# Expedient Synthesis of Peptides Containing $N^{\varepsilon}$-Carboxymethyllysine

Meder Kamalov,[a,b] Sunghyun Yang,[a] Paul W. R. Harris,[a] Garth J. S. Cooper,[b,c,d,e,f] Margaret A. Brimble*[a,b,c]

[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
Fax +64(9)3737422; E-mail: m.brimble@auckland.ac.nz

[d] Centre for Advanced Discovery and Experimental Therapeutics (CADET), Central Manchester University Hospitals, NHS Foundation Trust, Oxford Road, Manchester, M13 9WL, UK

[e] Institute of Inflammation and Repair, The University of Manchester, Oxford Road, Manchester, M13 9WL, UK

[f] Department of Pharmacology, Division of Medical Sciences, University of Oxford, Mansfield Road, Oxford, OX1 3QT, UK

*Received: 11.03.2014; Accepted: 09.05.2014*

**Abstract:** Accumulation of advanced glycation endproducts (AGEs) is responsible for the development and progress of diabetes- and age-related complications. Synthesis of specific chemical probes is key for the detailed understanding of biochemical properties of AGEs and their precise roles in the progression of disease. We herein report the expedient synthesis of such probes in the form of peptides site-specifically glycated by the major lysyl AGE, $N^{\varepsilon}$-carboxymethyllysine (CML). The facile and economical incorporation of CML into peptide sequences by using the nosyl group has been achieved in a single step on resin. This new method is a substantial improvement over the existing syntheses of CML-containing peptides in that it does not require the use of expensive reagents or elaborate purification techniques. The impact of CML on the proteolytic stability of the host peptide has been investigated using trypsin digest studies.

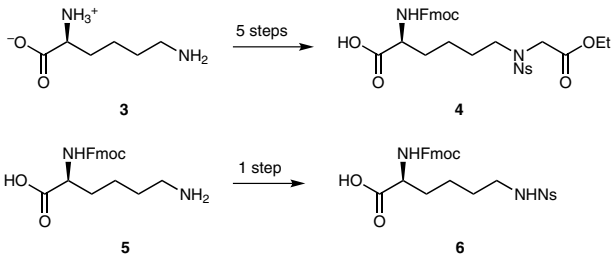
**Key words:** peptides, solid-phase synthesis, protecting groups, advanced glycation endproducts, $N^{\varepsilon}$-carboxymethyllysine

Advanced glycation endproducts (AGEs) are a structurally diverse family of modified amino acids formed via non-enzymatic reaction between sugars and proteins.[1] $N^{\varepsilon}$-Carboxymethyllysine (CML), a key member of the AGE family, is known to irreversibly accumulate in organ tissues over time,[2] and this accumulation is significantly accelerated in patients with diabetes mellitus, who also experience earlier onset of organ damage, such as cataract or heart disease.[3] Considerable interest has been aroused on the characterisation and use of CML as a biomarker for disease diagnosis and for evaluation of novel therapeutic strategies.[4]

Normal remodelling of the connective tissues requires enzymes to degrade fibrillar collagen. Several studies have highlighted the impact of AGE accumulation on the properties of collagen with regards to enzymatic proteolysis.[5–7] However, *in vitro* glycation in such studies is often achieved by incubating the protein in high-glucose solutions over prolonged periods. The resulting 'glycated' collagen shows lower levels of solubility and digestibility with trypsin, which is then assumed to indicate a connection between AGE accumulation and fibrosis observed in diabetic heart and kidney diseases. As a result, the precise impact of particular AGEs on the properties of the host proteins with regards to proteolysis is not well characterised. Therefore, peptides site-specifically glycated by AGEs serve as valuable probes for the investigation of AGE-impact on host protein proteolysis.

