A new glycation product ‘norpronyl-lysine,’ and direct characterization of cross linking and other glycation adducts: NMR of model compounds and collagen

Peter T. B. BULLOCK*, David G. REID*, W. YING CHOW*, Wendy P. W. LAU* and Melinda J. DUER*

*Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, U.K.

Synopsis

NMR is ideal for characterizing non-enzymatic protein glycation, including AGEs (advanced glycation endproducts) underlying tissue pathologies in diabetes and ageing. Ribose, R5P (ribose-5-phosphate) and ADPR (ADP-ribose), could be significant and underinvestigated biological glycating agents especially in chronic inflammation. Using [U-13C]ribose we have identified a novel glycoxidation adduct, 5-deoxy-5-desmethylpronyl-lysine, ‘norpronyl-lysine’, as well as numerous free ketones, acids and amino group reaction products. Glycation by R5P and ADPR proceeds rapidly with R5P generating a brown precipitate with PLL (poly-L-lysine) within hours. ssNMR (solid-state NMR) identifies several crosslinking adducts such as the newly identified norpronyl-lysine, in situ, from the glycation reaction of 13C5-ribose with collagen. The same adducts are also identifiable after reaction of collagen with R5P. We also demonstrate for the first time bio-amine (spermidine, N-acetyl lysine, PLL) catalysed ribose 2-epimerization to arabinose at physiological pH. This work raises the prospect of advancing understanding of the mechanisms and consequences of glycation in actual tissues, in vitro or even ex vivo, using NMR isotope-labelled glycating agents, without analyses requiring chemical or enzymatic degradations, or prior assumptions about glycation products.

Key words: ADP-ribose, advanced glycation endproducts, Maillard reaction, ribose, ribose-5-phosphate, solid-state NMR

INTRODUCTION

Glycation (non-enzymatic glycosylation) is the reaction of sugars to form covalent adducts with proteins. It occurs wherever sugars and proteins are in contact, usually initially through so-called Maillard reactions (see Supplementary Data Section S1; available at http://www.bioscirep.org/bsr/034/bsr034e096add.htm for summaries of these, and other glycation reactions and products and their abbreviations). Initial products react further to form AGEs (advanced glycation endproducts), including protein modifications, sugar degradation products and protein crosslinks, the latter implicated in numerous conditions including diabetes, atherosclerosis, osteoarthritis and cataracts [1,2]. The Maillard reaction with glucose in vivo is slow at physiological temperature and pH due to its low reactivity, so only long-lived structural proteins such as collagen and elastin are significantly affected by crosslinking arising from glycation by this sugar [2].

Collagen crosslinking is a poorly understood pathological consequence of glycation, leading to irreversible degeneration of mechanical properties, such as stiffer, more brittle connective tissue fibres. Both crosslinking and simple modifications of amino acid side-chains can lead to abnormal biochemical functionality of collagen by changing the charge profile of the collagen fibrils, with unpredictable consequences [3]. For instance, modification of lysine residues to the relatively...
common AGEs N\textsuperscript{ε}-CML [(carboxymethyl)lysine] and N\textsuperscript{ε}-CEL [(carboxethyl)lysine] changes a positively charged side-chain.

Research has focused on glycation by glucose, assumed to be the most relevant sugar in both non-pathological tissues and hyperglycaemic conditions. However, glucose is not an efficient glycating agent as it overwhelmingly adopts its thermodynamically stable pyranose form, rather than the reactive open-chain aldehyde form. Pentoses such as ribose and especially R5P (ribose-5-phosphate) are much more potent glycating agents, R5P reacting with amines 150-fold faster than glucose [4,5]. R5P and the mechanistically almost indistinguishable ADPR (adenosine diphosphate ribose) are both released into the extra-cellular matrix on cell necrosis, so are present in chronic inflammation where there is long-term cellular damage, and so could be potent biologically significant and, to date, under-investigated glycating agents. Some established reaction pathways to AGEs are not available to R5P or ADPR due to phosphoesterification of the terminal ribosyl 5-carbon; equally, it is possible that there are glycation products of R5P or ADPR that do not occur for glucose. The aim of this work is to identify the possible glycation products of ribose, R5P and ADPR in both model systems and collagen itself.

