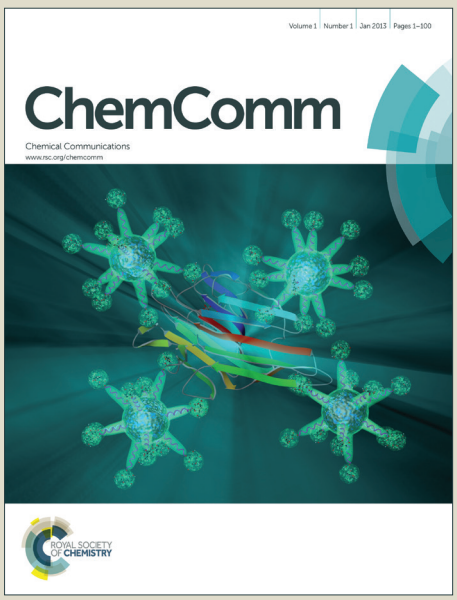
ChemComm

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

Journal Name

RSCPublishing

COMMUNICATION

Site-Specific Cross-linking of Collagen Peptides by Lysyl Advanced Glycation Endproducts

Cite this: DOI: 10.1039/xoxxoooooox

M. Kamalov,[a,b] P. W. R. Harris,[a] G. J. S. Cooper[b,c,d] and M. A. Brimble*[a,b,c]

Received ooth January 2012,
Accepted ooth January 2012

DOI: 10.1039/xoxxoooooox

www.rsc.org/

Cross-linking of proteins by advanced glycation endproducts (AGEs) causes a host of pathological conditions but their exact roles are unknown. Cross-linking lysyl AGEs were synthesized and incorporated into two types of collagen peptides. The utility of these cross-linked peptides for biochemical investigations was demonstrated by proteolysis studies and circular dichroism.

When in contact with sugars and sugar-degradation products, proteins can be irreversibly modified, giving rise to a diverse family of amino acid derivatives collectively known as advanced glycation endproducts (AGEs). Tissue accumulation of AGEs *in vivo* is continuous throughout a person's lifetime and is implicated in the pathogenesis of an array of chronic, age-related conditions, such as diabetes mellitus, cardiovascular, and Alzheimer diseases.[1] However, it has not been determined whether or not AGEs directly cause these detrimental effects or are biomarkers of other poorly understood underlying processes.[2]

AGEs also form during food preparation[3] and while some AGEs can be absorbed into the circulation, how the gastrointestinal tract handles AGE-proteins is still unknown.[4] The negative impact of AGEs on enzymatic proteolysis of the host proteins has also been described,[5] however glycation in such studies was effected by non-specific incubation of a protein in the presence of glucose. As a result, the impact of specific AGEs on proteolysis is poorly understood. A key factor preventing in-depth studies of AGEs is the lack of molecular probes, such as peptides site-specifically glycated by AGEs.[6]

Glyoxal lysine dimer (GOLD) and methylglyoxal lysine dimer (MOLD) are cross-linking lysyl AGEs that have been shown to accumulate in lens crystallins and in collagen with diabetes and ageing.[7] In uraemia, levels of GOLD and MOLD are significantly elevated compared to age-matched controls but,[8] as with other AGEs, their exact roles are uncertain. MOLD was first characterized by Brinkmann *et al.* from reaction between methyl glyoxal and N^{α}-hippuryllysine at pH 7.4, where MOLD formed as a single major product in only 15% yield.[9] Subsequent syntheses of GOLD and MOLD were similarly carried out by incubating lysine, or an α-protected lysine derivative, with either glyoxal or methylglyoxal for prolonged periods.[10] These syntheses required tedious purifications by HPLC hence the low yields. Linetsky and Shipova developed a new approach via alkylation of imidazole or methylimidazole with 5-(4-bromobutyl)hydantoin followed by hydrolysis.[11] Recently, Esposito *et. al.* synthesized MOLD by reacting N^{α}-Boc-lysine with formaldehyde, methyl glyoxal, and acetic acid for 1 h, however the final yield of the purified product was not reported.[12]