Currently, there are two strategies for incorporating CML into peptide sequences. One strategy uses a preformed building block and CML-containing peptides have been synthesised via this building block strategy both by us[8] and by others.[9] In this approach a suitably protected CML derivative is assembled via conventional organic synthesis and subsequently incorporated into peptides via autonomous solid-phase peptide synthesis (SPPS). This strategy suffers from poor step economy and low overall yield as three orthogonal protecting groups are required to facilitate the SPPS.

The alternative strategy for synthesis of CML-containing peptides relies on the site-specific alkylation of the resin-bound peptide containing a free lysyl ε-amine.[10] However, this procedure has not been documented fully and yields for CML-peptides made by this approach have not been reported.




**Scheme 1** Building blocks for CML incorporation into peptides

This work reports a straightforward and cost-effective synthetic procedure to access CML-containing peptides. Our method uses the 2-nitrobenzenesulfonyl (Ns) protect-

*SYNLETT* 2014, 25, 1835–1838
Advanced online publication: 06.06.2014
DOI: 10.1055/s-0033-1378305; Art ID: st-2014-d0234-l

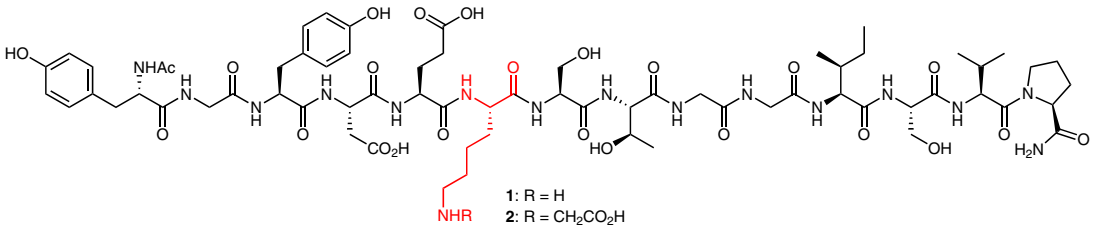
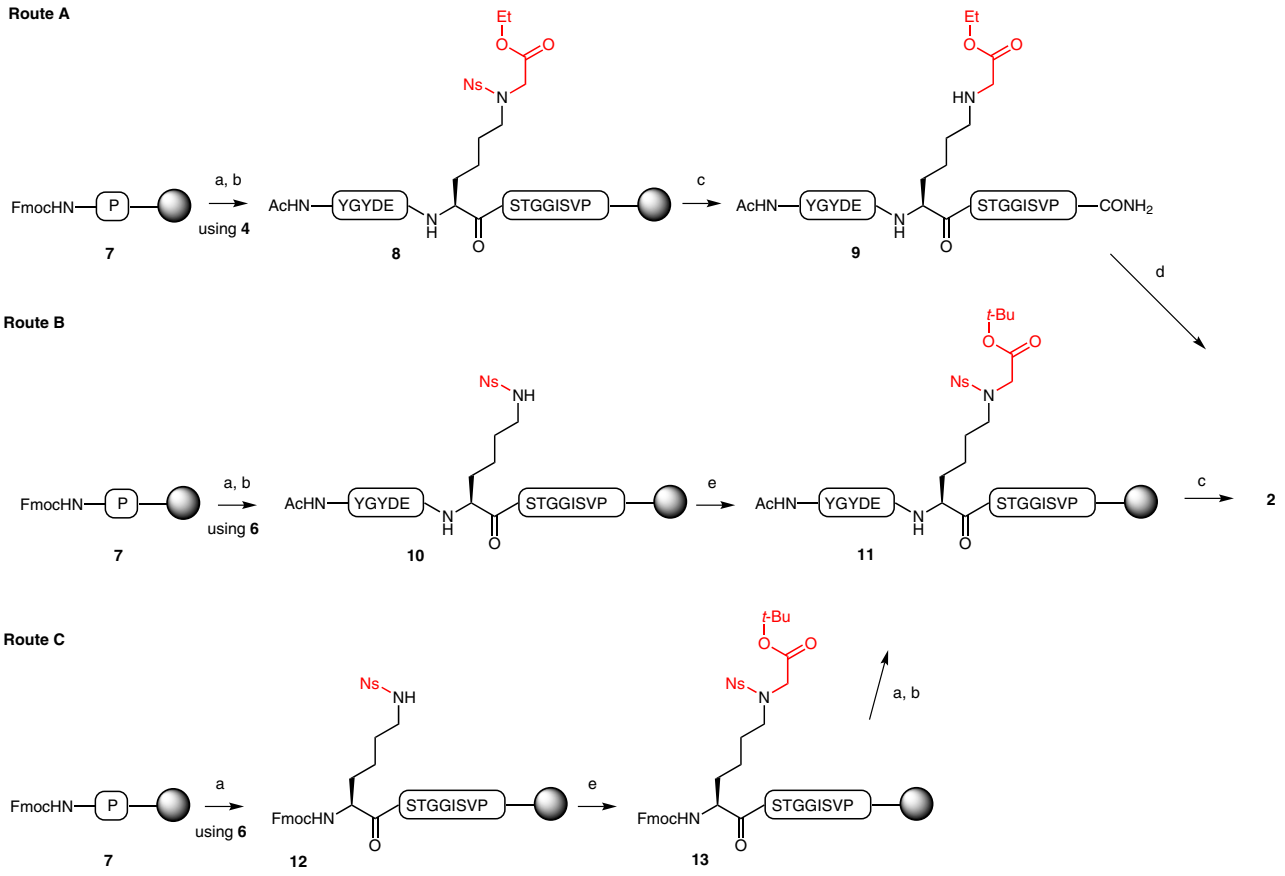