Glycation products can be detected by a variety of techniques in different tissues and experimental models systems [6–20], including solution state NMR (hereafter simply referred to as ‘NMR’ in distinction to ssNMR (solid-state NMR)), which can detect soluble AGEs [12,21–23]. Although all these previous approaches yield valuable chemical and mechanistic insights, they are only applicable with considerable prior sample preparation to the large insoluble structural proteins, which are significant pathogenic targets of tissue glycation. ssNMR, however, bypasses many of these difficulties and can access a wealth of chemical information from these solid materials in a close to native state. Thus, in this work we develop NMR spectroscopic and labelling approaches to analysing mixtures of glycation products, in particular those from ribose, R5P and ADPR, and to exploring the products of glycation of collagen itself. The approaches we develop could find applications in the study of tissue glycation in hyperglycaemic diseases such as diabetes.

**MATERIALS AND METHODS**

Reagents were from Sigma–Aldrich except [U-\textsuperscript{13}C]ribose, which was from Cambridge Isotope Laboratories. PLL (poly-L-lysine) hydrobromide was molecular mass 4,000–15,000 Da. Reagent salt forms were: Spermidine trihydrochloride, disodium R5P, sodium ADPR. Dry Type I collagen (bovine Achilles tendon) was rehydrated by soaking overnight in aqueous acetic acid (10%), homogenized with a hand held plunger homogenizer, and washed with distilled water and centrifuged repetitively until no longer acidic, and stored wet at 4 °C until used. All incubations were performed at 37 °C in 50 mM potassium phosphate at pH 7.4 in 5% D\textsubscript{2}O/H\textsubscript{2}O to which a few crystals of sodium azide had been added to ensure sterility, unless otherwise indicated, and progress was monitored periodically by NMR. Glycation products were shown to be reproducible by repeating each experiment.

**Glycation of model amines**

Composition of individual reaction mixtures in 1.05 ml (unless stated otherwise) solution are summarized below. To facilitate periodic analyses by NMR these reactions were performed in 4 mm outer diameter glass NMR tubes.

- **α-N-acetyl lysine**
  - AcLys 9.5 mg, ribose 7.5 mg; AcLys 9.5 mg, [U-\textsuperscript{13}C]ribose 7.5 mg; AcLys 9.5 mg R5P 13 mg.

- **Spermidine**
  - Spermidine 13 mg, ribose 7.5 mg; Spermidine 13 mg, R5P 13 mg.

- **PLL**
  - PLL 12 mg, ribose 7.5 mg; PLL 12 mg, [U-\textsuperscript{13}C]ribose 7.5 mg; PLL 12 mg, ADPR 28 mg.

- **Controls**
  - Ribose (7.5 mg), [U-\textsuperscript{13}C]ribose (7.5 mg), and R5P (12 mg), were each incubated for a month; the spectrum of each was unchanged after this time.

**Glycation of samples for ssNMR**

**Large scale glycation of PLL**

550 mg PLL and 750 mg R5P were incubated in 45 cm\textsuperscript{3} buffer containing sodium azide for a day. The resultant brown precipitate was washed with 40 ml water and centrifuged, three times, and air dried overnight for ssNMR.

**Collagen incubations under biomimetic conditions**

2 g wet mass of collagen was used in all incubations. A few crystals of sodium azide were added to incubations for sterility. All glycated solids were washed with 40 ml water and centrifuged, three times, and air dried overnight for ssNMR.

Collagen and R5P: a mixture of collagen and 60 mg R5P disodium salt was made up to 4 cm\textsuperscript{3} with water and incubated at 310 K for 14 days. Phosphate buffer (1 M, 2 ml) was added after 8 days. The glycated collagen was distinctly yellow at this point.

Collagen and [U-\textsuperscript{13}C]ribose: a mixture of collagen and [U-\textsuperscript{13}C]ribose (35 mg) was made up to 4 cm\textsuperscript{3} with phosphate buffer (50 mM) and incubated for 69 days, at which point the collagen was distinctly orange. The pale yellow supernatant was collected for NMR analysis.

Collagen and R5P (pH 9, sonicated): a mixture of collagen (2 g, wet) and R5P (60 mg) was made up to 4 cm\textsuperscript{3} with phosphate buffer (100 mM, pH 9.0 attained with NaOH), sonicated
and incubated at 310 K for 11 days, at which point the collagen was yellow.

