column packed with Ni-NTA resin (Qiagen Valencia, CA), and washed with 10 column volume of the Tris-HCl buffer (pH = 8.0, 50 mM) containing 500 mM NaCl and...
50 mM imidazole. The pure recombinant αA crystallin was flush out from the column with 5 column volume of the Tris-HCl buffer (pH = 8.0, 50 mM) containing 500 mM NaCl and 250 mM imidazole. The resulting eluent was concentrated to 2 mL by centrifuging in 10 kDa MWCO filter (Millipore, Billerica, MA), and dialyzed against 500 mL of chelex-pretreated sodium phosphate buffer (5 mM, pH = 7.4) containing 50 μM DTPA at 4 °C for 2 days with three times buffer change. Spectra/Por 7 dialysis tubing (Spectrum laboratories, Rancho Dominguez, CA, 12-14 kDa MWCO) was used. The purities of the proteins were checked by SDS-PAGE. 4-12% Bis-tris gel (Invitrogen, Carlsbad, CA). All the crystallin solution was freezing-dried, and kept at –80 °C.

**Mutagenesis of αA Crystallin.** Mutant plasmids for the expression of mutant αA Crystallin, K166/R, K166/A, R163/A and K166/A, R163/A were constructed with a Quickchange Site-Directed Mutagenesis Kit (Stratagene Cloning Systems). The wild type αA Crystallin plasmid was used as the template. The primers of K166/R mutant were 5′-CCGAGGTGGGCTCCCTCTCCCAC-3′ and 5′-GGGGAGGGAGGCCCACCTCCGG-3′. The primers of K166/A mutant were 5′-AGCCGAGGCTGGGCTCCTCCCCAC-3′ and 5′-GGGGAGGGAGGCCCACCTCCGGCT-3′. The primers of R163/A mutant were 5′-GGGCTTCTCCTCCGCAGACGCGGATG-3′ and 5′-CATCCCCGTGTCGGCGGAGGAGGGGAGG-3′. The primers of K166/A, R163/A mutant were 5′-GATCCCCGTGTCGGCGGAGGAGGCGGCCACCTCAGGC-3′ and 5′-GGGGAGGGAGGCCCACCTCCGGCT-3′. The mutagenesis was confirmed by DNA sequencing at Biotic Solutions, Inc. The mutated plasmids were transformed into XL1-Blue supercompetent cells. The plasmids containing the mutants were transformed into the *Escherichia coli* cell line BL21-Gold (DE3)pLysS (Stratagene) which was stored at -
80 °C. The cloning, overexpression and purification of the αA crystallin mutants were the same as those of wild type αA crystallin.

**Preparation of glycated proteins.** Two batches of the glucose incubation were carried out. The first incubation consisted of wild type αA crystallin (40 mg/mL) being incubated with a 1:1 mixture of 250 mM D-glucose and 250 mM D-(13C6)glucose in 100 mM Chelex-pretreated sodium phosphate buffer containing 1 mM DTPA at 37 °C for 4 weeks. The second incubation consisted of wild type αA crystallin and its mutants being incubated with 500 mM normal D-glucose in 100 mM Chelex-pretreated sodium phosphate buffer containing 1 mM DTPA or the same solution without glucose as control at 37 °C for 4 weeks respectively. After incubation, glucose was removed by dialysis against 4 L of sodium phosphate buffer (5 mM, pH = 7.4) at 4 °C for 2 days with four time buffer change.

**Enzymatic digestion for peptide mapping.** About 1 mg of glycated αA crystallins or control crystallins was dissolved in 0.2 mL urea solution (8M, with 50 mM NH₄HCO₃, pH = 7.8). 10 μL of DTT solution (0.5 M) was added into it, and the resulting solution was incubated at 37 °C for 1 hour. Then 40 μL of iodoacetamide solution (0.5 M) was added. The solution was incubated at 37 °C in the dark for another 1 hour, and diluted with 1.8 mL NH₄HCO₃ solution (50 mM, pH = 7.8). Then 10 μg of sequencing grade trypsin in 100 μL of CaCl₂ solution (50 mM) was added into the diluted solution. After the resulting solution was incubated at 37 °C for 4 hours, additional 10 μg of fresh trypsin was added to the solution, and digestion was further carried out for another 4 hours. The digest was quenched with 5 μL of formic acid. The tryptic peptide solutions were desalted with Micro Tip C18 column (the Nest Group, Inc. Southborough, MA)
following the manufacturer protocol. The desalted peptide solutions were kept at -80 °C until being used.

Matrix-assisted laser desorption and ionization mass spectrometry (MALDI). The digested peptides were analyzed by using a prOTOF 2000 MALDI O-TOF mass spectrometer (PerkinElmer, Inc. Waltham, MA). The desalted sample solutions were mixed with an equal volume of α-cyano-4-hydroxycinnamic acid (CHCA) (Fluka, St. Louis, MO) matrix solution (20 mg/mL in 50% acetonitrile, 0.1% TFA). 1 μL of the mixture was applied onto the laser target plate and was air-dried before being introduced into the mass spectrometer. Laser-desorbed positive ions were analyzed after acceleration with 19 kV in the reflectron mode for the peptide digest. Other instrument settings were the following: laser energy at 75%; laser rate, 100.0 Hz; declustering potential, 25.0 V; cooling flow, 200 mL/min; mass range, 500.0-10000.0 Da; focus flow, 200.0 mL/min.

Liquid chromatography-tandem mass spectrometry (LC/MS/MS) analysis for peptide mapping. LC-MS/MS analyses of the proteolytic digests were performed using a quadrupole ion trap mass spectrometer (model LTQ) from Thermo-Finnigan (San Jose, CA) coupled with an Ettan MDLC system (GE Healthcare, Piscataway, NJ), chromatographed with a gradient of 0-60% acetonitrile/0.1% formic acid for 50 min. The spectra were acquired by data dependent methods, consisting of a full scan and then MS/MS on the ten most abundant precursor ions at a collision energy of 35%. The previously selected precursor ions were repeated twice during 45 s and then were excluded for 180 s.

AGEs determination by GS/MS and LC-MS/MS. 500 μg of αA crystallins (glycated or control) were dissolved in 500 μL PBS, and sequentially digested at 37 °C.
for 24-h consecutive intervals by the addition of the following enzymes: 18 μg of proteinase K, 18 μg of peptidase, 18 μg of Pronase (Roche Applied Science), and 1.8 μg of aminopeptidase M (Roche Applied Science). Chloroform and toluene (1.5 μl) were added to every tube as antimicrobial agents. Digestion efficiency was determined by the ninhydrin reaction and expressed as the percentage of total leucine equivalents assayed in the hydrolysate. Digestibility varied from 61 to 97%. Enzymatically digested samples were measured by electron spray positive ionization-mass spectrometric multiple reaction monitoring (ESI-MRM) by using LC-MS/MS system composed of a 2690 Separation module with a Quattro Ultima triple quadrupole mass spectrometry detector (Water-Micromass, Manchester, UK) following the previously published procedure by Thornalley et al. Analytes released by self-digestion of proteases in assay blanks were subtracted from analytic estimates.

2.6. References


Chapter 3

Anaerobic vs aerobic pathways of carbonyl and oxidant stress in human lens and skin during aging and in diabetes: a comparative analysis
3.1. Introduction

A central challenge in studying diabetic complication is related to the difficulty of relating a specific type of damage by carbonyl stress or oxidative stress, to the underlying metabolic pathway responsible for the damage. One modification product can originate from multiple metabolic pathways. For example, it has been reported that large amounts of methylglyoxal-derived hydroimidazolone (MG-H1) are present in the diabetic lens;\textsuperscript{1} and methyglyoxal (MGO), the major metabolite accounting for MGH1, can originate from glycolysis, oxidation of threonine, ketone bodies,\textsuperscript{2} or ascorbic acid.\textsuperscript{3} Additionally, carboxymethyllysine (CML) can originate from metal or peroxynitrite-catalyzed oxidation of fructose-lysine,\textsuperscript{4} ascorbic acid, lipid peroxidation,\textsuperscript{5} and any pathway involving glyoxal/glycolaldehyde formation.\textsuperscript{6} Pentosidine, the first fluorescent sugar-derived cross-linking product,\textsuperscript{7} easily forms from pentoses, ascorbic acid, or oxidation of fructose-lysine.\textsuperscript{8} The rationale underlying these selective formations of AGEs is that the local chemical environment in the biological system dictates the pathway of deleterious metabolite formation.

In this study, two proteins, lens crystallins and skin collagen, were selected for comparative analysis of the impact of various chemical environments on levels of advanced glycation and oxidation products. Both of these proteins are long-lived: the half-life of skin collagen is about 15 years,\textsuperscript{9} while that of lens crystallins is almost lifetime.\textsuperscript{10} This allows us to observe accumulation of glycation or oxidative damage without considering the effect of protein turnover rates. On the other hand, these two proteins are surrounded by quite different chemical environments in the biological system. First, lens crystallins are surrounded by an avascular cellular system with relative low oxygen tension (8-20 mm Hg O\textsubscript{2}),\textsuperscript{11} whereas the skin is highly vascularized, with
collagen being exposed to 30–40 mm Hg O₂ tension. Second, the lens is rich in ascorbic acid and glutathione, two antioxidants that are present in millimolar concentrations, but in micromolar concentrations in plasma and extracellular fluids. Finally, the skin samples used in this study are from non-sun-exposed skin, and the lens crystallins are exposed to UV/VIS irradiation, which is a potential trigger of oxidative stress and carbonyl stress. It should also be noted that lens crystallins are water-soluble proteins whereas skin collagen is water-insoluble under physiological conditions. The latter is well documented with its rigid triple-helical structure, and is highly-enriched glycine and hydroxy proline residues.

3.2. Results

3.2.1. The preparation of isotope-labeled AGE standards.

[^13C₆]-Fructosyl lysine ([^13C₆]-FL) and[^13C₆]-glucosepane were prepared from[^13C₆]-glucose as described.[^14] d₄-Methionine sulfoxide (d₄-MetSOX) and d₄-6-hydroxy norleucine were synthesized from deuterium-labeled amino acids.[^15] d₂[^13C₁]-Carboxyethyl lysine (d₂[^13C₁]-CEL), d₂-glyoxal hydroimidazolone (d₂-G-H1), d₃-methylglyoxal hydroimidazolone (d₃-MG-H1), and d₃-2-aminoadipic acid, were prepared by a new method, which introduced the isotope-labeling by hydrogen-deuterium (H-D) exchange (Scheme 1). Specifically, the commercially available unlabeled 2-aminoadipic acid was converted to the ester 1, and the resulting compound was incubated with sodium methoxide solution in methanol-OD, which would produce H-D exchange at the C2 and C5. After removing the protecting groups and purification, 2,5,5[^2H₃]-2-aminoadipic acid was obtained with 77% isotopic purity (the major isotopic impurity was d₂-2-
aminoadipic acid). The synthesis of isotope-labeled CEL started with $^{13}$C$_1$-pyruvic acid, which only produced 1 Da increase of molecular weight. Then the resulting $^{13}$C$_1$-pyruvic acid was processed with the same method as for the $d_2$-2-aminoadipic acid preparation to give $d_2$-$[^{13}$C$_1$]-CEL. As for $d_2$-G-H1, it was made from refluxing of Boc-ornithine and 5-hydro-2-methylthio-4-imidazolone in ethanol-OD, which resulted in H-D exchange at the imidazolone ring. $d_3$-MG-H1 was prepared by the same method, but H-D exchange on 5-hydro-5-methyl-2-methylthio-4-imidazolone did not introduce enough isotope-labeling. Thus Boc-$d_2$-ornithine was synthesized. For the synthesis of G-H1 and MG-H1, ethanol was used as the reaction solvent instead of methanol, resulting in higher yield (10%) than that reported (4-6%).

Scheme 1. The synthesis routes for $d_3$-2-aminoadipic acid (A), $d_2$-$[^{13}$C$_1$]-CEL (B), $d_2$-G-H1 (C) and $d_3$-MG-H1 (D).
3.2.2. From glucose to glucosepane.