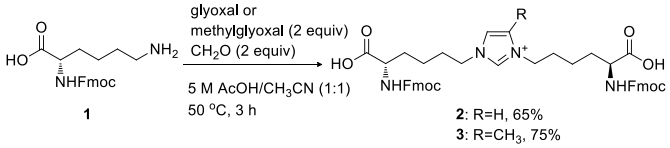
Several groups have incorporated non-AGE crosslinks into peptides. For example, Hutton *et al.* have successfully synthesised amyloid-β peptides containing the non-AGE crosslink diaminopimelic acid.[13] The coupling of each crosslink with the resin-bound peptide requires special optimisation, where factors such as resin loading and choice of the coupling agents play important roles. Incorporation of the cross-linked AGEs, GOLD and MOLD into peptides has only been reported by Yamada *et al.*, where only small quantities of the cross-linked peptide were obtained after incubation of the monomeric peptides with glucose or glyoxal for up to 7 months.[14] This approach is highly inefficient and does not allow for the targeted introduction of cross-links at specific sites within peptides.

Herein, we report efficient syntheses of GOLD and MOLD as Fmoc-protected building blocks suitable for solid phase peptide synthesis (SPPS). Optimization for coupling these building blocks to the resin-bound peptides was also achieved. Importantly, these two cross-linked AGEs were then successfully incorporated into specific sites of two types of collagen peptides. To our knowledge this is the first example of a cross-linked AGE building block being introduced into a peptide sequence using SPPS. Furthermore, the impact of GOLD and MOLD on the 3D structures of the host peptides and on their susceptibility to proteolysis has been investigated. These

methods have wide application for the future study of AGEs and their roles in the pathogenesis of ageing-related diseases.

Fmoc based SPPS was adopted in the present work due to its widespread use[15] hence we initially focused on the synthesis of $Fmoc_2$GOLD **2** and $Fmoc_2$MOLD **3** (Scheme 1). N^{α}-Fmoc-lysine, **1**, (2 equiv.) was reacted with glyoxal (1 equiv) and formaldehyde (1 equiv) in a mixture of 5 M AcOH - acetonitrile (1:1) at 50 °C for 3 h to afford **2** in 65% yield (see SI for details).

In an analogous fashion, MOLD derivative **3** was cleanly generated in 75% yield on a gram scale by a reaction of **1** (1 equiv) with methylglyoxal (2 equiv) and formaldehyde (2 equiv).



Scheme 1. Synthesis of GOLD and MOLD building blocks

With building blocks **2** and **3** in hand, our attention turned to their incorporation into peptides. Collagen comprises two sets of tertiary structures; the triple helical fibrous region and the non-helical telopeptide region, the latter of which is commonly cross-linked.[16] We therefore aimed to incorporate the cross-linking AGEs, GOLD and MOLD, into peptides that mimic both forms of collagen: the collagen telopeptides (CTPs) and triple-helical collagen model peptides (CMPs). The sequence used as a model for a CTP (YGYDEKSTGGISVP) was sourced from the telopeptide region of human type I α1 collagen and was chosen because it contained a lysyl residue in position 6.[17]

The native CTP sequence was synthesized by standard automated Fmoc SPPS (see SI for details). Standard couplings were carried out by Fmoc SPPS on polystyrene aminomethyl resin (0.9 mmol/g loading) up to the serine residue in position 8, which was then Fmoc-deprotected. The resulting resin-bound peptide **4** (Scheme 2) was used to evaluate several coupling conditions to incorporate $Fmoc_2$GOLD **2**. A 2:1 ratio of coupling reagent and **2** was employed recognizing the dicarboxylic acid in **2**. Following coupling with **4**, the Fmoc-groups were deprotected with piperidine and acylated with the next amino acid in the sequence, Fmoc-Glu(O^tBu)-OH. Small aliquots of the resultant intermediate were subsequently cleaved and analyzed for the presence of cross-linked peptides by LC-MS (see Table 1 and Figure SI-3).

