Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry Letters 14 (2004) 5677–5680
Bioorganic & Medicinal Chemistry Letters

Intact glycation end products containing carboxymethyl-lysine and glyoxal lysine dimer obtained from synthetic collagen model peptide

Hiroaki Yamada,[a] Tomoko Sasaki,[a] Sachiko Niwa,[a] Tohru Oishi,[a] Michio Murata,[a,*] Toru Kawakami[b] and Saburo Aimoto[b]

[a]*Department of Chemistry, Graduate School of Science, Osaka University, 1–16 Machinakeyama, Toyonaka, Osaka 560-0043, Japan*
[b]*Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan*

Received 13 June 2004; revised 18 August 2004; accepted 18 August 2004
Available online 15 September 2004

Abstract—Advanced glycation end products (AGE) are accumulated in human tissues when long-lived proteins are glycated due to hyperglycemia and/or aging. In this study, we synthesized a collagen model peptide, Ac-(Pro-Hyp-Gly)$_5$-Pro-Lys-Gly-(Pro-Hyp-Gly)$_5$-Ala-NH_2 to investigate intact AGEs in peptides. The peptide formed a stable triple helix structure, and was subjected to glycation reactions with glucose, ribose and glyoxal. Besides carboxymethyl-lysine in the peptide, a conjugated form linked with glyoxal lysine dimer (GOLD) was detected upon treatment with glyoxal. This is the first example of intact glycation-derived dimers of peptides retaining intrinsic protein structures.

Nonenzymatic glycation reactions of proteins lead to formation of advanced glycation end products (AGEs). Hyperglycemia inherent to diabetes patients accelerates accumulation of AGE, which consequently causes diabetes complications. Among long-lived proteins susceptible to glycation, collagen, the most abundant protein family in human tissues, is potentially implicated in AGE-related illnesses such as osteoarthritis,[1,2] cardiovascular failures,[3,4] and renal disorders.[5,6] Chemical and biochemical investigations have been carried out to identify the structural entities of AGEs and their formation mechanisms. In the last two decades, the structures of glycation moieties and their linkages between proteins have been extensively studied to accumulate the chemical knowledge on this complicated reactions and their resultant products. As seen for the diagnostic use of hemoglobin A1c, the initial glucose conjugation to an amino group in proteins triggers a further glycation reaction sequence. In addition, other carbonyl equivalents are known to play a crucial role in forming AGE in human tissues. Dicarbonyl agents represented by glyoxal and methylglyoxal can be generated through the primary metabolic pathway of carbohydrates as well as the oxidative degradations of unsaturated fatty acids.[7–9] Glyoxal-derived AGEs have therefore been extensively investigated, and some of them turned out to be heavily involved in formation of AGEs. Among these, carboxymethyl-lysine (CML, **1**)[10] is one of the best known AGE, and often utilized as a diagnostic marker for protein glycation in patients. Besides these *N*-substituted lysine and arginine, crosslinkers of proteins comprising glycated lysine/lysine or lysine/arginine have been attracting much attention since protein crosslinking greatly damages physiological functions, and potentially causes AGE-related complications. There are several crosslinkers reported from glycated proteins such as crossline, DOGDIC, GODIC, GOLD, MOLD, pentosidine, and so on. In particular, GOLD (**2**), glyoxal lyse dimer, generated from two molecules of lysine and glyoxal, has been found from human tissues as one of the aging-related AGEs.[11]

Most of AGEs including CML[10] and GOLD[12] were initially discovered from reactions between amino acids and sugars; that is, the detailed structure studies of AGEs have so far been based on products obtained from *N*-protected lysine and *N*-protected arginine in the presence of glucose, glyoxal or other glycating agents. Some of these AGEs such as CML, pentosidine, and GOLD[7,11,13,14] have been also detected from human tissues. Despite a great advance brought by these biochemical studies, little data are available as to how

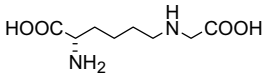
Keywords: Advanced glycation end products; Collagen.
* Corresponding author. Tel./fax: +81 66850 5774; e-mail: murata@ch.wani.osaka-u.ac.jp

0960-894X/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmcl.2004.08.044

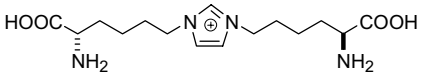
closely glycation reactions of proteins in human can be reproduced by model systems using an amino acid–sugar mixture. Previous studies have demonstrated that acid treatments such as hydrolysis of peptides results in degradation of some AGEs.[15–17] To know what is exactly happening in protein glycation under physiological conditions, intact AGEs should be examined without chemical or enzymatic degradation. Glycation of natural proteins, however, results in an extremely complexed mixture of AGEs, making their structure elucidation practically impossible. To address these problems, we attempted to prepare a simple and small model peptide of collagen that is expected to provide a limited number of glycation products. In this letter, we report the synthesis of lysine-containing collagen model peptide **3** and its glycation products.