It has been well documented that glucose can adduct lysine residues to form fructosyl-lysine. The resulting Amadori product can further form other advanced glycation end products (AGEs) through several metabolic pathways or revert to lysine by fructosamine 3- phosphokinase (FN3K).\(^\text{17}\) As shown in Fig. 1, fructosyl-lysine only mildly accumulated with aging in both lens and skin samples, and strongly correlated with diabetes. As a cross-linking product originating from fructosyl-lysine, glucosepane also increased in diabetic lens and skin samples, and accumulated mildly with aging in lens. However, glucosepane showed much stronger correlation with aging in skin than in lens. It is also should be noted that both fructosyl lysine and glucosepane levels in skin were much higher than those in lens.
**Figure 1.** Age-related changes in levels of fructose-lysine (Panel A, C and E) and glucosepane (Panel B, D and F) in insoluble lens crystallins and skin collagen-rich fraction. Symbol ● and ○ indicate the levels of fructosyl-lysine (Panel A) or glucosepane (Panel B) in non-diabetic lens and skin samples respectively; symbol ▲ indicates the levels of fructosyl-lysine (Panel C and E) or glucosepane (Panel D and F) in both diabetic lens and skin samples. The bold solid lines in Panel A and C indicate the regression line for fructose-lysine in non-diabetic lens, $y = 265 + 2x$, $n = 17$, $r = 0.46$, $P =$
0.064 (NS), and those in Panel B and D indicate the regression line for glucosepane in non-diabetic lens, \( y = 44 + 4x, n = 25, r = 0.72, P < 0.0001 \). The bold dash lines in Panel A and E indicate the regression line for fructose-lysine in non-diabetic skin, \( y = 1562 + 7x, n = 12, r = 0.46, P = 0.11 \) (NS), and those in Panel B and F indicate the regression line for glucosepane in non-diabetic skin, \( y = 20 + 34x, n = 14, r = 0.82, P < 0.0001 \). All the regression lines were generated through linear fitting. And both the upper bound and lower bound of each regression line (two light solid lines) were also generated in Panel C, D, E and F, representing the range in which the AGE levels of non-diabetic samples are located with 95% confidence. The accumulation rates of fructosyl-lysine between skin and lens are statistically the same (P > 0.2, NS). NS, nonsignificant.

With the incubation of bovine type I collagen with glucose (0 mM, 5 mM, 25 mM), the kinetics of glycation of collagen was carefully investigated (Fig. 2). Both fructosyl-lysine and glucosepane levels were elevated as a function of the concentration of glucose and the incubation time. It is interesting that the formation of both fructosyl-lysine and glucosepane were not significantly impacted by the presence of transition metal ions, whereas CML, MG-H1 and GH1 were clearly elevated by the presence of transition metal ions. It should be noted that the incubation solutions were refreshed once every week to keep the glucose concentrations constant. CML showed accumulation process on collagen like MG-H1 and G-H1, rather than fructosyl-lysine, suggesting that CML was primarily derived from the glyoxal produced by glycolysis. The oxidation of fructosyl-lysine to CML had weak effects on the levels of fructosyl lysine and CML, which in turn did not significantly impact the formation of glucosepane.
Figure 2. Kinetic studies on collagen glycation. Bovine type I collagen was incubated with glucose (0 mM, 5 mM, 25 mM) in sodium phosphate solutions with/without chelation of transition metal ions for 0, 2, 4, 8, 14, or 16 weeks. Each incubation was in duplicate. Fructosyl-lysine, glucosepane, CML and aminoadipic acid (AAA) are
expressed as mmol/mol lysine; MG-H1 and G-H1 are expressed by are expressed as mmol/mol arginine. No CEL was detected in any of the collagen samples in these incubations.

3.2.3. Glycoxidation vs glyoxal-dependent mechanism of CML formation.

CML can originate from multiple pathways including glycoxidation of fructose-lysine or modifications of glyoxal/glycolaldehyde, which can be derived from lipid oxidation or serine oxidation by myeloperoxidase. For this reason, we measured both CML and the glyoxal-derived hydroimidazolone (G-H1), the latter was assumed to originate exclusively from glyoxal. Chapter 2 already showed that when αA crystallin, one of the major proteins in lens, was incubated with glucose, CML primarily came from the glycoxidation of fructosyl-lysine even under transition metal ion free conditions, whereas no obvious G-H1 was detected. Fig. 2 shows that when collagen was incubated with glucose, CML and G-H1 displayed similar formation processes. Both were increased with the increased glucose concentration and the incubation time when transition metal ions were present, and inhibited upon removal of the transition metal ions. This strongly supports the conclusion that glucose auto-oxidation played an important role in the formation of CML.  

CML shows similar age-related accumulations both in the human lens and skin (P > 0.25, NS) (Fig. 3). However, G-H1 is barely detectable in lens samples, whereas it accumulates with aging in skin. The high CML and G-H1 levels in diabetics patients over 60 years old are attributed to concomitant renal failure, which is known to dramatically increase levels of these two markers.  

There are two exceptional data points with high
G-H1 levels in the diabetic collagen samples, which are derived from the skin of younger diabetic individuals with renal failure (not shown). It is surprising to observe that there are no obvious increases in either CML or G-H1 in the type 2 diabetic samples, while our previous studies showed that CML is elevated by two to four fold in skin samples from younger type 1 diabetic patients. Pentosidine is assumed to be a glycoxidation product. Like CML and G-H1, there is no significant effect of diabetes on the pentosidine levels. Both in lens and skin, pentosidine levels are quite low (below 50 pmol/mg protein), which suggests it is not the major cause of cross-linking AGE.
**Figure 3.** Age-related changes in CML (Panel A, D and G), G-H1 (Panel B, E and H), and pentosidine (Panel C, F and I) in insoluble lens crystallins and skin collagen-rich fraction. Symbol ● and ○ indicate the AGE levels (CML in Panel A, G-H1 in Panel B and pentosidine in Panel C) in non-diabetic lens and skin respectively; symbol ▲ indicates the AGE levels in diabetic lens and skin. Regression lines were generated as the method described in Fig. 1. The bold solid lines in both Panel A and D indicate the regression line for CML in non-diabetic lens, \(y = 152 + 10x, n = 25, r = 0.54, P = 0.004\); those in both Panel B and E indicate the regression line for G-H1 in non-diabetic lens \(y = 0.6 + 0.02x, n = 24, r = 0.15, P = 0.47\) (NS); and those in both Panel C and F indicate the regression line for pentosidine in non-diabetic lens, \(y = 0.16x - 3.3, n = 22, r = 0.65, P = 0.001\). The bold dash lines in both Panel A and G indicate the regression line for CML in non-diabetic skin, \(y = 5x - 84, n = 13, r = 0.78, P = 0.001\); those in both Panel B and H indicate the regression line for G-H1 in non-diabetic skin, \(y = 6x - 137, n = 13, r = 0.86, P < 0.0001\); and those in both Panel C and I indicate the regression line for pentosidine in non-diabetic skin, \(y = 0.14 + 0.33x, n = 15, r = 0.73, P = 0.002\), and with three outliers. The light solid lines in Panel D-I indicate the ranges in which the AGE levels of non-diabetic samples are located with 95% confidence. CML: levels are significantly higher in lens vs skin (\(P < 0.0001\)), but the accumulation rates are not significantly different (\(P > 0.25\)). GH-1: the accumulation rate and levels are significantly higher in skin vs lens (\(P < 0.0001\)). Pentosidine: levels are significantly more elevated (\(P < 0.0001\)) for skin vs lens, but the accumulation rates are not significantly different (\(P > 0.05\)). The effect of diabetes (Panel D-I) is not significant (\(P > 0.05\)) for all three parameters for both tissues.
3.2.4. Origin and significance of methylglyoxal modifications in lens and skin.

Previous work by Ahmed et al. revealed very high levels of methylglyoxal-derived hydroimidazolones (MG-H1 and MG-H2) in lens.\textsuperscript{1} Comparative measurements of MG-H1 and CEL, both of which are derived from methylglyoxal (MGO), confirmed the production of very high amounts of MGO in lens compared to skin (Fig. 4). However, MG-H1 and CEL levels and accumulation rates are significantly higher in lens than in skin. These facts suggest that another major source of methylglyoxal contributes to the high level of MG-H1 in lens. Considering the existence of ascorbic acid in millimolar concentration in the lens, we hypothesized that methylglyoxal AGEs might be derived from ascorbate. We incubated bovine lens protein homogenate with 5 mM ascorbic acid for 7 days. This led to a 19-, 8-, 7-, and 6-fold increase in CML, MG-H1, G-H1, and CEL vs the control, respectively (Fig. 5). Similar results were observed for incubation with dehydroascorbic acid (DHA), supporting a role for ascorbate as a source of methylglyoxal in the lens.
Figure 4. Age-related changes in levels of MG-H1 (Panel A, C and E) and CEL (Panel B, D and F) in insoluble lens crystallins and skin collagen-rich fraction. Symbol ● and ○ indicate the AGE levels (MG-H1 in Panel A and CEL in Panel B) in non-diabetic lens and skin respectively; symbol ▲ indicates the AGE levels in diabetic lens and skin. The bold solid lines in both Panel A and C indicate the regression line for MG-H1 in non-diabetic lens, $y = 69x - 181$, $n = 25$, $r = 0.88$, $P < 0.0001$; and those in both Panel B and
D indicate the regression line for CEL in non-diabetic lens, $y = 254 + 3x$, $n = 25$, $r = 0.34$, $P = 0.099$ (NS). The bold dash lines in both Panel A and E indicate the regression line for MG-H1 in non-diabetic skin, $y = 5x - 66$, $n = 14$, $r = 0.64$, $P = 0.014$; and those in both Panel B and F indicate the regression line for CEL in non-diabetic skin, $y = 2x - 16$, $n = 14$, $r = 0.59$, $P = 0.027$. The light solid lines in Panel C-E indicate the ranges in which the AGE levels of non-diabetic samples are located with 95% confidence. The accumulation rate is significantly ($P < 0.0001$) greater for lens vs skin for MG-H1, but not for CEL ($P > 0.6$). Levels are significantly higher in lens vs skin for both parameters ($P < 0.0001$). The effect of diabetes (Panel C-E) is not significant ($P > 0.1$) for levels of MG-H1 and CEL in both tissues.

**Figure 5.** Formation of AGEs in bovine lens crystallin homogenate incubated with 5 mM ascorbic acid or DHA for 7 days in 100 mM phosphate buffer containing 1 mM DTPA chelating agent. All values are elevated in the presence of ASA and DHA compared to control.
It is surprising that diabetes did not further enhance MG-H1 and CEL levels, either in lens or in skin, except for four data points in the diabetic lens. As for CEL in skin collagen, its formation is irrelevant for glycolysis or glycoxidation. When collagen was incubated with glucose in vitro, CEL was not detected in significant amounts (Fig. 2). However, MG-H1 and CEL in αA crystallin, or MG-H1 in bovine collagen (Fig. 2) are obviously impacted by glucose concentrations.

3.2.5. Metal-catalyzed and hydrogen peroxide-mediated oxidation.

The mechanism and extent of lysyl residue oxidation in lens and collagen have been recently described by Sell et al. in detail,\textsuperscript{15b, 19} and therefore we show here only the comparison of the regression lines. Fig. 6 shows that the absolute levels of allysine were much lower in lens than in skin but increased with age (P < 0.0001), whereas they were in a steady state in skin. Allysine oxidizes to the stable 2-aminoadipic acid, whose levels are about 10 × higher in skin and the rate of increase of 2-aminoadipic acid accumulation is over 5 × faster in skin than in lens (P < 0.0001). When bovine collagen was incubated in glucose, 2-aminoadipic acid was only elevated with incubation time and glucose concentration if transition metal ions were present, as CML and G-H1 did (Fig. 3). This was consistent with the results of the previous studies, showing that diabetes impacts on the formation of 2-aminoadipic acid in collagen.\textsuperscript{15b}
Figure 6. Age-related accumulation rates of allysine, 2-aminoadipic acid, pentosidine, and ornithine in skin vs lens. The regression lines and 95% confidence intervals of prediction are shown for both skin (dashed lines) and lens (solid lines with shading). (Top) The data for skin (nonsepsis) and lens are replotted from Sell et al.\textsuperscript{15b} \textsuperscript{20} and Fan et al.\textsuperscript{19} respectively. Regression lines are allysine, skin, \( y = 71 - 0.02x \), \( P = 0.94 \) (NS), and lens, \( y = 5 + 0.7x \), \( P < 0.0001 \); 2-aminoadipic acid, skin, \( y = 78 + 3x \), \( P < 0.0001 \), and lens, \( y = 28 + 0.6x \), \( P = 0.042 \). (Bottom) Regression lines for skin pentosidine, see Fig. 3. Ornithine, skin, \( y = 1281 + 32x \), \( P < 0.0001 \), and lens, \( y = 3445 + 67x \), \( P = 0.033 \).
3.2.6. The hierarchy of protein damage during aging.