**Figure 1** Structures of the target peptides **1** and **2**

ing group for the lysyl ε-amine during SPPS as it is orthogonal to the conditions used for 9-fluorenylmethyloxycarbonyl (Fmoc) solid-phase chemistry.[11,12] Importantly, the Ns group also allows specific monoalkylation of the ε-amine.[13] We demonstrate that synthesis of CML-containing peptides via Ns alkylation is facile and can be accomplished both during SPPS on Fmoc-protected intermediate peptide or after completion of SPPS on the full-length resin-bound peptide. We then show the impact of the incorporation of CML on the host peptide digestibility by the key proteolytic enzyme trypsin, and validate the capacity of this AGE to limit proteolysis.

In order to demonstrate the practical applicability of the new synthetic approach to append CML directly to peptides, a sequence (YGYDEKSTGGISVP) was chosen from the collagen telopeptide (CTP) region of human type I α1 collagen.[14] Such telopeptide regions of collagen are common glycation and proteolysis sites *in vivo*.[15,16] It was decided to compare the protease digestions of the native peptide **1** with the CML-peptide **2**, in which the lysine was substituted for CML (Figure 1).

Peptide **1** was first synthesised via automated microwave SPPS, using HBTU/DIPEA for the coupling steps and 20% piperidine for deprotection of the Fmoc groups.[17] CML building block **4** (Scheme 1) was synthesised in five




**Scheme 2** Synthesis of CML-containing peptide **2** by building block approach (A), on-resin alkylation post-synthesis (B), and on-resin alkylation mid-synthesis (C). *Reagents and conditions:* (a) Fmoc-SPPS: (i) Fmoc removal: 20% piperidine in DMF; (ii) coupling: Fmoc-AA-OH, HBTU, DIPEA; (b) (i) *N*-terminal Fmoc removal: 20% piperidine in DMF; (ii) $Ac_2O$, DIPEA; (c) (i) Ns removal: 2-mercaptoethanol, DBU, DMF, r.t., 15 min; (ii) peptide release: TFA–TIS–$H_2O$, r.t., 2 h; (d) 0.1 M aq NaOH, 1 h; (e) *tert*-butyl bromoacetate, DIPEA, DMF, r.t., overnight.

steps starting from lysine **3** following the conditions we described earlier.[8]

This building block was then incorporated into the CML-containing peptide **2** via SPPS (Route A, Scheme 2), where resin-bound peptide **8** was synthesised following standard conditions from Fmoc-proline **7** attached to aminomethyl resin via a rink amide linker.[17] Cleavage of peptide **8** afforded peptide **9**, which underwent deesterification to give the target CML-containing peptide **2**. The HPLC profile of the crude peptide **2** synthesised via this traditional building block approach is shown in Figure SI-1A (see the Supporting Information).