A collagen model was designed as **3**, which possesses several features to facilitate AGE analysis. First, the model peptide of 34 residues comprises the repeated tripeptides of Pro-Hyp-Gly that stabilize the triple helix structure of collagen.[18] Secondly, one lysine residue is introduced in place of Hyp 17 as a single site for glycation, which should simplify the profile of AGEs. Thirdly, the N- and C-termini of the peptide are substituted with *N*-acetyl and amide groups, respectively, to reduce ionic repulsion.

Collagen model peptide was prepared by the combined use of liquid-phase and solid-phase syntheses. Tripeptide Boc-Pro-Hyp(Bzl)-Gly was obtained by a solution phase method as shown in Scheme 1. Coupling between



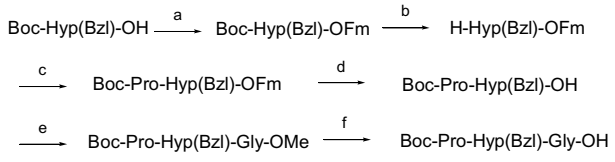
Carboxymethyl-lysine (CML) **1**



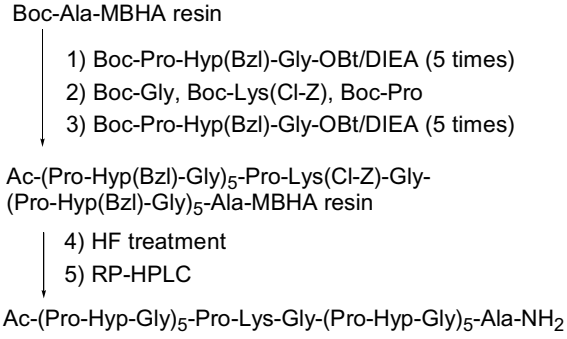
Glyoxal lysine dimer (GOLD) **2**

Ac-(Pro-Hyp-Gly)$_5$-Pro-**Lys**-Gly-(Pro-Hyp-Gly)$_5$-Ala-NH$_2$

Collagen Model Peptide **3**



Scheme 1. Reagents and conditions: (a) FmOH, WSCl/HCl, DMAP, CH_2Cl_2, rt, 12 h; (b) HCl, dioxane–CH_2Cl_2, rt, 1 h (89% for two steps); (c) Boc-Pro-OH, WSCI/HCl, HOBt, DIEA, DMF, rt, 12 h (80%); (d) piperidine, DMF, rt, 3 h (78%); (e) Boc-Gly-OCH_3, WSCI/HCl, HOBt, DIEA, DMF, rt, 40 min (84%); (f) KOH aq THF/H_2O 5:1 rt, 1 h (98%).



Scheme 2. Conditions for coupling between Boc-amino acid and peptide on resin for (1)–(3): (a) washed with *N*-methylpyridone and then CH_2Cl_2; (b) 50% TFA/CH_2Cl_2, rt, 5 min and then 25 min; (c) washed with CH_2Cl_2 and then *N*-methylpyridone; (d) dicyclohexylcarbodiimide-activated Boc-amino acid or Boc-Pro-Hyp(Bzl)-Gly, *N*-methylpyridone, rt, 90 min; (e) washed with CH_2Cl_2 and then *N*-methylpyridone; (f) Ac_2O, *N*-methylpyridone, rt, 5 min. (4) Anhydrous HF, anisole, 0 °C, 90 min. (5) See Fig 1.