The above data illustrate how two different environments affect AGE formation and protein oxidation in aging. However, they do not compare the extent of each modification. We have therefore graphed in Fig. 7 the average levels of each AGE for subjects above the age of 65 years. It should be noticed that most of AGEs derived from arginine such as MG-H1 can spontaneously hydrolyze to ornithine.20 As shown in Fig. 6, ornithine accumulates with aging both in lens and skin. However, it is still surprising to observe that ornithine is the most prevalent modification both in lens and skin (Fig. 7), suggesting that the assault by glycation on arginine residues results in the most extensive damage to proteins. In lens, the second and third most prevalent modifications are MG-H1 and CML respectively, in contrast with fructosyl-lysine and glucosepane in skin. This difference of modification priority may be influenced by protein structure. Fig. 8 shows that human alpha A crystallin converts fructosyl-lysine to CML much more efficiently than bovine collagen. As a consequence, less fructosyl-lysine can form glucosepane.

Total damage by the AGEs is at least 17 nmol/mg crystallin and 8 nmol/mg collagen in lens and skin, respectively. Assuming a mean molecular weight of 20 and 100 kDa for each protein, this amounts to about 35 and 250 mol% modification of old crystallin and triple-helical collagen, respectively. Thus, damage by carbonyl stress in old age is far from trivial. Most other AGEs, such as MOLD, GOLD, GODIC, MODIC, DOGDIC, and argpyrimidine are present in very low concentrations.21 Quite strikingly, 2-aminoadipic acid was the fourth most prevalent form of damage in collagen, implying considerable oxidation catalyzed by metals and hydrogen peroxide during aging that was
reflected in a four-fold increase in methionine sulfoxide compared to lens. These data were from Sochaski et al.\textsuperscript{22}

\textbf{Figure 7.} Quantitative comparisons of mean levels ± SD for AGEs and lysine oxidation levels reached at age ≥ 65 years for insoluble lens crystallins and skin collagen-rich fraction. Levels shown for argpyrimidine were reproduced from noncataractous lens samples of Padayatti et al.\textsuperscript{23}
Figure 8. The glycoxidation efficiency in crystallin vs in collagen. Both human alpha A crystallin and bovine type I collagen were incubated with 500 mM glucose in chelex-treated sodium phosphate buffer for 4 weeks. The ratio of CML/fructosyl lysine is much higher in crystallin than in collagen, whereas the ratio of CEL/FL is almost same in both proteins. 5.2% fructosyl lysine formed glucosepane in collagen, and only 1.3% in crystallin.

3.3. Discussion

As reported, we observed that multiple metabolic pathways can contribute to the formation of AGEs, suggesting that impact of glycation on different protein substrates can be variable depending on the biological environment. By comparing AGE levels in the anaerobic, ascorbic acid-rich lens environment with those in the aerobic skin
environment, a better understanding of glycation pathways in biological system was obtained.

3.3.1. The glycoxidation of fructosyl lysine to CML is favored in lens.

It is surprising that the oxidative products such as CML are present in much higher levels in the anaerobic lens environment than the aerobic skin environment, whereas the non-oxidative products such as fructosyl-lysine and glucosepane are much higher in skin (Fig. 1 and 3). It is true that there are high levels of antioxidants such as glutathione and ascorbic acid in human lens, and ascorbic acid may directly attack the lysine residues through C2 on dehydroy ascorbate (DHA) or 2, 3-diketogulonic acid. However, antioxidants can not interrupt the formation of fructosyl-lysine and glucosepane. One potential explanation for the lower level of fructosyl-lysine in lens is the existence of fructosamine 3-phospho- kinase (FN3K), which can degrade fructosyl-lysine to lysine and 3-deoxyglucosone. The lower fructosyl-lysine level results in lower level of glucosepane. However, this explanation is unlikely because the FN3K deglycation byproduct, 3-deoxyglucosone modifies proteins very quickly resulting the formation of 3-deoxyglucosone-derived hydroimidazolone (3-DG-H) and imidazolium cross-link (DOGDIC). But the DOGDIC level in lens is 40-fold lower than that of glucosepane. Therefore, FN3K activity may be not as high as thought in lens.

Considering the high level of CML, we postulated another potential explanation for the low amounts of fructosyl-lysine and glucosepane in lens, i.e., the glycoxidation of fructosyl-lysine to CML is much more favored for lens proteins compared to skin collagen. As shown in Fig.8, much more fructosyl-lysine was oxidized to CML in human
alpha A crystallin than in bovine collagen, and consequently more fructosyl lysine was converted to glucosepane. We postulate that hypothetically attributed this high reactivity of oxidation of fructosyl lysine to the REEK motif in alpha crystallin, which is not available in collagen.

It should be noticed that the oxidation products such as 2-AAA and methionine sulfoxide also show more obvious correlations with diabetes in lens, even though their absolute levels are much higher in skin (Table. 1). Hydrogene peroxide, a potential byproduct of glycoxidation, can contribute to the formation of 2-AAA and methionine sulfoxide. These facts provide more support to our hypothesis that the crystallins favor glycoxidation more than collagen do, which results in the much lower fructosyl lysine and glucosepane levels in lens than in skin.

If the hypothesis that the glycoxidation rates in lens and in skins are different is true, it may imply an evolutionary effect. The cross-links of proteins may have more malicious effect on the functions of lens proteins than on those of skin collagens. So lens proteins developed a structure, the REEK motif, that can catalyze the glycoxidation and thereby inhibit sugar-derived cross-link formation.
Table 1 Effect of diabetes on the relationship between age-related accumulation of protein modifications in lens and skin as reflected by Spearman correlation coefficients ranked according to R value as a function of age and presence of diabetes.

<table>
<thead>
<tr>
<th>AGE</th>
<th>With age</th>
<th>R</th>
<th>P</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lens</td>
<td>Allysine</td>
<td>0.77</td>
<td>&lt;0.0001</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>MG-H1</td>
<td>0.76</td>
<td>&lt;0.0001</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>Glucosepane</td>
<td>0.60</td>
<td>0.001</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>MetSOX</td>
<td>0.53</td>
<td>0.003</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>CML</td>
<td>0.51</td>
<td>0.007</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Pentosidine</td>
<td>0.49</td>
<td>0.014</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>2-Aminoacidipic acid</td>
<td>0.42</td>
<td>0.036</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Ornithine</td>
<td>0.38</td>
<td>0.036</td>
<td>31</td>
</tr>
<tr>
<td>Skin</td>
<td>Pentosidine</td>
<td>0.87</td>
<td>&lt;0.0001</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>Glucosepane</td>
<td>0.84</td>
<td>&lt;0.0001</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>CML</td>
<td>0.79</td>
<td>&lt;0.0001</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>Ornithine</td>
<td>0.78</td>
<td>&lt;0.0001</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>CEL</td>
<td>0.76</td>
<td>&lt;0.0001</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>MG-H1</td>
<td>0.76</td>
<td>0.001</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>G-H1</td>
<td>0.72</td>
<td>0.004</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>MetSOX</td>
<td>0.68</td>
<td>0.008</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2-Aminoacidipic acid</td>
<td>0.61</td>
<td>&lt;0.0001</td>
<td>49</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>AGE</th>
<th>With diabetes</th>
<th>R</th>
<th>P</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lens</td>
<td>Fructose-lysine</td>
<td>0.83</td>
<td>&lt;0.0001</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Glucosepane</td>
<td>0.73</td>
<td>&lt;0.0001</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>MetSOX</td>
<td>0.46</td>
<td>0.001</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>2-Aminoacidipic acid</td>
<td>0.41</td>
<td>0.007</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>CEL</td>
<td>0.34</td>
<td>0.025</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>MG-H1</td>
<td>0.32</td>
<td>0.03</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>Ornithine</td>
<td>0.25</td>
<td>0.08 [NS]</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>G-H1</td>
<td>0.24</td>
<td>0.12 [NS]</td>
<td>44</td>
</tr>
<tr>
<td>Skin</td>
<td>Fructose-lysine</td>
<td>0.65</td>
<td>&lt;0.0001</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>2-Aminoacidipic acid</td>
<td>0.29</td>
<td>0.005</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>Glucosepane</td>
<td>0.37</td>
<td>0.001</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>Pentosidine</td>
<td>0.26</td>
<td>0.012</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>CEL</td>
<td>0.14</td>
<td>0.21 [NS]</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>CML</td>
<td>0.13</td>
<td>0.23 [NS]</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>MetSOX</td>
<td>0.07</td>
<td>0.68 [NS]</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>MG-H1</td>
<td>0.04</td>
<td>0.81 [NS]</td>
<td>34</td>
</tr>
<tr>
<td></td>
<td>G-H1</td>
<td>0.01</td>
<td>0.95 [NS]</td>
<td>34</td>
</tr>
</tbody>
</table>
3.3.2. Glucosepane is a better marker for diabetes than CML.

The other striking finding in the comparative studies is that only fructosyl-lysine and glucosepane are significantly affected by diabetes, whereas CML, MG-H1, G-H1 and CEL are not obviously impacted by diabetes (type 2). This is different from our previous studies in type 1 diabetic patients, in which skin CML levels were elevated. The medication information for these two studies are not available, but it is well known that the major treatment for type 1 diabetes is the injection of insulin to lower the blood glycemic level, and many drugs for type 2 diabetes such as metformin themselves are also scavengers of dicarbonyl species, which may lower MGO and GO to the normal levels. On the other hand, chelating reagents are also considered to be anti-diabetic drugs. CML and MG-H1 levels may not be elevated by hyperglycemia if the amounts of redox active transition metal ions are decreased by chelation. The formation of fructosyl-lysine and glucosepane may be intensified if glycoxidation was inhibited. Glucosepane in skin may be a better marker than CML for assessment of the progress of type 2 diabetes. It is also interesting to consider using anti-crosslinking reagents as supplements to the diabetic medication.

3.3.3. Multiple sources for dicarbonyl species formation.

In lens, an obvious source of methylglyoxal is glycolysis, which is active in lens despite the absence of synthetic and mitochondrial machinery. However, the MG-H1 level is much higher in lens than skin. A potential explanation is that ascorbic acid levels in lens are quite high (1–3 mM), and substantial amounts of MG-H1 may come from ascorbate degradation. Fig. 5 Shows that when bovine lens crystallin homogenate was
incubated with 5 mM ascorbic acid or DHA, CML, MG-H1, G-H1 and CEL in lens proteins were dramatically elevated. This suggests that ascorbic acid is a potential source of dicarbonyl species in lens. It should be noticed that CML levels are higher than MG-H1 level in Fig. 5, whereas the opposite result is obtained in Fig. 7. We thought the high concentration of glutathione in lens can inhibit the glycoxidation of ascorbylated lysine to CML, while MGO can still originate from ascorbic acid degradation, which resulted in the large amounts of MG-H1 in lens.

Unlike MG-H1, G-H1 level is extremely low in lens, suggesting that the glyoxal level in lens is quite low. On the other hand, the G-H1 level is higher than MG-H1 (Fig. 7) in skin. This result is contradictory to Fig. 2, which shows that the formation of MG-H1 was favored over G-H1 when collagen was incubated with glucose. This may be explained by the fact that the major source of glyoxal in biological systems is lipoxidation. In the lens, the high level of glutathione may inhibit lipoxidation, which results in the extremely low levels of G-H1.