Next, revised conditions for selective CML incorporation by alkylating the resin-bound peptide sequence were explored. First, Fmoc-Lys(Ns)-OH **6** (Scheme 1) was synthesised from commercially available Fmoc-Lys-OH **5** in a single step following conditions reported in literature.[18] **6** was incorporated into the peptide sequence via automated SPPS using analogous conditions as for peptides **1** and **2** (Route B, Scheme 2).[17]

Upon completion of peptide synthesis but prior to its cleavage, the *N*-terminus of the resin-bound peptide was acetylated with acetic anhydride, and the resulting peptide **10** was subjected to alkylation with *tert*-butyl bromoacetate in the presence of DIPEA to afford peptide **11**.[19] The nosyl group on peptide **11** was then smoothly deprotected in the presence of 2-mercaptoethanol and DIPEA (15 min),[20] and the resulting peptide cleaved with TFA–TIS–$H_2O$ (38:1:1). The crude peptide was purified by HPLC to afford **2** in 30% overall yield with >98% purity. The HPLC profile of the target peptide (Figure SI-1B) confirmed that it was identical to **2** prepared using the original building block approach (Figure SI-1A).

With a successful method for on-resin alkylation of complete peptide in hand, our attention turned to whether the Fmoc group was stable to the required alkylation conditions. Stability of the Fmoc group to such conditions would indicate whether CML could be incorporated efficiently during SPPS, as post-synthetic on-resin modification is not always feasible. This is especially the case if multiple modifications are necessary.

Peptide **12** was synthesised via standard SPPS (Route C, Scheme 2) on resin up to the point of attachment of **6**, and then alkylated with *tert*-butyl bromoacetate in the presence of DIPEA.[19] The resulting peptide **13** underwent further SPPS and the Ns group was removed[19] prior to TFA-mediated release of the peptide from the resin. The crude peptide was analysed by HPLC (Figure SI-1C), which showed a similar profile to that of the peptide obtained by post-synthetic alkylation (Figure SI-1B), or the building block approach (Figure SI-1A). Purification afforded the target peptide in 26% yield and >96% purity.

Examination of HPLC profiles of crude peptide **2** prepared by the traditional building block approach, alkylation post-synthesis, or alkylation mid-synthesis, demonstrates not only the reliability of this novel synthetic strategy but also its orthogonality to Fmoc/*t*-Bu peptide synthesis. Furthermore, comparable yields were obtained in all cases.

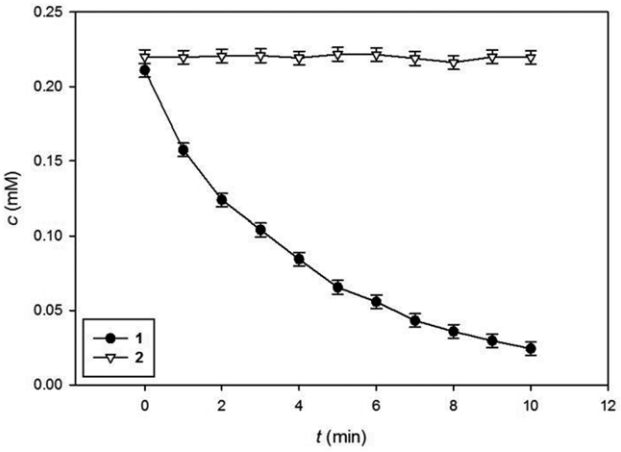



**Figure 2** Concentrations of CTP peptides **1** and **2** during the trypsin digest, extrapolated from HPLC

With CML-modified CTPs in hand, the impact of the presence of an AGE (CML) on the proteolytic digestibility of host peptide was investigated. Peptide **2** and the native peptide **1** were subjected to a bovine trypsin digest. Solutions of the protease were incubated in the presence of peptides and aliquots were removed, quenched, and analysed by HPLC.[21]