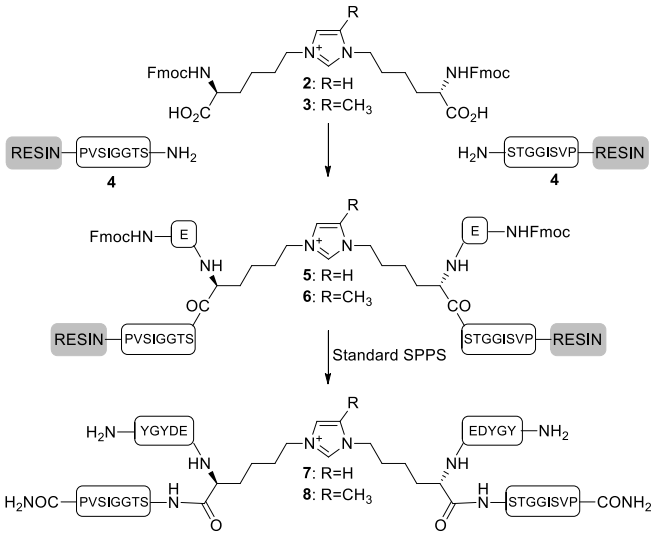
Table 1. Coupling of **2** with the CTP peptide **4**

Entry	Coupling Agent	Resin Loading [mmol/g]	Reagent Ratio[a]	Yield of **5**[b] [%]
1	HOAt/DIC	0.9	2:1:2	3
2	Oxyma/DIC	0.9	2:1:2	4
3	PyBOP	0.9	2:1:2	24
4	PyBOP	0.35	2:1:2	25
5	PyBOP	0.09	2:1:2	29
6	PyBOP	0.09	2:1:1	37
7	HBTU	0.9	2:1:2	20
8	HATU	0.9	2:1:2	30
9	HATU	0.09	2:1:1	37
10	HATU	0.09	2.2:1:2	90

[a] Coupling agent:**2**:**4**, calculation based on the resin loading and assuming 100% conversion in previous couplings. [b] Measured by HPLC at 214 nm (Figure SI-3)

Use of diisopropylcarbodiimide (DIC) as a coupling agent resulted in minimal addition of **2** onto **4**, with deletion product (Fmoc-ESTGGISVP-NH_2) accounting for 64% of the product mixtures when either HOAt or Oxyma were employed as additives (Table 1, entries 1 and 2). Uronium-based coupling agents afforded better incorporation of the cross-linking amino acid **2**, 20% with HBTU (entry 7) and 30% with HATU (entry 8). However, use of the phosphonium coupling agent, PyBOP, gave the cleanest conversion of **4** to **5** in 24% yield (entry 3). Use of a lower resin loading of peptide **4** (0.35 mmol/g and 0.09 mmol/g) was next investigated. Coupling **2** to resin-bound peptide **4**, in this case using PyBOP, afforded **5** in better yield (entries 4 and 5). Resin loading of 0.09 mmol/g was optimal (29%). Peptide **4** with the lowest resin loading was next subjected to treatment with excess of the coupling mixture, affording **5** in 37% yield (entry 6). The analogous coupling with HATU also afforded **5** in 37% yield (entry 9). Finally, coupling with 2.2 equivalents of HATU relative to the building block **2** with low resin loading afforded **5** in 90% yield (entry 10)**.** The use of a small excess of an effective coupling agent, namely HATU, in conjunction with low resin loading was key to efficient incorporation of the cross-linking building block onto the resin-bound peptide.

The resulting dimeric peptide **5** was subjected to elongation of the peptide sequence via standard SPPS, cleaved and purified by semi-preparative HPLC to give the desired homodimeric peptide **7** with >98% purity (Figure SI-6).



Scheme 2. Synthesis of homodimeric CTPs **9** and **10** incorporating GOLD and MOLD.