Boc proline and Hyp(Bzl)-OFm (*O*-benzyl-hydroxyproline fluorenylmethyl ester) followed by deprotection with piperidine gave Boc-Pro-Hyp(Bzl)-OH. Further coupling with Gly-OMe provided Boc-Pro-Hyp(Bzl)-Gly-OMe. After hydrolyzing methylester, the tripeptide was purified by crystallization as a dicyclohexylamine salt. The solid-phase synthesis of peptide **3** using Boc-Pro-Hyp(Bzl)-Gly-OH was carried out as depicted in Scheme 2. Starting from Boc-Ala-MBHA resin (MBHA resin: 4-methylbenzhydrylamine resin), Boc-Pro-Pro-Gly was condensed five times according to a standard Boc solid-phase method and then Boc-Gly, Boc-Lys(Cl-Z), and Boc-Pro were successively condensed to furnish 19-mer peptide. Further coupling with Boc-Pro-Hyp(Bzl)-Gly five times provided the protected peptide resin, which was then treated with anhydrous HF containing 10% anisole to furnish collagen model **3**.[19] The HPLC elution profile of peptide **3** showed two peaks, implying that **3** undergoes association/dissociation between monomeric and trimeric forms under the HPLC conditions; although one at 18.5 min may probably be due to the triple helix and the other at 21.1 min due to monomers, their assignment is not completed.

The transition temperature (T_m) from triple helix to monomeric forms was determined by a differential scanning calorimeter (DSC) to be 69.8 °C, which significantly exceeded that of (Pro-Hyp-Gyl)$_{10}$, 60.1 °C.[20] This high T_m of **3** assured the stability of triple helix during a long glycation period, typically three months at 37 °C. The molecular weight of the triple helix is about 9.2 kDa, which can be regarded as a small protein. Collagen model **3** was subjected to glycation reactions using three different agents, glucose, ribose, and glyoxal,[21] where incubation times were varied depending on their reactivity; three month at 37 °C for glucose, four weeks at 37 °C for ribose, 24 h at 40 °C for glyoxal. In the MALDI-TOF spectra of their glycation products, an ion peak corresponding to CML-containing **3** was commonly observed at *m*/*z* 3142 (Fig. 2**a** and **c**); a spectrum for ribose is not shown since it is similar to Figure 2**a** except for a lack of

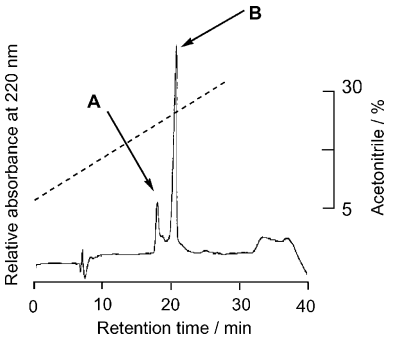


Figure 1. HPLC elution profile of peptide **3**. Column: Cosmosil 5C18-MS (10 × 250 mm) at a flow rate of 2.5 mL/min. Eluent: aqueous acetonitrile containing 0.1% TFA with linear gradient elution as depicted in figure. Peaks **A** and **B** gave rise to an identical molecular mass in MALDI-TOF, suggesting one of them in the triple helix form.

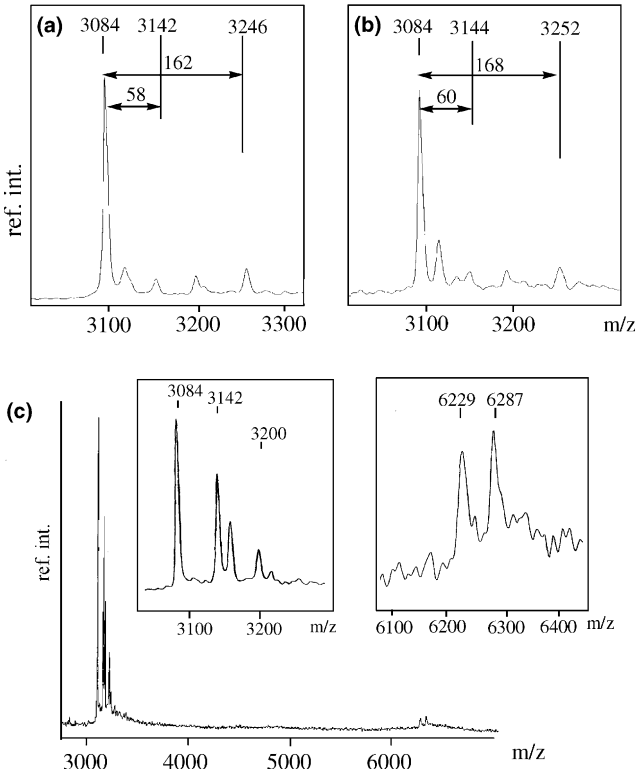


Figure 2. MALDI-TOF spectra of peptide **3** after glycation treatment. (**a**) Glycation products from **3** and glucose; (**b**) those from **3** and $^{13}C_6$-glucose; (**c**) those from **3** and gyoxal. An ion peak at *m/z* 3142 in **a** or inset spectra of **c** corresponds to CML-bearing **3**, and that at *m/z* 6229 corresponds to a dimer of **3** linked with GOLD. An ion peak at 6287 in **c** matches a further glyoxal adduct of the GOLD-linked dimer.