Under the experimental conditions, trypsin digested 90% of the native peptide **1** in 10 minutes ($t_{1/2}$ = 3 min), while the concentration of CML-modified peptide **2** remained unchanged (Figure 2). Because the presence of CML slows down the digestion of host peptides by trypsin, it is possible that this AGE is predominantly absorbed in peptide-bound form. Investigation of the subsequent metabolic transit of CML-modified peptides within a suitable animal model is currently underway.

A convenient and economical method for site-specific glycation of peptides by CML has been developed. This strategy does not require the use of expensive reagents or elaborate manipulation of protecting groups, and applicability of this method to standard Fmoc SPPS has been demonstrated. Trypsin digestion studies of both native and CML-modified peptides showed that CML dramatically influences the rate of enzymatic proteolysis of the host peptide.

## Acknowledgment

Authors thank the Maurice Wilkins Centre for Molecular Biodiscovery for generous financial support.

**Supporting Information** for this article is available online at http://www.thieme-connect.com/products/ejournals/journal/10.1055/s-00000083.

## References and Notes


(1) Kamalov, M.; Harris, P. W. R.; Cooper, G. J. S.; Brimble, M. A. *Chem. Commun.* **2014**, *50*, 4944.
(2) Dunn, J. A.; Patrick, J. S.; Thorpe, S. R.; Baynes, J. W. *Biochemistry (Moscow)* **1989**, *28*, 9464.
(3) Rabbani, N.; Thornalley, P. *Amino Acids* **2012**, *42*, 1087.
(4) Beisswenger, P. *Amino Acids* **2012**, *42*, 1171.
(5) Fu, M. X.; Requena, J. R.; Jenkins, A. J.; Lyons, T. J.; Baynes, J. W.; Thorpe, S. R. *J. Biol. Chem.* **1996**, *271*, 9982.
(6) Golub, L. M.; Greenwald, R. A.; Zebrowski, E. J.; Ramamurthy, N. S. *Biochim. Biophys. Acta* **1978**, *534*, 73.
(7) Hamlin, C. R.; Kohn, R. R.; Luschin, J. H. *Diabetes* **1975**, *24*, 902.
(8) Woods, T. M.; Kamalov, M.; Harris, P. W.; Cooper, G. J.; Brimble, M. *Org. Lett.* **2012**, *14*, 5740.
(9) Gruber, P.; Hofmann, T. *J. Pept. Res.* **2005**, *66*, 111.
(10) Xue, J.; Rai, V.; Singer, D.; Chabierski, S.; Xie, J.; Reverdatto, S.; Burz, D. S.; Schmidt, A. M.; Hoffmann, R.; Shekhtman, A. *Structure* **2011**, *19*, 722.
(11) Demmer, O.; Dijkgraaf, I.; Schottelius, M.; Wester, H.-J.; Kessler, H. *Org. Lett.* **2008**, *10*, 2015.
(12) Monfregola, L.; Leone, M.; Calce, E.; De Luca, S. *Org. Lett.* **2012**, *14*, 1664.
(13) Fukuyama, T.; Jow, C.-K.; Cheung, M. *Tetrahedron Lett.* **1995**, *36*, 6373.
(14) Chu, M.-L.; de Wet, W.; Bernard, M.; Ding, J.-F.; Morabito, M.; Myers, J.; Williams, C.; Ramirez, F. *Nature (London)* **1984**, *310*, 337.
(15) Woodley, D. T.; Yamauchi, M.; Wynn, K. C.; Mechanic, G.; Briggaman, R. A. *J. Invest. Dermatol.* **1991**, *97*, 580.
(16) Paul, R. G.; Bailey, A. J. *Int. J. Biochem. Cell Biol.* **1996**, *28*, 1297.