**NMR**

All $^{13}$C and $^{31}$P experiments were performed in H$_2$O/2H$_2$O using standard methodology on a Bruker 11.8 Tesla Avance-500 standard bore spectrometer, at frequencies of 500.1 MHz (1H), 125.6 MHz ($^{13}$C) and 202 MHz ($^{31}$P). Pulse-acquire and $^{13}$C-$^{13}$C COSY, was performed with continuous broadband decoupling. $^{13}$C and $^{31}$P chemical shifts were referenced, respectively, to internal TSP (trimethylsilapentane) sulfonate sodium salt, and external 85% phosphoric acid, at 0 ppm.

**ssNMR**

All measurements were performed on a Bruker 9.4 T Avance-400 wide bore spectrometer, with a standard Bruker MAS 4 mm double resonance probe operating at 400.4 MHz, 100.6 MHz and 162.1 MHz for 1H, $^{13}$C and $^{31}$P, respectively. Samples were in 4 mm zirconia rotors and spun at 12.5 kHz at the ‘magic’ angle of 54.7° unless stated otherwise. CP (cross-polarization) from 1H was used to enhance signal intensity ($^{1}$H π/2 pulse 2.5 μs, CP field strength 70 kHz, CP time 2.5 ms, spine64 broadband decoupling (100 kHz field) during acquisition, repetition time 2 s. SQ–DQ (single quantum–double quantum) $^{13}$C-$^{13}$C ssNMR correlation was achieved using POST-C7 (permutationally offset stabilized C7) [24] methodology and MAS (magic angle spinning) frequency of 10 kHz. $^{13}$C and $^{31}$P spectra were referenced to the glycine Cα signal at 43.1 ppm, and to hydroxyapatite at 2.6 ppm, respectively.

**RESULTS**

Many protein gyations occur through lysyl amines, so in this work these processes were modelled by using the simpler amines PLL, AcLys and spermidine as models before examining the glycation of collagen. Glycation reactions with ribose, R5P and ADPR all produced numerous AGEs and AGE analogues, which we categorize as (i) sugar modifications, (ii) protein side chain modifications and (iii) cross links.

**Sugar modifications**

The majority of Maillard reactions do not in fact consume the amine [5], rather the amine catalyses sugar degradations and rearrangements, which would not occur in the absence of amines.

**Catalytic epimerization of ribose**

All the amines catalyse epimerization of ribose into arabinose, readily shown by the replacement of ribose NMR signals by signals characteristic [25] of the two anomer forms of arabinose as the reactions proceed (Supplementary Data Section S2 Figure S2; available at http://www.bioscirep.org/bsr/034/bsr034e096add.htm). This result is significant because, although previously observed in strong alkali [26], such rearrangements have not been reported before in such mild biomimetic conditions as those used in this work. This suggests that the amine is integral to the mechanism of epimerization, with the reaction mechanism in Supplementary Data Section S1 Scheme 1 explaining the observation if the Amadori rearrangement (Step 2) is significantly reversible.

**Generation of carboxylic acids**

Simple carboxylic acids were common Maillard products detected in all reaction systems, Supplementary Data Section S3 Table S1 (available at http://www.bioscirep.org/bsr/034/bsr034e096add.htm) summarizes likely assignment of $^{13}$C solution-state NMR signals from possible 1- and 2-carbon atom acids resulting from oxidative degradation of ribose and R5P. Formic, acetic and glycolic acids were abundant AGEs from all reaction mixtures, with oxalic acid also common although generated in smaller quantities; the occurrences of these acids is summarized in Supplementary Data Section 3 Table S2. Acetic and glycolic acids are unlikely to have any pathological effects, formic acid is metabolized rapidly, but even so its generation by glycation in the eyes may have pathological consequences due to its high ocular toxicity. Oxalic acid generation in sufficient concentrations may lead to pathological precipitation of calcium oxalate, the most significant component of kidney stones.

**Norfuraneol**

This was observed in all ribose/amine incubations, but not in R5P/amine incubations. A full $^{13}$C solution-state NMR assignment of $^{13}$C norfuraneol is given in Supplementary Data Section S4 Table S3 (available at http://www.bioscirep.org/bsr/034/bsr034e096add.htm). The lack of norfuraneol as an AGE in reactions of R5P was expected [5] as R5P lacks the C5 hydroxyl group necessary for the final cyclization step that produces norfuraneol (see Online Supplemental Material Scheme S2).