The reaction conditions optimized for coupling of **2** were next used for incorporating **3** into the same peptide sequence. HATU (2.2 equiv) and **3** (1.0 equiv.) were reacted with peptide **4** (2.0 equiv.) attached to low-loading aminomethyl resin. The resin bound peptide **6** was then subjected to further standard SPPS and purified by HPLC. The target peptide **8** was obtained in >98% purity (Figure SI-7).

The telopeptide domain of collagen is a common target of proteolytic enzymes during both tissue remodeling and digestion.[18] The same region is also a common cross-linking and glycation site: hence it is important to investigate the effects of AGEs on collagen proteolysis. To demonstrate the utility of the CTPs cross-linked with

GOLD and MOLD in studies of peptide proteolysis, **7** and **8** together with the native CTP peptide were subjected to digestion by bovine trypsin. A solution of the protease was incubated at 37 °C in the presence of native CTP. Aliquots were removed every minute, quenched with 1 M HCl, and analyzed by HPLC for relative peptide concentrations. Analogous incubations were performed with peptides **7** and **8,** and the results are given in Figure SI-10. Under the experimental conditions, trypsin digested 90% of the native CTP peptide in 10 minutes (Figure SI-10A, graph i; Figure SI-10B), while the HPLC profiles of **7** (Figure SI-10A, graph ii) and **8** (Figure SI-10A, graph iii) and their relative concentrations (Figure SI-10B) remained unchanged. This result demonstrates the dramatic effect that the presence of GOLD or MOLD cross-links has on host protein proteolysis, specifically by trypsin as representative of the proteases present in the digestive tract.

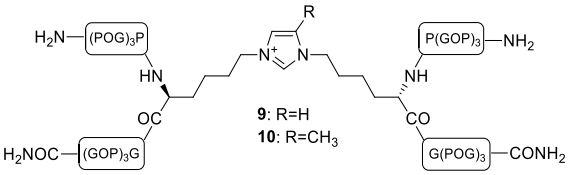


Figure 1. Structures of CMPs with cross-linking AGEs, where O denotes 4R-hydroxyproline residues.

Having developed conditions for successful coupling of GOLD and MOLD with a CTP, our attention turned to incorporating these cross-linking structures into CMPs, which are model peptides for investigating the triple-helical structure specific to collagen. CMPs cross-linked with GOLD **9** or MOLD **10** (Figure 1) were prepared according to procedure described above for CTPs in 17% yield (**9**) and >98% purity (Figure SI-8), and 15% yield (**10**) and >98% purity after HPLC (Figure SI-9).

The key feature of CMPs is their propensity to form triple helices in solution,[19] rendering them suitable models to study the structure and function of the native collagenous triple helices on which they are modelled. The CD profiles of **9** and **10** along with the peptide Ac-(Pro-Hyp-Gly)$_3$-Pro-Lys-Gly-(Pro-Hyp-Gly)$_3$–NH$_2$ (CMP-K) were analyzed in the range 190-260 nm at 20 °C (Figure SI-11). The CD spectrum of CMP-K displayed a maximum at 222 nm and a minimum at 198 nm suggesting that the peptide formed a triple helix. By comparison, the maxima and minima in the CD spectra of **9** and **10** are shifted with small maxima at 225 nm and minima at 202 nm suggesting that the cross-linked peptides formed random coils. This result is consistent with the dimeric structure of peptides **9** and **10**, where GOLD and MOLD covalently hold the two chains together preventing the formation of triple helices.

The efficient synthesis of these useful cross-linking GOLD- and MOLD-amino acid building blocks when combined with standard SPPS methods, has enabled a detailed analysis of the effects of GOLD- and MOLD-mediated cross-linking on the properties of collagenous peptides with precisely specified sequences. This important advance enables the research community to develop an improved understanding of the roles played by these cross-links in the pathophysiology of collagen as it impacts the pathogenesis of ageing-related diseases.

In conclusion, we have developed efficient syntheses of two cross-linking AGEs, GOLD and MOLD, and incorporated them into two sets of collagen peptides. The utility of these cross-linked peptides to probe AGE biochemistry has been demonstrated by a trypsin digest study of cross-linked CTPs and by circular dichroism studies of the tertiary structures of cross-linked CMPs. Investigation of the precise metabolic fate of lysyl cross-linking AGEs in suitable animal models, as well as analysis of their relationship to human proteases that target collagen is currently underway.