(17) **Peptide Synthesis**: SPPS was performed via the Fmoc strategy on Rink Amide resin using a Biotage Alstra peptide synthesiser on 0.1-mmol scale. The Fmoc group was deprotected with 20% piperidine in DMF for 2 min + 3 min at 60 °C. The coupling step was performed with Fmoc-AA-OH (5 equiv) in DMF (0.2 M), HBTU in DMF (4.5 equiv, 0.45 M) and DIPEA (10 equiv) for 5 min at 75 °C. The final Fmoc group was removed and the amine was acetylated by $Ac_2O$ in the presence of DIPEA at r.t. for 10 min. The peptides were released from resin with concomitant removal of the side-chain protecting groups by treatment with TFA–TIS–$H_2O$ (38:1:1, 5 mL) at r.t. for 2 h. Peptides were precipitated with cold $Et_2O$, isolated by centrifugation, washed in cold $Et_2O$, dissolved in MeCN–$H_2O$ (1:1) containing 0.1% TFA and lyophilised. The peptides were analysed for purity by LCMS using a Zorbax C3 column (3.5 μm; 3 × 150 mm; Agilent) at 0.3 mL/min using a linear gradient. The solvent system used was A (0.1% formic acid in $H_2O$) and B (0.1% formic acid in MeCN). Purification of crude peptides was performed by semipreparative HPLC using a Gemini C18 column (10 μm; 250 × 10 mm; Phenomenex) at 5 mL/min using a shallow linear gradient. The solvent system used was A (0.1% TFA in $H_2O$) and B (0.1% TFA in MeCN). The resulting purified peptides were analysed by the LCMS system used for crude peptide analysis. The purities were extrapolated from integrating the peaks corresponding to peptide **1** (13.1 min) and to peptide **2** (12.7 min).
(18) Halpin, D. R.; Lee, J. A.; Wrenn, S. J.; Harbury, P. B. *PLoS Biol.* **2004**, *2*, e175.
(19) **Alkylation Conditions**: Resin-bound peptide **10** or **12** (0.1 mmol, 1 equiv) was swollen in $CH_2Cl_2$ (30 min), then in DMF (10 min), and drained. A solution of *tert*-butyl bromoacetate (70 μL, 0.5 mmol, 5 equiv) with DIPEA (175 μL, 1 mmol, 10 equiv) in DMF (5 mL) was added in one portion and the reaction mixture was shaken overnight at r.t. to afford peptides **11** or **13**, respectively, which were drained and washed with DMF.
(20) **Ns Deprotection**: Resin-bound peptide **8** or **11** (0.1 mmol, 1 equiv) was swollen in $CH_2Cl_2$ (30 min), then in DMF (10 min) and filtered. A solution of 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) (90 μL, 0.6 mmol, 6 equiv) and 2-mercaptoethanol (42 μL, 0.6 mmol, 6 equiv) in DMF (5 mL) was added in one portion and the mixture was shaken for 15 min, drained, and washed with DMF..
(21) **Trypsin Digest**: Bovine trypsin (0.3 mg, type XI, 9090 units/mg, Sigma) was dissolved in $H_2O$ (1 mL) and 3.3 μL (9 units) of this solution diluted to 1 mL using Tris buffer (pH 8.0) and incubated at 37 °C for 30 min. Substrate peptide (0.21 μmol) was added in one portion and 50 μL aliquots removed every minute, quenched with 1 M HCl (50 μL) and analysed by analytical reverse phase-HPLC using a Luna C18(2) column (3μ; 150 × 3 mm; Phenomenex) at 0.3 mL/min using linear gradient. The concentrations were extrapolated from integrating the peaks corresponding to peptide **1** (14.9 min) and to peptide **2** (15.0 min).