**Protein side chain modifications**

Amadori compound

All incubations of ribose or R5P with amines yielded a small transient $^{13}$C NMR signal at 208–210 ppm (Online Supplementary Material Section S5; available at http://www.bioscirep.org/bsr/034/bsr034e096add.htm), most likely due to the Amadori product, ribulosesylsine (or 5-phospho-ribulosesylsine, see Supplementary Data Scheme S1), which is predicted to be present at low concentrations according to steady-state kinetics. The signal appears as a doublet of doublets (J = 46 and 40 Hz) in the $^{13}$C$_{3}$-ribose/AcLys system, and shows $^{13}$C-$^{13}$C COSY correlations to signals at 57 and 79 ppm in the $^{13}$C$_{3}$-ribose/PLL system, as expected for the ribulosesyl-carbonyl carbon [27]. This suggests that the progress of the Maillard reaction in vitro
P. T. B. Bullock and others

Scheme 1  Postulated mechanism of formation of norpronyl-lysine from the Amadori products of both ribose and R5P with lysine

Note the similarity to the established mechanism of formation of norfuraneol in the cyclization, rearrangement and elimination of water (Online Supplemental Material Section S1 Scheme S2).

can be monitored by the evolution of this ca. 209 ppm signal, since once it has disappeared, the initial glycation step (whereby free ribose becomes covalently bound to the amine) must be complete.

CML and CEL

The anticipated AGE, CML, was readily identified by $^{13}$C NMR spectroscopy as a pair of signals at 52.1 and 174.3 ppm ($^{13}$C-enriched ribose experiments yielded two doublets, $J = 53$ Hz). In the liquid state this provides a very straightforward diagnostic for CML. However, since in a protein the peptide $\alpha$-carbons and carbonyls resonate at $\sim 50$ and $\sim 175$ ppm, respectively, in the solid state there is no chance of resolving CML signals from those of the protein without selective $^{13}$C enrichment. No evidence was found for CEL as a glycation product with ribose, RSP or ADPR; since the methyl carbon of CEL gives a distinctive resonance at $\sim 15$ ppm [28], the absence of such a signal implies that very little CEL was generated.

Acetylation

$^{13}$C-$^{13}$C COSY analysis of $^{13}$C$_2$-ribosylated PLL reveals a correlation between 23 and 177 ppm. This is characteristic of an acetyl amide, $N^\epsilon$-acetyl lysine, a known AGE.

Discovery of a novel pyrrolinone reductone AGE

A set of five COSY correlated $^{13}$C resonances were well resolved in [U-$^{13}$C]ribosylated PLL (Figure 1) and in [U-$^{13}$C]ribosylated AcLys, at 59 ↔ 187 ↔ 130 ↔ 176 ↔ 14 ppm due to a previously unreported pyrrolinone reductone [29] AGE for which we propose the name ‘norpronyl-lysine’ for its similarity to the established AGE pronyl-lysine [30]. A full assignment of the $^{13}$C NMR spectrum of norpronyl-lysine including through-bond coupling constants ($J_{CC}$ values) is given in Table 1.

Norpronyl-lysine is postulated to form by 5-elimination from the Amadori product [31] followed by cyclization, rearrangement and dehydration (Scheme 1). This mechanism implies that RSP should undergo conversion to norpronyl-lysine more readily than ribose as elimination of a phosphate group proceeds.

86

© 2014 The Author(s) This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC-BY) (http://creativecommons.org/licenses/by/3.0/) which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.
Novel AGE, and protein crosslink characterization, by NMR

Figure 1 13C–13C COSY correlation spectrum of [U-13C]ribose glycated PLL

Portions of the 2D 13C–13C COSY correlation spectrum acquired from PLL after 10 days incubation with 13C5-ribose. The data is presented as a contour plot in which the diagonal (bottom left to top right) represents a conventional 1D spectrum 'viewed from above'. The off-diagonal 'cross peak' features connect signals between carbon atoms which are directly bonded to each other, at chemical shifts (i.e. frequencies) which are effectively Cartesian coordinates of the two bonded atoms. Cross peaks corresponding to the norpronyl group (shown) are assigned and connected by solid vertical and horizontal lines.