ion at *m/z* 3246. The presence of CML **1** in **3** was further confirmed by amino acid analysis for its hydrolysates.[22] When, in lieu of glucose, $^{13}C_6$-glucose was used for glycation reactions, the products gave rise to ion peaks at *m/z* 3144 (Fig. 2**b**), clearly indicating the incorporation of C_2 unit from glucose. The amount of CML formed in glucose-treated **3** was estimated to be 10% of lysine based upon HPLC quantification[23] for acid hydrosates, which roughly corresponds to that by MALDI-TOF (Fig. 1**a**). On the other hand, the yield of CML from natural collagen[23] was estimated to be less than 0.1% of total lysine residues under the same glycation conditions although a trace amount of CML was detected by LC–MS. This low yield is probably because further conversion of CML proceeded in natural collagen to generate more complicated AGEs. AGE with 162/168 mass-unit higher ion was detected (Fig. 2**a/b**), which may correspond to a fructosyl derivative of **3**, an Amadori product of glucose-conjugated **3**. Among the AGEs obtained with glyoxal, ion peaks corresponding dimers of **3** appeared at *m/z* 6200–6300 (Fig. 2**c**). In particular, an ion peak at *m/z* 6229 corresponds to dimer of **3** linked with GOLD. The generation of a GOLD moiety was confirmed by LC–MS analysis.[24] To our knowledge, the present result provides the first example of observation of an intact glycation-derived peptide dimer with intrinsic higher structures. In the glycation reactions with glucose and ribose, AGEs in early stage such as CML and fructosyl-lysine were formed, which suggested that a longer incubation period is necessary for more advanced AGEs including crosslinkers. In contrast, glyoxal elicited dimerization of the peptide due to its higher reactivity.

As Reiser et al. have reported,[25] the glycation of type I collagen occurs in a sequence specific manner, which evidently indicates that AGE formation reactions can never be reproduced with amino acid derivatives. The present model peptide **3** is much smaller than collagen, and may not replicate the higher structure such as the super coil and long fibrous assemblage. It does, however, form the triple helix characteristic of the natural protein. The dominant formation of GOLD in the peptide may be in parallel with the higher content of GOLD (and MOLD) than other dicarbonyl-derived crosslinkers in human plasma and tissues.[11,14,26] Moreover, we have observed some differences between amino acid derivatives and model peptide **3** in preliminary experiments; the formation rate of CML appears higher with **3** than with *N*-Boc-lysine.[27] The present findings suggest the utility of protein mimics for questing true AGEs occurring in human tissues. Further glycation reactions using collagen model peptides are currently underway.

Acknowledgements

We are grateful to Dr. Nobuaki Matsumori in authors' laboratory (M.M.) for discussion; and to Dr. Katsumasa Iijima, Nippi Research Institute of Biomatrix for information on collagen glycation. This work was supported in part by a Grant-In-Aid for Scientific Research on Priority Area (A) (No. 12045243) from MEXT, Japan.

References and notes

1. Uchiyama, A.; Ohishi, T.; Takahashi, M.; Kushida, K.; Inoue, T.; Fujie, M., et al. *J. Biochem.* **1991**, *110*, 714–718.
2. Vierzijl, N.; DeGroot, J.; Oldehinkel, E.; Bank, R. A.; Thorpe, S. R.; Baynes, J. W., et al. *Biochem. J.* **2000**, *350*, 381–387.