**3.3.4. Lysyl oxidation is more prevalent in skin than in lens.**

Allysine levels are higher in skin than in lens and do not increase with age (Fig. 6). Whereas its end product 2-aminoadipic acid also increases in the aging lens. The absolute level in old skin is ~10× higher than in lens, and the age-related increase is much more dramatic in skin. In the lens, the magnitude of the insult is kept in check by glutathione, which is present at ~4 mM levels, but only at 50 μM in serum. A role for hydrogen peroxide was postulated with the demonstration that catalase suppressed allysine formation.
3.4. Conclusions

The comparative and comprehensive analysis of AGE accumulation in two tissues exposed to drastically different oxygenation conditions has provided new insights into the metabolic nature, magnitude, and insult of carbonyl stress to long-lived proteins. A major conclusion is that the aging skin is much more prone to oxidative damage than the lens. Modifications, especially the cross-links, of the lens proteins, become important for the formation of senile cataracts. Although several other AGEs and other oxidation products would have to be added to provide a more complete picture, we do not expect major changes in the AGE hierarchy presented above, because many of the other products measured in skin or lens, such as MOLD, GOLD, GODIC, MODIC, DOGDIC, and DOGDIC-ox, are present in minor amounts compared to the major modifications listed in Fig. 7. The latter includes ornithine, MG-H1, and fructoselysine in lens and ornithine, glucosepane, and fructose-lysine in skin.

3.5. Experimental section

**General.** Standards of L-2-aminoadipic acid and d₈-lysine (DL-[3,3,4,4,5,5,6,6-²H₈]lysine), d₃ methionine (L-[methyl-²H₃]methionine), and d₄-methionine (DL-[3,3,4,4-²H₄]methionine) were purchased from Sigma–Aldrich (St. Louis, MO, USA); L-[guanidine-¹⁵N₂]arginine was purchased from Cambridge Isotope Laboratories (Andover, MA, USA), CEL (Nε-carboxyethyllysine) and d₄-CML (Nε-carboxymethyllysine-4,4,5,5-d₄) were from Drs. Susan Thorpe and John Baynes (Department of Chemistry and Biochemistry, University of South Carolina, Columbia, SC, USA); CML was from
Chelex 100 Resin was bought from Bio-Rad (Hercules, CA). All other chemicals were purchased from Sigma. Deionized water (18.2 MΩ cm⁻¹) was used for all experiments.

**Synthesis of isotopically labeled AGE standards**

*d₃-2-Aminoadipic acid*. Thionyl chloride (0.5 mL) was added dropwise to a solution of normal 2-aminoadipic acid (161 mg, 1 mmol) in methanol (10 mL). The resulting solution was stirred 15 min at r.t, and evaporated. The formed dimethyl aminoadipate was dissolved in THF/water (1:1, 10 mL) and tert-butyl dicarbonate (240 mg, 1.1 mmol) was added into the solution after the pH was adjusted above 9 with Na₂CO₃. The reaction mixture was stirred for an additional 6 h, and filtered. The filtrate was extracted with ethyl acetate (3 x 20 mL). The combined organic layers were dried over anhydrous sodium sulphate, filtered, and solvents were evaporated. The resulting oil was added into a sodium methoxide solution in methanol-OD (1M, 10 mL) and stirred overnight at r.t. The reaction was quenched by adding D₂O (10 mL) and the resulting solution was stirred for an additional 30 min at r.t., then acidified by HCl (0.5 N, pH 4), and extracted with ethyl acetate (3 x 40 mL). The combined organic layers were dried over anhydrous sodium sulphate, filtered, solvents were evaporated, and the residue was chromatographed on a silica gel column with ethyl acetate/hexanes/acetic acid (1: 3 : 0.02) as eluent. The fractions containing the product of interest were combined, and the solvents were evaporated under reduced pressure, then the residue was dissolved in 1 mL TFA and stirred for 0.5 hour at r.t. The solution was evaporated to give 49 mg of 2,5,5-[³H₃]-2-aminoadipic acid (18.5% yield). ¹H NMR (400 MHz, D₂O) δ (ppm): 1.4-1.5
(2H), 1.7-1.75 (2H). Mass spectrum (M+H\(^+\)), m/z 165.72, fragments m/z: 147.70; 119.78; 101.76; 100.76; 99.82. Isotopic purity: 76.93%

\(d_3\)-N\(\epsilon\)-(5-Hydro-5-methyl-4-imidazolon-2-yl)-ornithine \((d_3\text{-MG-H1})\). Boc-glutamine (246 mg, 1 mmol) was added into a solution of acetic anhydride: pyridine (1:1 v/v, 2 mL). The resulting solution was stirred at r.t. for 4 h, then concentrated, and chromatographed on silica gel with the eluent of hexanes/ethyl acetate/acetic acid (2:1:0.1). The resulting nitrile-containing compound was dissolved in acetic acid-OD (10mL), and stirred under deuterium (1 atm) with PtO\(_2\) (10 mg) at r.t. for 24 h. Then the reaction mixture was filtered, and the solvent was evaporated from the filtrate. The resulting crude Boc-\(d_2\)-ornithine was used for the following reaction without purification.

5-Hydro-5-methyl-2-methylthio-4-imidazolone, synthesized as previously published method,\(^{16}\) was deuterated in sodium ethoxide solution in ethanol-OD (0.5 M, 5 mL). The solution was refluxed with the crude Boc-\(d_2\)-ornithine for 6 h. The resulting Boc-\(d_3\)-MG-H1 was purified by semi-preparative RP-HPLC with a flow rate 4 mL/min and a gradient of 0.1% TFA in 5-50% acetonitrile in water from 0 to 30 min. The eluate was collected at 17-18 min, freeze-dried and stirred in TFA (1 mL) for 1 h at r.t. to give 33 mg of \(d_3\)-MG-H1 (10% yield). \(^1\)H NMR (400 MHz, D\(_2\)O) \(\delta\) (ppm): 1.32 (s, 3H), 1.6-1.7 (2H), 1.78-1.85 (2H), 3.74 (t, 1H, \(J = 6\) Hz). Mass spectrum (M+H\(^+\)), m/z 232.67, fragments m/z: 215.69; 214.62; 187.67; 169.63; 159.54; 118.62; 115.61; 72.74. Isotopic purity: 79.47%.

\(d_2\)-N\(\epsilon\)-(5-Hydro-4-imidazolon-2-yl)-ornithine \((d_2\text{-G-H1})\). \(d_2\)-G-H1 was prepared similarly as \(d_3\)-MG-H1 where Boc–ornithine was used instead of Boc-\(d_2\)-ornithine. The resulting Boc-\(d_2\)-G-H1 was purified by semi preparative RP-HPLC as described above. The eluate was collected at 15-16 min, freeze-dried and stirred in TFA (1mL) 1 h at r.t. to
give 31 mg of $d_2$-G-H 1 (10% yield). $^1$H NMR (400 MHz, D$_2$O) $\delta$ (ppm): 1.55-1.61 (2H), 1.78-1.86 (2H), 3.25 (t, 2H, $J = 5.1$ Hz), 3.84 (t, 1H, $J = 6.2$ Hz), $^{13}$C NMR (150 MHz D$_2$O) $\delta$ (ppm): 23.6, 26.9, 41.5, 49.2, 52.7, 157.2, 172.0, 174.3. Mass spectrum (M+H$^+$), $m/z$ 217.80 fragments $m/z$: 198.84; 171.63; 153.98; 125.93; 115.88; 101.94; 70.10. Isotopic purity: 84.18%

$2, 2'$-[^2$H$_2$]$-2'$-[^13$C$_1$]-N$^e$-Carboxyethyl-lysine ($d_2$-[^13$C$_1$]-CEL). $^{13}$C$_1$-pyruvic acid (0.2 mmol) and Boc-lysine (100 mg, 0.4 mmol) were dissolved in sodium phosphate buffer (2 mL, 200 mM, pH 7.4). Then sodium cyanoborohydride (200 mg, 3 mmol) was added in batches. After stirring at r.t. for 2 h, the pH of the reaction solution was adjusted to 8.5. Then tert-butyl dicarbonate (230 mg, 1 mmol) was added into the solution. the resulting solution was stirred at r.t. for 4 h, and then acidified by 0.1 N HCl to a pH below 4 and extracted with ethyl acetate (3 x 20 mL). The combined organic layers were dried over anhydrous sodium sulphate and filtered. The filtrate was evaporated in vacuum, and dissolved in anhydrous DMF (2 mL). Anhydrous K$_2$CO$_3$ (500 mg) and CH$_3$I (500 mg) was added into the resulting solution, and stirred for 1 h. Then the reaction was quenched by adding water (10 mL), the resulting mixture was extracted with ethyl acetate (3 x 20 mL). The combined organic layers were dried over anhydrous sodium sulphate, filtered, and concentrated. The resulting oil was converted to $d_2$-[^13$C$_1$]-CEL (20 mg, 45% yield) as described above for the conversion of amino adipic acid into $d_2$-amino adipic acid. $^1$H NMR (400 MHz, D$_2$O) $\delta$ (ppm): 1.31 (d, 3H, $J_{CH} = 131$ Hz), 1.28-1.32 (2H), 1.51-1.57 (2H), 1.65-1.71 (2H), 2.85 (m, 2H). $^{13}$C NMR (100 MHz, D$_2$O) $\delta$ (ppm): 14.3, 21.5, 25.3, 29.2, 45.5, 172.4; Mass spectrum (M+H$^+$), $m/z$ 222.00, fragments $m/z$: 159.79; 131.65; 85.78; 84.78. Isotopic purity: 88.84%.
Human lens and skin samples. A total of 54 human lenses (n = 23 with type 1 or 2 diabetes, and n = 31 normal) were obtained from the National Disease Research Interchange (Philadelphia, PA, USA). Lenses were stored at −80 °C until use. The delipidated insoluble protein fraction was prepared as described earlier. The collagen samples used in this study originated from a collection of n = 117 “highly insoluble” collagen fractions extracted from human abdominal skin and stored in dry form at −20 °C. Additional insoluble abdominal skin collagen samples obtained at autopsy from n = 14 nondiabetic and n = 20 mostly type 2 diabetic patients were from Dr. Marc Halushka, Department of Pathology, Johns Hopkins Medical Institutions (Baltimore, MD, USA).

The kinetics study of glycation on collagen. 360 mL of soluble bovine type I collagen (4.1 mg/mL) was aliquoted into 72 dialysis tubings (Spectrum laboratories, Rancho Dominguez, CA, 1 kDa MWCO). Six groups of incubations were carried out at 37 °C. In each group, 12 samples were incubated in 500 mL solutions. The solutions were 200 mM chelex-treated sodium phosphate buffer with 2 mM diethylene triamine pentaacetic acid (DTPA) and 0 mM, 5 mM, or 25 mM glucose, and 200 mM normal sodium phosphate buffer with 0 mM, 5 mM, or 25 mM glucose. All incubation buffers were at pH 7.4, and refreshed every week. Two samples were taken from each incubation group at the following time points: 0 week, 1 week, 2 weeks, 8 weeks, 14 weeks, 16 weeks. Each sample was dialyzed against deionized water (4 x 1 L) at 4 °C for 2 days, then transferred to tubes, and freeze-dried. All the samples were stored at −20 °C.

Intensive glycation on both human alpha A crystallin and bovine collagen. Bovine collagen (20 mg) was incubated in chelex-treated sodium phosphate buffer (100 mM pH 7.4) with 500 mM glucose and 1 mM DTPA at 37 °C for 4 weeks. All the
incubation buffers contain 0.02% sodium azide as antibiotics. The glycated collagen was washed with deionized water (10 x 2 mL), and freeze-dried. The glycation of human alpha A crystallin with 500 mM glucose was described in chapter 2. All the samples were stored at –20 °C.

**Exhaustive enzymatic digestions.** All the lens protein samples were proteolytically digested by the method described in chapter 2. All the collagen samples were digested by the following method: every 1 mg of collagen sample was suspended in 200 μL of buffer H (0.02 M HEPES and 0.1 M calcium chloride pH 7.5). Then 28 units of collagenase, 20 mU of peptidase, 200 mU pronase, 40 mU amino peptidase and 10 mU prolidase were sequentially added in 5 days. All the digestions were carried out at 37 °C. And chloroform and toluene (1.5 μl) were added to every tube as antimicrobial agents. The digestion efficiency was determined by hydroxyproline assay.