Notes and references

[a] School of Chemical Sciences, The University of Auckland, 23 Symonds Street, Auckland, New Zealand
[b] Maurice Wilkins Centre for Molecular Biodiscovery, The University of Auckland, 3 Symonds Street, Auckland, New Zealand
[c] School of Biological Sciences, The University of Auckland, 3 Symonds Street, Auckland, New Zealand
[d] Centre for Advanced Discovery and Experimental Therapeutics, Central Manchester University Hospitals, NHS Foundation Trust, and the School of Biomedicine, The University of Manchester, Manchester UK, and Department of Pharmacology, Division of Medical Sciences, University of Oxford, Oxford, UK

Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/c000000x/

1 H. Vlassara and G. E. Striker, *Nat. Rev. Endocrinol.*, 2011, **7**, 526–539.
2 J. W. Baynes, *Exp. Gerontol.*, 2001, **36**, 1527–1537.
3 W. Cai, J. Uribarri, L. Zhu, X. Chen, S. Swamy, Z. Zhao, F. Grosjean, C. Simonaro, G. A. Kuchel, and M. Schnaider-Beeri, *Proc. Natl. Acad. Sci.*, 2014, 201316013.
4 J. M. Ames, *Mol. Nutr. Food Res.*, 2007, **51**, 1085–1090.
5 M. Brownlee, A. Cerami, and H. Vlassara, *N. Engl. J. Med.*, 1988, **318**, 1315.
6 T. M. Woods, M. Kamalov, P. W. Harris, G. J. Cooper, and M. Brimble, *Org. Lett.*, 2012, **14**, 5740–5743.
7 P. Chellan and R. H. Nagaraj, *Arch. Biochem. Biophys.*, 1999, **368**, 98–104.
8 H. Odani, T. Shinzato, J. Usami, Y. Matsumoto, E. Brinkmann Frye, J. W. Baynes, and K. Maeda, *FEBS Lett.*, 1998, **427**, 381–385.
9 E. Brinkmann, K. J. Wells-Knecht, S. R. Thorpe, and J. W. Baynes, *J. Chem. Soc. Perkin Trans. 1*, 1995, 2817–2818.
10 K. J. Wells-Knecht, E. Brinkmann, and J. W. Baynes, *J. Org. Chem.*, 1995, **60**, 6246–6247.
11 M. D. Linetsky and E. V. Shipova, *Amino Acids*, 2007, **32**, 285–289.
12 D. Esposito, S. Kirchhecker, and M. Antonietti, *Chem. – Eur. J.*, 2013, **19**, 15097–15100.
13 W. M. Kok, D. B. Scanlon, J. A. Karas, L. A. Miles, D. J. Tew, M. W. Parker, K. J. Barnham, and C. A. Hutton, *Chem. Commun.*, 2009, 6228–6230.
14 H. Yamada, T. Sasaki, S. Niwa, T. Oishi, M. Murata, T. Kawakami, and S. Aimoto, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 5677–5680.
15 L. A. Carpino and G. Y. Han, *J. Org. Chem.*, 1972, **37**, 3404–3409.
16 S. Viguet-Carrin, J. P. Roux, M. E. Arlot, Z. Merabet, D. J. Leeming, I. Byrjalsen, P. D. Delmas, and M. L. Bouxsein, *Bone*, 2006, **39**, 1073–1079.
17 M.-L. Chu, W. de Wet, M. Bernard, J.-F. Ding, M. Morabito, J. Myers, C. Williams, and F. Ramirez, *Nature*, 1984, **310**, 337–340.
18 M. P. Drake, P. F. Davison, S. Bump, and F. O. Schmitt, *Biochemistry (Mosc.)*, 1966, **5**, 301–312.
19 R. S. Erdmann and H. Wennemers, *J. Am. Chem. Soc.*, 2010, **132**, 13957–13959.