3. Brownlee, M.; Vlassara, H.; Kooney, A.; Ulrich, P.; Cerami, A. *Science* **1986**, *232*, 1629–1632.
4. Varan, S.; Foiles, P.; Founds, H. *Arch. Biochem. Biophys.* **2003**, *419*, 89–96.
5. Makita, Z.; Radoff, S.; Rayfield, E. J.; Yang, Z.; Skolnik, E., et al. *New. Engl. J. Med.* **1991**, *325*, 836–842.
6. Mene, P.; Festucia, F. In *Advanced Gycation End-Products in Nephrology*; Berlyne, G., D'Angelo, A., Favaro, S., Gambaro, G., Eds.; Karger: Basel, 2000; pp 22–32.
7. Nagaraj, R. H.; Shipanova, I. N.; Faust, F. M. *J. Biol. Chem.* **1996**, *271*, 19338–19345.
8. Ahmed, M. U.; Thorpe, S. R.; Baynes, J. W. *Biochem. J.* **1997**, *324*, 565–570.
9. Monnier, V. M. *Arch. Biochem. Biophys.* **2003**, *419*, 1–15.
10. Ahmed, M. U.; Thorpe, S. R.; Baynes, J. W. *J. Biol. Chem.* **1986**, *261*, 4889–4894.
11. Brinkmann-Frye, B. E.; Degenhardt, T. P.; Thorpe, S. R.; Baynes, J. W. *J. Biol. Chem.* **1998**, *273*, 18714–18719.
12. Wells-Knecht, K. J.; Brinkmann, E.; Baynes, J. W. *J. Org. Chem.* **1995**, *60*, 6246–6247.
13. Sell, D. R.; Monnier, V. M. *J. Biol. Chem.* **1989**, *264*, 21597–21602.
14. Odani, H.; Shinzato, T.; Usami, J.; Matsumoto, Y.; Brinkmann-Frye, E.; Baynes, J. W.; Maeda, K. *FEBS Lett.* **1998**, *427*, 381–385.
15. Lederer, M. O.; Bühler, H. P. *Bioorg. Med. Chem.* **1999**, *7*, 1081–1088.
16. Iijima, K.; Murata, M.; Takahara, H.; Irie, S.; Fujimoto, D. *Biochem. J.* **2000**, *347*, 23–27.
17. Biemel, K. M.; Friedl, D. A.; Lederer, M. O. *J. Biol. Chem.* **2002**, *277*, 24907–24915.
18. Inoue, K.; Sakakibara, S.; Prokop, D. J. *Biochem. Bophys. Acta* **1976**, *420*, 133–141.
19. Identity of **3** was confirmed by MALDI-TOF MS, which revealed the $(M+H)^+$ at *m/z* 3084.
20. Uchiyama, S.; Kai, T.; Kajiyama, K.; Kobayashi, Y.; Tomiyama, T. *Chem. Phys. Lett.* **1997**, *281*, 92–96.
21. Collagen model **3** (490 nM, 1.5 mg/mL) or natural collagen (acid-soluble collagen from calf skin, SIGMA, 1.5 mg/mL) was incubated in 50 mM phosphate buffer (pH 7.5) in the presence of 3 mM NaN_3 with D-glucose (0.2 M) or $^{13}C_6$-D-glucose (0.2 M) at 37 °C for three months. Glycation with ribose was carried out in the same manner except for use of ribose or 1,2-$^{13}C_2$-ribose and the incubation time of four weeks. Glycation with glyoxal was performed in 50 mM phosphate buffer (pH 8.5) with glyoxal (30 mM) at 40 °C for 24 h. The glycated **3** in these three preparations was then passed through a PD-10 column to remove sugar or glyoxal.
22. CML in **3** was detected in the hydrolysate (6 M HCl at 110 °C for 24 h), which was subjected to amino acid analysis (L-8500, Hitachi). CML is known to be stable in these hydrolysis conditions and obtained as 20–30% of a lysine residue in the reaction with glyoxal.
23. Collagen model **3** and natural collagen were incubated with glucose[21] and hydrolyzed with 6 M HCl at 110 °C, 24 h. After phenylthio-carbamyl (PTC) derivatization with phenylisothiocyanate, CML and lysine were determined by HPLC on an ODS column eluting with 0.1% TFA-containing aqueous 15–35% MeCN with monitoring at 254 nm.
24. GOLD **2** was identified in the hydrolysate by LC–MS; glyoxal-treated **3** was hydrolyzed with 4 M HCl, 110 °C, 24 h. After PTC derivatization, LC–MS analysis monitoring at *m/z* 597 clearly demonstrated the presence of PTC_2-GOLD. LC–MS was carried out using an ion-trap MS instrument (LCQ-DECA, Thermo-Quest) with the ion spray ionization mode, which was not suitable for quantitative analysis. The yield of GOLD either in **3** or natural collagen could not be determined by UV-monitoring HPLC analysis for PTC derivatives[23] because of the presence of co-eluting multiple impurities.
25. Reiser, K. M.; Amigable, M. A.; Last, J. A. *J. Biol. Chem.* **1992**, *267*, 24207–24216.
26. Chellan, P.; Nagaraj, R. H. *Arch. Biochem. Biopys.* **1999**, *368*, 98–104.
27. Details will be published elsewhere.