**AGE determination by GC/MS and LC/MS.** Amounts of 6-hydroxynorleucine, 2-aminoadipic acid, CML, and CEL were determined in acid hydrolysates of processed lens samples and derivatized as their trifluoroacetyl methyl esters by selected-ion monitoring GC/MS as previously described. FL, MetSOX, G-H1, MG-H1, and glucosepane in enzymatically digested sample were measured by electrospray positive-ionization–mass spectrometric multiple reaction monitoring by using an LC/MS/MS system composed of a 2690 separation module with a Quattro Ultima triple-quadrupole mass spectrometry detector (Water-Micromass, Milford, MA, USA) following the previously published procedure by Ahmed and Thornalley. Analytes released by autodigestion of proteases in assay blanks were subtracted from analytic estimates.
**Statistical methods.** Linear regression analysis and Spearman's correlations were computed using SPSS software (version 11.5 or 16). Testing for homogeneity of variance was done using either the F test or the Burr–Foster Q test, as previously described.\textsuperscript{28} The statistical comparisons of regression lines for differences in their slopes, i.e., representing the accumulation rates between parameters, and line elevations were done by the methods of Snedecor and Cochran. Statistical probability was determined by SPSS software or by using a critical-F calculator hosted online at [www.danielsoper.com](http://www.danielsoper.com). Significance was considered at P > 0.05, and not significant (NS) is indicated where necessary.

3.6. References


16. Ahmed, N.; Thornalley, P. J., Chromatographic assay of glycation adducts in human serum albumin glycated in vitro by derivatization with 6-aminoquinolyl-N-


Chapter 4

Mechanistic insight into EKODE modification of protein side-chains
4.1. Introduction

The non-enzymatic oxidation of the polyunsaturated fatty acids (PUFAs) represents an imbalance between reactive oxygen spices (ROS) and the regulatory system in biological systems.\(^1\) The resulting lipid hydroperoxides (LOOHs) can further evolve into a series of carbonyl species such as acrolein, malondialdehyde (MDA), 4-hydroxy-nonenal (HNE) and 9-hydroxy-12-oxo-10-dodecenoic acid (HODA), which can covalently bind to proteins and DNA.\(^2\) During the last two decades, the intensive research on HNE, a product of linoleic acid peroxidation, revealed its strong correlations to vascular and neurodegenerative diseases.\(^3\) The cytotoxicity of HNE originates from its ability to form covalent adducts with nucleophilic residues such as lysine, histidine and cysteine.\(^2,4\) It should be noted that HNE shows striking reactivity to cysteine, and its stable Michael adduct to cysteine accounts for much of the dysfunction of active enzymes caused by HNE.\(^5\) The major metabolism pathway for elimination of 4-HNE was presumed to be caused via glutathionylation by glutathione S-transferase (GST) and subsequent excretion through the kidneys,\(^6\) as well as through hepatocyte metabolism through the alcohol dehydrogenase (ADH) and aldehyde dehydrogenase (ALDH) pathways,\(^7\) though little was known about the latter process. Recently, in collaboration with Zhang et al.,\(^8\) we found that 4-hydroxynonanoate (the 4-hydroxy acid derived from 4-HNE) is primarily catabolized through a pathway that phosphorylates the C-4 hydroxyl and isomerizes it to a C-3 hydroxy compound, which is catabolized through beta-oxidation.

HNE is derived from the metal-catalyzed fragmentation of linoleic acid hydroperoxides.\(^9\) However this is only one of many pathways for linoleic acid
peroxidation. Recently, the Sayre lab revealed that under mild oxidation conditions (Fe(II) + ascorbic acid), linoleic acid also can be polyoxgenated while retaining the complete length of its acyl chain. The resulting compounds (Fig. 1) are named epoxyketoctadecenoic acids (EKODEs), which causes the major modification of apomyoglobin by linoleic acid non-enzymatic oxidation products. The fact that the EKODE-peptide adducts in mass spectrometry was intensified after NaBH₄ reduction suggests the major chemistry underlying the modifications is Michael addition, which was further confirmed by NMR studies on reactions of trans-5,6-epoxy-3(E)-octen-2-one, a model compound of EKODE I, with imidazole.

![Figure 1. The EKODE isomers and their model compounds.](image)

Even though a series of studies had previously revealed that EKODEs stimulate steroid hormone production and secretion and increase cellular calcium in biological
systems, we still are fundamentally interested in their pathological involvement as potential toxins contributing to age-related diseases, as a consequence of their covalent adduction to protein side chains. Our previous studies assumed that the Michael addition of histidine accounts for the major modifications of apomyoglobin by EKODEs, and that the reactions of EKODE II isomers were much faster than those of EKODE I isomers. However, the Michael addition product of EKODE II is not confirmed yet. Moreover, the reactions of EKODEs to cysteine have not been elucidated. They likely represent the major modifications by EKODEs in biological systems and have deleterious consequences like HNE. Our lab has successfully used amino acid surrogates to study the mechanism of HNE and ONE modifications to protein side chains. In the present study, we continued applying this strategy to reveal the chemistry of EKODE modifications, and then to extend the acquired knowledge to the modification of amino acids.

4.2. Results and discussion

4.2.1. Preparation of EKODE model compounds.

We simplified all the trans EKODE isomers to two model compounds trans-5,6-epoxy-3-(E)-undecen-2-one (1) and trans-5,6-epoxy-2-(E)-undecen-4-one (2) (Fig. 1 and Scheme 1), models of EKODE I and II respectively. Also trans-5,6-Epoxy-3-(E)-octen-2-one (3) was prepared. The synthesis of 1 and 3 followed our previous studies. The synthesis of compound 2 was started with epoxy-octanal 4, which was elongated by Grignard reagent 5. The resulting alcohol 6 was oxidized by pyridinium chlorochromate (PCC) to give 2. The low yield of 2 implied the instability of EKODE II. The synthesis of cis-EKODE model compounds was not pursued, based on the presumption that the
modification modes of cis-EKODEs were analogous to those of trans-EKODEs. 1-Butanethiol and imidazole were applied as the surrogates of cysteine and histidine respectively. Except where stated elsewise, all the model studies were carried out in the sodium phosphate buffer (50 mM, pH 7.4, containing 20% CH₃CN) to mimic physiological conditions.

Scheme 1. The synthesis of EKODE model compounds.
In this study, we also revealed an unusual modification of cysteine by an EKODE II model. To explore the underlying mechanism, we also synthesized three analogs of compound 2: compound 7, 8, and 10 (Scheme 1). The synthesis route of 7 was similar to that of 2 except for the use of different Grignard reagents. The synthesis of 10 followed the published method.13

4.2.2. The reactions of EKODE I models with butanethiol.

To understand the chemistry of EKODE I modifications to thiol residues, the reaction of model compound 1 with butanethiol (1:1) was first examined. After a 3 day incubation under dark conditions at 37 °C, compound 1 was mostly converted to a new product with strong UV absorbance, which was identified as a epoxy-ring opening product 11 (Scheme 2). A similar product 12 was also obtained upon incubation of compound 3 with butanethiol at r.t. for only 12 h. The fact that no significant quantity of 11 was detected in the reaction of 1 to butanethiol at r.t. suggests that a steric hindrance impedes the reaction.

The attack of thiol groups on epoxy rings has been widely observed, and most of these reactions require bases or Lewis acids as catalyst.14 The position of thiol attack is normally close to electron-deficient groups with the least steric hindrance.15 Heating or microwave irradiation can accelerate the reaction.16 In our study, the reaction was carried out in neutral solution, and the reaction of 3 with butanethiol takes place even without heating. Based on this fact, we postulated that the Michael addition product of EKODE I is the precursor of the thiolysis product, and the conversion to the final product was facilitated by an intramolecular [1, 2] shift. (Scheme 3)
Scheme 2. The reactions of butanethiol with EKODE I model compounds.

\[
\begin{align*}
\text{O} & \quad \text{BuSH} \\
\text{1} & \quad \text{BuSH} \\
\end{align*}
\]

Scheme 3. Proposed mechanism for the reaction of thiol with EKODE I model

This hypothesis was verified by monitoring the reaction process spectrophotometrically, in particular with the chromophore change of \( \lambda_{232 \text{ nm}} \) (Fig. 2).

When compound 1 (1 mM) was incubated with butanethiol (1 mM) at 37 °C, the absorbance of the solution dropped rapidly in first 3 h, suggesting that the Michael addition dominated during this period. After 3 h, the epoxy ring opening reaction
replaced the Michael addition, resulting in the recovery chromophore (Fig. 2. A). The major final product was 11, which can further react with butanethiol to form 14. The plateau of the chromophore recovery was achieved in 30 hours, implying UV irradiation by the spectrophotometer promotes the epoxy ring opening reaction. When both the concentrations of 1 and butanethiol were decreased to 0.1 mM, the Michael addition of butanethiol to 1 was slowed. But the light intensity of the UV-spectrometry was not decreased. Notably the rate of epoxy-ring opening reaction was not changed significantly, and was always higher than that of the Michael addition of compound 1 (Scheme 3). The drop of UV absorbance resulted from the Michael addition of butanethiol to compound 11. Therefore, the solution chromophore decreased slowly without any recovery. When the butanethiol concentration was increased to 1 mM, the Michael addition rate was elevated due to the fact that the ratio of butanethiol/1 was increased to 10:1. Consequently, the chromophore recovery was observed again because the stability of the Michael adduct 14 is lower than that of 1 (Fig. 2. B).

When butanethiol was replaced with N-acetyl-cysteine methyl ester, the Michael addition was accelerated by the higher reactivity of the cysteine thiol residue. This resulted in the rapid drop of the UV-absorbance at the beginning phase. The chromophore began to recover after 12 hours of incubation, even though the recovery was quite slow. This suggests that the thiolysis of EKODE 1 was strongly impacted by a steric effect (Fig. 2. C). In addition, 1 (1 mM) was incubated with N-acetyl-cysteine methyl ester (1 mM) in sodium phosphate buffer (50 mM, pH 7.4) containing 5% d<sub>6</sub>-DMSO in the dark at 37 °C for 24 hours, the solution was monitored by 600 MHz proton nuclear magnetic resonance (1H NMR) (Figure 3). The use of normal sodium phosphate buffer can inhibit
the H-D exchange. Consequently, the recovery of the olefin peaks in the epoxy ring opening reaction can be observed. However, only one set of olefin peaks (6.23 ppm and 6.48 ppm) were observed, which corresponded to compound 1. And a singlet peak at 2.03 ppm corresponded to the CH\textsubscript{3} peak of C1 in the Michael adduct, the major product in the incubation. These observations suggested that only the starting material 1 and Michael adducts from 1 were present in the solution, and gave more support to the hypothesis that the thiolysis of EKODE I was promoted by light, and inhibited by steric effects.

4.2.3. The reactions of an EKODE II model with butanethiol.

Inspired by the EKODE I model reactions, I investigated the model reaction of 2 with butanethiol. Unlike 1, the Michael adduct of 2 to butanethiol is quite stable, and it was successfully isolated by extraction from the reaction mixture with ethyl acetate. The resulting Michael adducts 15 consists of two diastereomers generated in a 1:1 ratio. The adduct 15 can be converted into the epoxy ring opening product 16 by incubating the reaction mixture at r.t. for 48 h (Scheme4). This result is consistent with our previous studies showing that EKODE II is much more reactive than EKODE I.
Figure 2. UV-monitoring of the reactions of 1 with butanethiol or N-acetyl cysteine methyl ester. All the reactions were carried out in sodium phosphate buffer (50 mM, pH 7.4) containing 20% acetonitrile at 37 °C. The UV-absorbance was detected at 232 nm. A. the reactions of compound 1 with butanethiol (1:1) at 0.1 mM or 1 mM; B. the reactions of 1 (0.1 mM) with butanethiol with 1 fold, 2 fold and 10 fold butanethiol respectively; C. the reaction of 1 to N-acetyl cysteine methyl ester (0.1 mM).
Figure 3. The reactions of 1 to N-acetyl cysteine methyl ester, monitored by 600 MHz NMR. The reaction was executed in normal sodium phosphate buffer containing 5% d₆-DMSO at 37°C. The bottom panel is the NMR spectrum of pure 1; the middle panel is the NMR spectrum after 1 hour of incubation; the top panel is the NMR spectrum after 24 h of incubation. No epoxy ring opening product was observed in NMR spectra of the reaction product mixtures.

It is interesting that thiol was added at C6 of compound 2 instead of C5. This was confirmed by unambiguous ¹H NMR and ¹H-¹H correlation spectroscopy (HHCOSY).
(Fig. 4). The peaks at 3.52 ppm correspond to the proton of the OH group, which was easily confirmed by the fact that treating 16 with CH₃OD before NMR analysis removed the peaks completely. The correlation between the OH peaks and the peaks at 4.48 ppm suggests that the OH group is on C5, which is supported by the fact that there is no correlation of peaks at 4.48 ppm with peaks in the range 1-2 ppm.