Table 1 Full assignment of 13C5-norpronyl-lysine as observed after glycation of AcLys by [U-13C]ribose

<table>
<thead>
<tr>
<th>carbon atom</th>
<th>( \Delta ) (ppm)</th>
<th>Multiplicity</th>
<th>J (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>59.3</td>
<td>dd</td>
<td>43, 19</td>
</tr>
<tr>
<td>2</td>
<td>186.6</td>
<td>ddd</td>
<td>73, 43, 14</td>
</tr>
<tr>
<td>3</td>
<td>129.9</td>
<td>ddd</td>
<td>77, 73, 19</td>
</tr>
<tr>
<td>4</td>
<td>176.6</td>
<td>ddd</td>
<td>77, 44, 14</td>
</tr>
<tr>
<td>5</td>
<td>13.6</td>
<td>d</td>
<td>44</td>
</tr>
</tbody>
</table>

By comparison with the ribose/PLL system it appears to be due to norpronyl-lysine. The change in chemical shift of the carbonyl group between liquid state and solid state (\( \approx 5 \) ppm) is not unexpected, and is attributable to packing effects in solids which are not operative in the liquid, and/or changes in hydrogen bonding between solid and solution state.

The biological effects of norpronyl-lysine can only be speculated at, but the very similar lysine modification proline-lysine may be a free radical scavenging antioxidant which inhibits development of colonic precancerous lesions [32]. On the other hand, as a protein modification it is likely to have a negative impact on the functionality of collagen and on collagen recycling by collagenases.

Crosslinking

PLL with R5P, and with ADPR

Reaction between R5P and PLL led to precipitation of a brown polymer after only a few hours, attributed to the formation of crosslinks between the strands of PLL. This dramatic glycation effect was also observed for ADPR and shown, by 31P and 13C NMR (Supplementary Data Section S6 Figures S4 and S5; available at http://www.bioscirep.org/bsr/034/bsr034e096add.htm, respectively), to proceed via elimination of ADP, consistent with the literature [5] and further corroborated by 31P ssNMR.
several overlapping components (Figure 2) consistent with both MOLD, as well as signal in the sugar region (60–80 ppm) from of these species.

glyoxal lysine dimer) [33] and GOLD (glyoxal lysine dimmer) signals consistent with the expected crosslinks MOLD (methyl- 

effectively only shows those which originate from [U-13C]ribose. 

Assignment was aided by 2D SQ–DQ correlations between cova-

lent bonds13C p a i r s (Figure 4), which demonstrated that many 

signals from13C atoms not directly bonded to another 13C and hence 

unambiguously derived from [U-13C]ribose.

Figure 2 13C ssNMR from the PLL–R5P reaction

13C ssNMR of PLL before (black solid lines, bottom) and after (red, dashed lines) incubation with R5P Glycation broadens the PLL peaks due probably to increased chemical diversity. The most significant new signals characterized by signal-to-noise ratios of greater than 5, are asterisked. Tentative assignments are: 130 and 10 ppm–MOLD; 70 ppm–secondary alcohol resonances from intact sugars or sugar degradation products bound into the poly-

mer, and/or lysyl modifications such as N-glyceryl lysine; 44 ppm could correspond to C, of a modified lysine such as MOLD (which would explain the broadness as many modifications are possible). 192 ppm is from norpronyl-lysine. See Online Supplemental Material Section S1 Figure S1 for chemical structures of glycation products.

(Supplementary Data Section S6 Figure S6). Glycation of PLL by RSP led to generation of reaction products with 13C ssNMR signals consistent with the expected crosslinks MOLD (methyl-glyoxal lysine dimer) [33] and GOLD (glyoxal lysine dimmer) [34], with the resonance at 10 ppm particularly diagnostic of MOLD, as well as signal in the sugar region (60–80 ppm) from several overlapping components (Figure 2) consistent with both of these species.

Glycation of collagen

Reactions with [U-13C]ribose

Incubation of collagen with [U-13C]ribose led to numerous re-

action products derived from the ribose, whose NMR signals could be separated from those of the unaltered collagen compon-

ents using double-quantum filtering in the NMR experiment (the 1D POST-C7 experiment; Figure 3). This procedure removes all signals from 13C atoms not directly bonded to another 13C so ef-


ehently only shows those which originate from [U-13C]ribose. Assignment was aided by 2D SQ–DQ correlations between cova-

lently bonded 13C pairs (Figure 4), which demonstrated that many signals are superpositions of signals from two or more AGEs.