Fringuelli et al. reported that thiolysis of epoxy ketones occurs selectively at the alpha position to ketones.¹⁷ The reasonable explanation to these two contradictory results is that 2 first reacted with butanethiol to form the Michael adduct 15, which was further converted to final product 16 through an intra-molecular [1, 5] shift. Moreover, the formation of the intermediate 17 required that C2-C5 and the sulfur atoms were in the same plane, which meant sulfur was closer to C6 compared to C5, and the attack on the beta position faced less steric hindrance (Scheme 5).

To confirm this theory, compounds 7, 8 and 10 were incubated with butanethiol with a 1:1 ratio at r.t. Alkyne 7 was converted into alkene 18 quantitatively in 4 hours. And the resulting product 18 was stable in the aqueous incubation solution for more than 2 days at 37 °C. This revealed that the Michael addition to alkyne 7 is much more favored than that to alkene 18, and the geometry of alkene 18 does not favor the formation of six-membered ring or five-membered ring intermediates (Scheme 6). This result supports the hypothesis that the elimination of the Michael adduct to form the six-membered ring intermediates is the key step for thiolysis of the epoxy-ring (Scheme 5). Meanwhile, 8 and 10, without the interruptions of Michael adducts, undergo the alpha thiolysis to form the products 19 and 20 respectively in 24 h (Scheme 6). These results are consistent to Fringuelli et al.’s result.
Scheme 4. The reaction of compound 2 to butanethiol.

Figure 4. The HHCOSY spectrum of compound 16.
Scheme 5. The proposed mechanism of the EKODE-II model epoxy ring opening reaction.

Scheme 6. The reactions of EKODE II analogues with butanethiol.

The spectrophotometric data also supports this hypothesis. When 2 (0.1 mM) was incubated with butanethiol (0.1 mM) at 37 °C, the UV absorption of the solution decreased rapidly corresponding to the higher Michael addition reactivity of the EKODE II model. But the consequent chromophore recovery was faster than that of compound 1, which supports the conclusion that the intermediate of the EKODE II model reaction is much more stable compared to the EKODE I model (Fig. 5). This difference comes from
the facts that a six-membered ring is more stable than a three-membered ring, and [1, 5] intra-molecular shift is favored compared to [1, 2] shift.

Figure 5. The reactions of 2 with butanethiol or N-acetyl cysteine methyl ester. (0.1 mM, 1:1).

It should be noted that compound 2 is not stable. It slowly degrades in aqueous solution at 37 °C. The mechanism is not clear yet, but the decrease of the chromophore suggests that a Michael addition is involved. But the thiolysis product 16 is quite stable in aqueous solution (Fig. 5). These facts suggest that inadequate detoxification of EKODE II by thiols may elongate the toxin’s lifespan.

The slower chromophore recovery in the cysteine reaction with compound 2 cannot be simply considered as a contribution of steric hindrance. The NMR studies showed that the final products are a mixture of compound 2, 16 and an unidentified compound (Fig. 6).
Figure 6. The reactions of 2 with N-acetyl-cysteine methyl ester (1 mM) in sodium phosphate buffer pH 7.4 containing 5% $d_6$-DMSO. The blue labeled peaks correspond to the olefin peaks of the starting material 2, the yellow labeled peaks correspond to the olefin peaks of epoxy ring open product 21, peaks at 5.8 ppm correspond to an unidentified compound.
4.2.4. The reactions of EKODE model compounds with imidazole.

Since the reaction of EKODE I with imidazole has been well studied by Lin et al.,\textsuperscript{10} only the reaction of the EKODE II model compound 2 with imidazole was investigated in this study. 1 mM 2 was incubated with 3 mM imidazole in sodium phosphate buffer (50 mM, pH 7.4) containing 20% CD$_3$CN, and the reaction was monitored by 600 MHz NMR (Fig. 7). The reaction started rapidly when the compounds were mixed together, and finished in 3 hours. The resulting products were clean Michael adducts with two diastereomers in almost a 1:1 ratio (Fig. 8). The two diastereomers are defined by the position of imidazole to the epoxy-ring as anti-adducts (Fig. 8. A) and syn-adducts (Fig. 8. B). Lin et al. revealed that both the reactions of EKODE I and 5,6-epoxy-2-heptenal (EH) (unpublished data) with imidazole showed stereoselectivity, which resulted in the anti-diastereomers being favored.\textsuperscript{10} This implied that EKODE II has less steric hindrance for Michael additions. The imidazole Michael adducts of EKODE II also can survive longer in incubations than EKODE II itself, and the thiolysis to the Michael adducts is still possible due to the epoxy ketone residues.
Figure 7. The reaction of 2 to imidazole. Top: 1 mM 2 and 3 mM imidazole were incubated in sodium phosphate buffer (50 mM pH 7.4) containing 20% CD$_3$CN for 3 hours; Middle: incubated for 15 min. Bottom: pure 2.
Figure 8. Identification of the two isomers of imidazole Michael adducts to an EKODE II model by HHCOSY. The peaks of Hb were buried in the water peaks.
4.2.5. The reactions of EKODE model compounds with lysine.

Model compounds 1 and 2 were incubated with 3 equivalents of N^α-acetyl lysine methyl ester or 20 equivalents of butylamine at 37 °C for 24 hours. No significant amounts of Michael adducts or Shiff base were detected. This result is consistent with the results of our previous proteomics study. In this study, linoleic acid was oxidized under mild oxidation conditions, no lysine residues were modified by EKODEs.

4.3. Conclusions

The foregoing results revealed that models of EKODE isomers can form stable thiolyisin products under physiological conditions. The resulting compounds can further covalently adduct cysteine or histidine residues. Considering the high possibility that the histidine-EKODE Michael adducts can further adduct cysteine residues, EKODEs likely are very important long-lifespan chronic toxins as a biological consequence of lipid peroxidation, and it is important to further investigate the protein damage originating from EKODEs.

4.4. Experimental section

**General.** Unless otherwise stated, the solvents and reagents were of commercially available analytical grade quality. ^1^H NMR spectra (400 or 600 MHz) and ^13^C NMR spectra (100 or 150 MHz) were recorded on Varian Gemini and Varian Inova instruments.

**trans-2,3-Epoxyoctanal (4).** A H_2O_2 solution (20 mL, 200 mmol) was added into a solution of trans-2-octenal (10 g, 80 mmol) in methanol (30 mL). The resulting solution
was cooled to 4 °C, and then NaHCO₃ (3 g, 25 mmol) was added. The resulting mixture was stirred further at r.t. for 3 h. Then the reaction was quenched by 5 % acetic acid to pH below 6. The solution was extracted with CH₂Cl₂ (3 x 200 mL). The combined organic layers were dried over anhydrous Na₂SO₄, and filtered. The solvents were evaporated from the filtrate, and the crude product was purified by silica gel chromatography with hexanes/EtOAc (9:1) as eluent, affording 4 as a colorless oil (8.4 g, 75%).¹H NMR (400 MHz, CDCl₃) δ 0.90 (t, 3H, J = 7.0 Hz), 1.22-1.75 (8H), 3.14 (dd, 1H, J = 6.4 Hz and 2.0 Hz), 3.23 (td, 1H, J = 5.3 Hz and 2.0 Hz), 9.02 (d, 1H, J = 6.4 Hz).

**trans-5,6-Epoxy-3-(E)-undecen-2-one (((E)-4-(3-pentyloxiran-2-yl)but-3-en-2-one) (1).** 1-Triphenylphosphoranylidene-2-propanone (1.1g, 3.5 mmol) was added into a solution of 2,3-epoxyoctanal 3 (280 mg, 2 mmol) in anhydrous CH₂Cl₂ (10 mL) at r.t., and stirred for 1 hour. Then the solution was concentrated and purified by flash chromatography on the silica gel column with hexanes/ethyl acetate (10:1) as eluent, providing a yellow oil 1 (200 mg, 55% yield). ¹H NMR (400 MHz, CDCl₃) δ 0.90 (t, 3H, J = 7.0 Hz), 1.30-1.34 (4H), 1.40-1.52 (2H), 1.62 (m, 2H), 2.26 (s, 3H), 2.91 (td, 1H, J = 5.6 Hz and 2 Hz), 3.21 (dd, 1H, J = 6.8 Hz and 1.6 Hz), 6.35 (d, 1H, J = 16 Hz), 6.47 (dd, 1H, J = 16 Hz and 6.8 Hz); ¹³C NMR (100 MHz, CDCl₃) δ 14.54, 22.68, 25.71, 27.58, 31.63, 31.90, 56.53, 61.24, 133.01, 145.31, 198.12;

**trans-5,6-Epoxy-3-(E)-octen-2-one (((E)-4-(3-ethylxiran-2-yl)but-3-en-2-one) (3).** The synthesis of 6 is basically same as the synthesis of 1.¹H NMR (400 MHz, CDCl₃) δ 1.01 (t, 3H, J = 7.4 Hz), 1.55-1.70 (2H), 2.25 (s, 3H), 2.89 (td, 1H, J = 5.5 Hz and 2.0 Hz), 3.23 (dd, 1H, J = 6.5 Hz and 2.0 Hz), 6.36 (d, 1H, J = 16.0 Hz), 6.49 (dd, 1H, J = 16.0 Hz and 6.5 Hz)
(E)-1-(3-Pentyloxiran-2-yl)but-2-en-1-ol (6). A solution of 1-propenyl-magnesium bromide 5 (20 mL, 10 mmol) was added dropwise into the solution of epoxy aldehyde 4 (2 g, 14 mmol) in ether (10 mL) at –20 °C in 1 hour. The resulting mixture was further stirred at r.t. for another hour. Then the reaction was quenched with the saturated NH₄Cl solution (40 mL). The aqueous layer was further extracted by ether (3 x 40 mL). The combined organic layers were dried over anhydrous Na₂SO₄ and filtered. The filtrate was concentrated, and the crude product was purified by silica gel chromatography with hexanes/ethyl acetate (5:1) as eluent, providing epoxy alcohol 6 (1.4 g, 77% yield) ¹H NMR (400 MHz, CDCl₃) 0.89 (t, 3H, J = 6.8 Hz), 1.29-1.37 (4H), 1.41-1.48 (2H), 1.50-1.64 (2H), 1.74 (dd, 3H, J = 7.2 Hz and 1.6 Hz), 1.9-2.1 (b, 1H), 2.82 (dd, 1H, J = 2.4 Hz and 2.4 Hz), 3.03 (td, 1H, J = 5.6 Hz and 2.4 Hz), 4.71 (dd, 1H, J = 8.4 Hz and 2.4 Hz), 5.36 (ddq, 1H, J = 10.8 Hz, 8.4 Hz and 1.6 Hz), 5.73(dqd, 1H, J = 10.8 Hz, 7.2 Hz and 1.2 Hz); ¹³C NMR (100 MHz, CDCl₃) δ 13.81, 14.21, 22.76, 25.86, 31.69, 31.74, 55.08, 60.32, 64.85, 128.16, 129.46

trans-5,6-Epoxy-2-(E)-undecen-4-one ((E)-1-(3-pentyloxiran-2-yl)-2-buten-1-one) (2). A mixture of PCC (2 g, 9 mmol) and celite (5g) was added to a solution of epoxy alcohol 6 (1 g, 5.4 mmol) and triethylamine (1 g, 10 mmol) in anhydrous CH₂Cl₂ (10 mL). The resulting suspension was stirred at r.t. for 1 hour, then filtered. The filtrate was concentrated and purified by silica gel chromatography with hexanes/ethyl acetate (20:1) as eluent, giving the epoxy ketone 2 (200 mg, 20 % yield). ¹H NMR (400 MHz, CDCl₃) 0.88 (t, 3H, J = 7.2 Hz), 1.30-1.36 (4H), 1.45-1.50 (2H), 1.57-1.68 (2H), 1.92 (dd, 3H, J = 7.0 Hz and 1.6 Hz), 3.05 (ddd, 1H, J = 6 Hz, 4.8 Hz and 2 Hz), 3.34 (d, 1H, J = 2 Hz), 6.25 (dq, 1H, J = 15.6 Hz and 1.6 Hz), 7.09 (dq, 1H, J = 15.6 Hz, 7.0 Hz); ¹³C
NMR (100 MHz, CDCl₃) δ 14.16, 18.81, 22.72, 25.71, 31.63, 32.00, 58.56, 59.20, 125.85, 145.79, 195.75

1-(3-Pentyloxiran-2-yl)but-2-yn-1-one (7). 1-Propynylmagnesium bromide (10 mL, 5 mmol) was added dropwise into the solution of epoxy aldehyde 4 (1 g, 7 mmol) in ether (10 mL) at −20 °C in 1 hour. The resulting mixture was then stirred at room temperature for an additional hour. Then the reaction was quenched with the saturated NH₄Cl solution (20 mL). The aqueous layer was then extracted with ether (3 x 20 mL). The combined organic layers were dried over anhydrous Na₂SO₄, filtered, and the solvents were evaporated. The resulting oil was dissolved in anhydrous CH₂Cl₂ (5 mL). Then a solution of Dess-Martin periodinane (0.3 M, 15 mL) and NaHCO₃ (0.84 g, 10 mmol) were added. The suspension was stirred at room temperature for another hour before the reaction was quenched with water (10 mL). Then saturated Na₂S₂O₃ solution (10 mL) was added. The resulting aqueous phase was then extracted by CH₂Cl₂ (2 x 20 mL). The combined organic layer was dried over anhydrous Na₂SO₄, filtered, concentrated and purified by silica gel chromatography with hexanes/ethyl acetate (20:1) as eluent, giving the epoxy ketone 7 (380 mg, 2.1 mmol, 42% yield). ¹H NMR (400 MHz, CDCl₃) 0.83 (t, 3H, J = 7 Hz), 1.27-1.42 (6H), 1.5-1.6 (m, 2H), 1.99 (t, 3H, J = 0.8 Hz), 3.19 (dt, 1H, J = 5.2 Hz, and 0.8 Hz), 3.23 (m, 1H); ¹³C NMR (100 MHz, CDCl₃) δ 4.53, 14.12, 22.67, 25.65, 31.58, 31.66, 59.26, 60.24, 77.54, 93.86, 184.63.

trans-5,6-Epoxy-undecan-4-one (8). A solution of 1-bromopropane (0.30 g, 2.5 mmol) in ether (5 mL) was added dropwise to a mixture of Mg turnings (0.10 g) and ether (5 mL), and stirred for 2 h. Then a solution of trans-2-octenal (260 mg, 2 mmol) in ether (10 mL) was added dropwise into the suspension, and the mixture was stirred at
room temperature for 1 hour before the reaction was quenched by saturated NH₄Cl solution (20 mL). The organic layer was washed with brine, and dried over Na₂SO₄ and filtered. The solvents were evaporated from the filtrate, and the residue was dissolved in anhydrous CH₂Cl₂ (10 mL). Then Dess-Martin reagent (0.3 M, 3 mL) was added to the solution. The mixture was stirred at r.t. for 30 min, and then quenched with saturated NaHCO₃ solution (15 mL). The aqueous phase was then extracted with CH₂Cl₂ (2 x 20 mL). The combined organic layers were dried over anhydrous Na₂SO₄ and filtered. The solvents were evaporated from the filtrate, and the residue was dissolved in anhydrous CH₂Cl₂ (10 mL). Then Dess-Martin reagent (0.3 M, 3 mL) was added to the solution. The mixture was stirred at r.t. for 30 min, and then quenched with saturated NaHCO₃ solution (15 mL). The aqueous phase was then extracted with CH₂Cl₂ (2 x 20 mL). The combined organic layers were dried over anhydrous Na₂SO₄ and filtered. The solvents were evaporated from the filtrate, and the residue was dissolved in methanol. H₂O₂ (1.0 mL, 30%, 10 mmol) and NaHCO₃ (840 mg, 10 mmol) were added into the solution. The resulting suspension was stirred at room temperature for 4 hours, and then filtered. The filtrate was diluted by brine (90 mL), and extracted with CH₂Cl₂ (3 x 50 mL). The combined organic layers were dried over anhydrous Na₂SO₄, and then filtered. The filtrate was concentrated by evaporation of solvents, and the residue was purified by silica gel chromatography with hexanes/ethyl acetate (10:1) as eluent to give trans-5,6-epoxy-undecan-4-one 8 (175 mg, 52% yield). ¹H NMR (400 MHz, CDCl₃) δ 0.90 (t, 6H, J = 7.0 Hz), 1.29-1.34 (4H), 1.42-1.47 (m, 2H), 1.54-1.65 (4H), 2.25 (m, 1H), 2.43 (m, 1H), 3.03 (td, 1H, J = 5.4 Hz and 0.8 Hz), 3.20 (d, 1H, J = 0.8 Hz); ¹³C NMR δ (100 MHz, CDCl₃) δ 13.89, 14.16, 16.80, 22.71, 25.68, 31.63, 32.03, 39.25, 58.57, 59.89, 208.15.

1-Phenyloct-2en-1-one (9). Amberlyst-15 (500 mg) was added to a mixture of phenylacetylene (500 mg, 5 mmol) and hexanal (500 mg, 5 mmol) in anhydrous CH₂Cl₂, and stirred for 4 hours at room temperature. Then the mixture was filtered. The resulting filtrate was concentrated by evaporation of solvents in vacuo and the residue was
purified by silica gel chromatography with hexanes/ethyl acetate (10:1) as eluent to afford a brown oil 9 (400 mg, 40% yield). $^1$H NMR (400 MHz, CDCl$_3$) $\delta$ 0.89 (t, 3H, $J$ = 7.2 Hz), 1.29-1.34 (4H), 1.48-1.52 (m, 2H), 2.30 (m, 2H), 6.86 (dt, 1H, $J$ = 15.2 Hz and 1.6 Hz), 7.05 (dt, 1H, $J$ = 15.2 Hz and 7 Hz), 7.43 (m, 2H), 7.51 (m, 1H), 7.91 (m, 2H); $^{13}$C NMR (100 MHz, CDCl$_3$) $\delta$ 14.19, 22.68, 28.07, 31.63, 33.03, 126.03, 128.69, 128.72, 132.77, 138.22, 150.34, 191.07.

**(3-Pentyloxiran-2-yl)(phenyl)methanone (10).** The ketone 9 (200 mg, 1 mmol) was dissolved in methanol/water (9:1, 10 mL). Then Na$_2$CO$_3$ (216 mg, 2 mmol) and H$_2$O$_2$ (0.4 mL, 30%, 4 mmol) were added. The resulting mixture was stirred at room temperature for 4 hours, and then filtered. The filtrate was diluted with brine solution (90 mL), and extracted with ethyl acetate (3 x 50 mL). The combined organic layers were dried over anhydrous Na$_2$SO$_4$, and filtered. The filtrate was concentrated by evaporation of solvents *in vacuo* and the resulting oil was purified by silica gel chromatography with hexanes/ethyl acetate (10:1) as eluent to afford a brown oil 10 (180 mg, 82% yield). $^1$H NMR (400 MHz, CDCl$_3$) 0.89 (t, 3H, $J$ = 7.2 Hz), 1.31-1.36 (4H), 1.450 (m, 2H), 1.72 (m, 2H), 3.13 (ddd, 1H, $J$ = 6 Hz, 4.8 Hz and 2 Hz), 4.01 (d, 1H, $J$ = 2 Hz), 7.49 (m, 2H), 7.61 (ddt, 1H, $J$ = 8 Hz, 6.8Hz, 1.2Hz), 8.01(m, 2H); $^{13}$C NMR (100 MHz, CDCl$_3$) $\delta$ 14.19, 22.73, 25.77, 31.71, 32.17, 57.62, 60.35, 128.48, 129.03, 134.04, 135.78, 194.95.

**/(E)-5-(Butylthio)-6-hydroxyundec-3-en-2-one (11).** A solution of butanethiol (45 mg, 0.5 mmol) and 1 (91 mg, 0.5 mmol) in acetonitrile (20 mL) was added to sodium phosphate buffer (50 mM, pH 7.4, 80 mL). The resulting solution was incubated at 37 °C for 72 hours, and then extracted with CH$_2$Cl$_2$ (3 x 100 mL). The combined organic layer was dried over anhydrous Na$_2$SO$_4$, and filtered. The filtrate was concentrated, and
purified by silica gel chromatography with hexanes/ethyl acetate (10:1) as eluent to provide 11. $^1$H NMR (400 MHz, CDCl$_3$) 0.78-0.88 (6H), 1.2-1.6 (12H), 2.24 (s, 3H), 2.38 (t, 2H, $J = 7.6$ Hz), 3.27 (dd, 1H, $J = 6$ Hz and 4 Hz), 3.82 (m, 1H), 5.96 (dd, 1H, $J = 15.6$ Hz), 6.67 (dd, 1H, $J = 16$ Hz and 10 Hz).

**E-5-(Butylthio)-6-hydroxyoct-3-en-2-one (12).** Butythiol (90 mg, 1 mmol) with 3 (142 mg, 1 mmol) was incubated at r.t. for 12 hours as the method described above to give compound 12 (150 mg, 65% yield). $^1$H NMR (400 MHz, CDCl$_3$) 0.87 (t, 3H, $J = 7.4$ Hz), 0.95 (t, 3H, $J = 7.4$ Hz), 1.32-1.39 (m, 2H), 1.47-1.54 (m, 2H), 2.27 (s, 3H), 2.42 (t, 2H, $J = 7.2$ Hz), 3.34 (dd, 1H, $J = 6$ Hz and 3.2 Hz), 3.7 (m, 1H), 6.00 (d, 1H, $J = 16$ Hz), 6.69 (dd, 1H, $J = 16$ Hz and 6 Hz); $^{13}$C NMR (100 MHz, CDCl$_3$) δ 13.03, 13.88, 22.14, 27.01, 28.02, 30.95, 31.67, 53.53, 74.24, 132.48, 143.45, 198.78.

**3-(Butylthio)-1-(3-pentyloxiran-2-yl)butan-1-one (15).** A solution of butanethiol (45 mg, 0.5mmol) and EKODE II model compound 2 in acetonitrile (20 mL) was added to sodium phosphate buffer (50 mM, pH 7.4, 80 mL). The resulting solution was stirred at room temperature for 1 hour, and then quickly extracted with CH$_2$Cl$_2$ (3 x 100 mL). The combined organic layer was dried over anhydrous Na$_2$SO$_4$, and filtered. The filtrate was concentrated, and purified by silica gel chromatography with hexanes/ethyl acetate (20:1) as eluent to almost quantitatively provide the Michael addition product 15 containing two isomers (syn:anti = 45:55). $^1$H NMR (400 MHz, CDCl$_3$) 0.82-0.88 (6H), 1.21(d, 3H, $J = 6.8$ Hz), 1.24-1.61 (12H), 2.31-2.49 (m, 1H), 2.47 (t, 2H, $J = 7.2$ Hz), 2.51-2.72 (m, 1H), 2.97-3.07(m, 1H), 3.15-3.23 (m, 2H).

**E-6-(Butylthio)-5-hydroxyundec-2-en-4-one (16).** Michael adduct 15 was dissolved in sodium phosphate buffer (50 mM, pH 7.4, 200 mL) containing 20%
acetonitrile, and stirred at room temperature for 2 days. Then the solution was extracted with CH$_2$Cl$_2$ (3 x 100 mL). The combined organic layer was dried over anhydrous Na$_2$SO$_4$, and filtered. The filtrate was concentrated, and purified by silica gel chromatography with hexanes/ethyl acetate (5:1) as eluent to give the epoxy ring opening product 16 (42 mg, 80% yield).$^1$H NMR (400MHz, CDCl$_3$) 0.86 (t, 3H, J = 6.8 Hz), 0.93 (t, 3H, J = 7.2 Hz), 1.21-1.33 (4H), 1.45 (m, 2H), 1.57 (m, 2H), 1.97 (dd, 3H, J = 7.2 Hz and 1.6 Hz), 2.6 (m, 2H), 2.94 (ddd, 1H, J = 10 Hz, 3.2 Hz and 3.2 Hz), 3.52 (d, 1H, J = 5.2 Hz), 4.48 (dd, 1H, J = 5.2 Hz and 3.2 Hz), 6.39 (dq, 1H, J = 15.6 Hz and 1.6 Hz), 7.08(dq, 1H, J = 15.6 Hz and 7.2 Hz); $^{13}$C NMR (100 MHz, CDCl$_3$) δ 13.91, 14.23, 18.87, 22.26, 22.65, 27.20, 29.32, 31.28, 31.73, 31.94, 49.56, 78.22, 127.06, 145.81, 199.17.

(E)-3-(Butylthio)-1-(3-pentyloxiran-2-yl)but-2-en-1-one (18). Compound 7 was incubated with butanethiol by the method described above to quantitatively give 18. $^1$H NMR (400 MHz, CDCl$_3$) δ 0.82-0.88 (6H), 1.28 - 1.61 (8H), 2.26 (s, 3H), 2.84 (t, 2H, J = 7.2 Hz), 2.98 (m, 1H), 3.19 (d, 1H, J = 1.6 Hz), 6.25 (s, 1H); $^{13}$C NMR (100 MHz, CDCl$_3$) δ 13.87, 14.18, 22.132, 22.74, 24.67, 25.74, 30.73, 31.27, 31.66, 32.13, 59.08, 59.90, 114.12, 164.14, 193.99.

5-(Butylthio)-6-hydroxyundecan-4-one (19). Compound 8 was incubated with butanethiol by the method described above to give 19 (40 mg, 76%). $^1$H NMR (400 MHz, CDCl$_3$) δ 0.83 (t, 3H, J = 7.2 Hz), 0.87 (t, 3H, J = 7.6 Hz), 1.28 - 1.42 (11H), 1.57 (m, 2H), 1.85 (m, 1H), 2.34 (m, 1H), 2.43 (m, 1H), 2.46 (m, 1H), 2.51 (m, 1H), 2.70 (m, 1H), 3.10 (d, 1H, J = 8.8 Hz), 3.91 (m, 1H); $^{13}$C NMR (100 MHz, CDCl$_3$) δ 13.84, 13.91, 14.28, 17.72, 22.16, 22.85, 25.38, 30.05, 31.71, 31.94, 33.87, 42.94, 57.54, 70.89, 208.24.
2-(Butylthio)-3-hydroxy-1-phenyloctan-1-one (20). Compound 10 was incubated with butanethiol, and the resulting product was purified as described above to afford a pale yellow solid 9 (37 mg, 60% yield). $^1$H NMR (400 MHz, CDCl$_3$) δ 0.83 (t, 3H, $J = 7.4$ Hz), 0.89 (t, 3H, $J = 7$ Hz), 1.2-1.7 (12H), 2.02 (m, 1H), 2.38 (m, 1H), 2.53 (m, 1H), 4.00 (d, 1H, $J = 8.8$ Hz), 4.19 (ddd, 1H, $J = 8.8$ Hz, 8.4 Hz and 2.8 Hz), 7.47 (m, 2H), 7.58 (ddt, 1H, $J = 8.4$ Hz, 8 Hz and 1.2 Hz), 7.99 (m, 2H); $^{13}$C NMR (100 MHz, CDCl$_3$) δ 13.79, 14.32, 22.14, 22.88, 25.57, 29.57, 31.54, 31.99, 33.723, 51.68, 71.18, 128.77, 128.84, 133.58, 136.42, 197.06.

Kinetic studies of the reactions of EKODE model compounds with butanethiol or N-acetyl cysteine methyl ester. A mixture of EKODE model compound 1 or 2 with butanethiol or N-acetyl cysteine methyl ester was incubated in sodium phosphate buffer (50 mM pH 7.4) containing 20% acetonitrile at 37 °C. The changes of the absorbance at 232 nm were used to monitor the reactions. The spectrohotometric data were recorded on LAMBDA 25 instruments (Perkin Wlmer, Shelton, CT). The reactions were monitored by repetitive spectral scanning from 200-300 nm. 10 mm path length quartz cuvettes were used for the low concentration reactions (the concentration of EKODE model compounds was 0.1 mM); while 1 mm path length quartz cuvettes were used for higher concentration reactions (the concentration of EKODE model compounds was 1 mM).

Incubation of EKODE II model compound 2 with imidazole. 1 mM 2 and 3 mM imidazole were incubated in phosphate buffer (50 mM, pH 7.4, 200 mL) containing 20% CD$_3$CN for 3 hours. The reaction was directly monitored by 600 MHz NMR without separation of products. The water peak was sequestered by pre-saturation and solvent substraction. The proton NMR spectrum showed that, besides imidazole, the major
compounds in the reaction solution are Michael adducts of 2 with imidazole, suggesting a mixture of diastereomers A and B.

A. $^1$H NMR (400 MHz, CDCl$_3$) δ 0.75-0.79 (3H), 1.20-1.29 (6H), 1.43 (d, 3H, $J = 6.6$ Hz), 1.49 (m, 2H), 1.95 (s, 3H), 2.89 (m, 1H), 2.93 (dd, 1H, $J = 19.2$ Hz and 4.2 Hz), 3.20 (dd, 1H, $J = 18$ Hz and 9 Hz), 3.50 (s, 1H).

B. $^1$H NMR (400 MHz, CDCl$_3$) δ 0.75-0.79 (3H), 1.20-1.29 (6H), 1.43 (d, 3H, $J = 6.6$ Hz), 1.49 (m, 2H), 1.95 (s, 3H), 2.74 (m, 1H), 2.93 (dd, 1H, $J = 19.2$ Hz and 4.2 Hz), 3.05 (dd, 1H, $J = 18$ Hz and 9 Hz), 3.41 (s, 1H).

4.5. References

1. Frankel, E. N., Chemistry of free radical and singlet oxidation of lipids. Prog Lipid Res 1984, 23 (4), 197-221.


3. (a) Butterfield, D. A.; Bader Lange, M. L.; Sultana, R., Involvements of the lipid peroxidation product, HNE, in the pathogenesis and progression of Alzheimer's disease. Biochim Biophys Acta 2010, 1801 (8), 924-9; (b) Mattson, M. P., Roles of the lipid peroxidation product 4-hydroxynonenal in obesity, the metabolic syndrome, and associated vascular and neurodegenerative disorders. Exp Gerontol 2009, 44 (10), 625-33.


Appendix

Preparation of the antigen to KODA-derived ketoamide (see Scheme 4 in Chapter 1 page 15).

**9-Oxodecanoic acid.** A mixture of cyclohexanone (9.8 g, 100 mmol), 30% H$_2$O$_2$ (5.5 mL, 50 mmol), perchloric acid (70 %, 0.35 mL) and water (10 mL) was added dropwise to a solution of 12% TiCl$_3$ (140 mL, 100mol), methyl vinyl ketone (15 mL, 180 mmol) and acetone (150 mL) at -78 °C. The resulting mixture was stirred at r.t. for 3 h, and then extracted with toluene (5 x 200 mL). The combined extracts were concentrated by evaporation of solvents to 200 mL, and washed with 10% NaOH (4 x 50 mL). The aqueous extracts were combined, acidified with 12N HCl, and then extracted with toluene (5 x 200 mL). The combined organic layers were dried over anhydrous Na$_2$SO$_4$, and filtered. The solvents were removed from the filtrate by evaporation to give 9-oxodecanoic acid (3.25g, 35% yield).$^1$ $^1$H NMR (400 MHz, CDCl$_3$) δ 1.24-1.36 (6H), 1.55-1.66 (4H), 2.14 (s, 3H), 2.35 (t, 2H, $J$ = 7.6 Hz), 2.42 (t, 2H, $J$ = 7.4 Hz).

**Benzyl 9-oxodecanoate.** Benzyl chloroformate (3 g, 18 mmol) was added to a mixture of 9-oxodecanoic acid (2.8 g, 15 mmol), triethylamine (2 g, 20 mmol) and DMAP (200 mg, 1.6 mmol) in CH$_2$Cl$_2$. The resulting solution was stirred at r.t. for 2 hour, and then concentrated by evaporation of solvents. The residue was purified by silica gel chromatography with hexanes/ethyl acetate (5:1) as eluent to give benzyl 9-oxodecanoate (3 g, 70% yield).

**11-((Benzyloxy)carbonyl)-4-oxoundecanoic acid.** Bromine (0.8 g, 5 mmol) was added to a solution of benzyl 9-oxodecanoate (1.4 g, 5 mmol) in benzyl alcohol (5 mL). The resulting solution was stirred at r.t. until the red color of the solution disappeared.
Then benzyl alcohol was removed from the solution by silica gel chromatography with hexanes/ethyl acetate (5:1) as eluent to give a mixture of benzyl 10-bromo-9-oxodecanoate and benzyl 8-bromo-9-oxodecanoate. The mixture was dissolved in DMF (20 mL). Then Meldrum’s acid (1.44 g, 10 mol), anhydrous NaOAc (820 mg, 10 mmol) and acetic acid (600 mg, 10 mmol) were added. The resulting solution was stirred at r.t. for 2 h, and then water (80 mL) was added. The solution was acidified with 1 N HCl, and extracted with ethyl acetate (3 x 50 mL). The combined organic layers were washed with water (2 x 30 mL), and solvents were evaporated. Then the residue was dissolved in DMF/water (9:1, 10 mL). The resulting solution was refluxed for 1 h, and then the solvents were evaporated. The residue was purified by silica gel chromatography with hexanes/ethyl acetate/acetic acid (5:1:0.1) as eluent to give 11-((benzyloxy)carbonyl)-4-oxoundecanoic acid (250 mg, 15% yield). ¹H NMR (400 MHz, CDCl₃) δ 1.24-1.36 (6H), 1.54-1.66 (4H), 2.34 (t, 2H, J = 7.2 Hz), 2.42 (t, 2H, J = 7.4 Hz), 2.62 (m, 2H), 2.70 (m, 2H), 5.11 (s, 2H), 7.35 (m, 5H); ¹³C NMR (100 MHz, CDCl₃) δ 23.87, 25.07, 27.97, 29.12, 29.14, 29.19, 34.89, 36.99, 42.86, 66.35, 128.42, 128.78, 136.30, 173.96, 178.69, 209.17.

4-Oxododecanedioic acid 1-(2,5-dioxo-1-pyrrolidinyl) ester. N-Hydroxy succinimide (30 mg, 0.26 mmol) and DCC (60 mg, 0.3 mmol) were added to a solution of 11-((benzyloxy)carbonyl)-4-oxoundecanoic acid (100 mg, 0.3 mmol) in CH₂Cl₂ (10 mL). The resulting mixture was stirred at r.t. for 2 h, and then the solvent was evaporated. The residue was dissolved in methanol, and Pd/C (10%, 10 mg) was added. The resulting mixture was shaked at r.t. under H₂ (30 psi) for 4 h, and then filtered. The filtrate was concentrated, and purified by silica gel chromatography with hexanes/ethyl acetate/acetic
acid (4:1:0.1) as eluent to give 4-oxododecanedioic acid 1-(2,5-dioxo-1-pyrrolidinyl)
ester (48 mg, 55% yield). $^1$H NMR (400 MHz, CDCl$_3$) $\delta$ 1.21-1.39 (6H), 1.51-1.66 (4H),
2.33 (m, 2H), 2.44 (m, 2H), 2.75-2.92 (8H); $^{13}$C NMR (100 MHz, CDCl$_3$) $\delta$ 23.83, 24.79,
25.26, 25.79, 29.01, 29.07, 29.13, 29.93, 34.18, 36.86, 42.75, 66.35, 168.51, 169.30,
195.77, 207.65.

**Modification of bovine serum albumin (BSA) or keyhole limpet hemocyanin**

(KLH) with 4-oxododecanedioic acid 1-(2,5-dioxo-1-pyrrolidinyl) ester. 4-
Oxododecanedioic acid 1-(2,5-dioxo-1-pyrrolidinyl) ester (20 mg, 0.06 mmol) was added
to BSA or KLH solution in DMF/PBS (1:4, 5 mL, 5 mg/mL). The resulting solution was
shaken at r.t. for 4 h, and then dialyzed against NH$_4$Cl buffer/methanol (3:1, 50 mM, pH
6.0, 4 x 0.5 L) for 24 h, and sodium phosphate buffer (50 mM, pH 7.4, 4 x 0.5 L) for
another 24 h. Then the protein solutions was transferred to tubes, and stored at -80 °C.

**Reference**

Bibliography


65. Mikulikova, K.; Eckhardt, A.; Pataridis, S.; Miksik, I., Study of posttranslational non-enzymatic modifications of collagen using capillary electrophoresis/mass


127. Xu, G.; Liu, Y.; Sayre, L. M., Polyclonal antibodies to a fluorescent 4-hydroxy-2-nonenal (HNE)-derived lysine-lysine cross-link: characterization and application to HNE-