The most intense signal in the 13C–13C SQ–DQ spectrum is from secondary alcohols at 65–75 ppm, such as the Amadori product ribuloselysine and certain AGEs such as DOPDIC (N6-

{2-[(4S)-4-ammonio-5-oxido-5-oxopentyl]amino}-5-{(2S)-2, 3-dihydroxypropyl]-3,5-dihydro-4H-imidazol-4-ylidene}-L-

lysinate) and glyceric acid. The strong correlation between 35

and 59 ppm is evidence for pentosinane and/or its hydrolysis product DOPDIC [11]. Correlations between 55 and 175 ppm are consistent with AGEs which share the imidazole lysine-arginine crosslink motif (Figure 4), the closely related pentosinane, and CML and CEL. The correlation between 17 and 57 ppm corresponds to a methyl group adjacent to an amine, function-
alized to give in turn a correlation between 23 and 174 ppm, characteristic of an acetyl group in an acid, ester or amide. It is probably due to Nα-acetyl lysine, as the alternative, an ester, has never been reported as a stable AGE. The correlation between 59 and 192 ppm cannot be explained by any of the established AGEs, but is entirely consistent with the newly characterized norpronyl-lysine. The weakest correlation between 56 and 102 ppm corresponds to the furanose form of ribuloselysine [27]. Its persistence after 69 days of reaction suggests the Amadori product is more stable in the solid state than in solution. NMR spectral assignments thus far are summarized in Table 2. (Two autocorrelations, of the 31 ppm, and the 88 ppm, signals to themselves or other atoms with very similar shifts, have not yet been assigned.)

13C NMR of the supernatant at the end of the incubation showed a variety of soluble AGEs, including large quantities of glycolic and oxalic acids (Supplementary Data Section S7 Figure S7, http://www.bioscierep.org/bsr/034/bsr034e096add.htm). In addition there were many new signals that did not correspond to anything seen in any of the model systems, particularly several signals between 20 and 40 ppm (Supplementary Data Section S7 Figure S8). Signals due to CML were observed, which suggests that some protein or amino acids have been released into solution.

Figure 3 13C ssNMR of the collagen-[U-13C]ribose reaction

13C ssNMR spectrum of 13C5-ribosylated collagen (69 days incubation, red, dashed middle trace), and 13C 1D SQ–DQ spectrum ‘filtered’ using the POST-C7 technique (blue, dashed upper trace), shown relative to unglycated collagen incubated for the same period (black, lower trace). Resolved new signals derived from the ribose are marked with aster-

isks, but there are broad areas of new signal throughout the range 10–215 ppm as is evident from the 1D SQ–DQ filtered spectrum, which shows only signals from 13C directly bound to another 13C and hence unambiguously derived from [U-13C]ribose.
Novel AGE, and protein crosslink characterization, by NMR

Figure 4 2D ssNMR characterization of [U-13C]ribose glycated collagen

13C ssNMR SQ–DQ (POST-C7) spectrum of [U-13C]ribose glycated collagen, in which correlations are seen only between carbon atoms directly bonded to each other. The output of this experiment is presented as a contour plot in which the horizontal axis corresponds to the 1D SQ–DQ spectrum (plotted above the 2D spectrum), and the vertical axis to a ‘double quantum’ dimension. In effect the horizontal chemical shift coordinate of each cross peak corresponds to the shift of one or other of the directly bonded carbons, and the vertical coordinate to the sum of both shifts. Relevant frequency pairings are indicated. Note that the correlation labelled ‘55–175 ppm’ corresponds to several correlations over a range of frequencies. Only species derived from the 13C5-ribose molecule will give signals in this spectrum. The inset depicts the imidazole lysine-arginine crosslink motif, which leads to a POST-C7 correlation between the carboxyimidamide carbon at \( \sim 175 \) ppm and the adjacent amino carbon at \( \sim 55 \) ppm. The possible AGEs with such a motif are DOPDIC (\( R = CH_2CH(OH)CH_2OH \)), which is derived directly from ribose, GODIC (\( R = H \)), derived from glyoxal and MODIC (\( R = Me \)), derived from methylglyoxal. Some signals (asterisked) correspond to ‘autocorrelations’ between signals from directly bonded carbons with very similar or identical chemical shifts. See Table 2 for assignments.

The soluble AGES produced from ribosylation of collagen bear limited resemblance to those produced by the PLL model. This shows that PLL does not completely mimic the Maillard reactions of collagen, and a more sophisticated model system is required to help determine the species present in the supernatant. In particular PLL does not replicate the considerable proportion of glycation-reactive arginyl residues in collagen (ca. 5% versus about 3.5% lysine in Type I collagen). It is also devoid of other amino acid functionalities, post-translational modifications and stable higher order structures (such as the collagen triple helices and fibril aggregates), any of which may exert unknown influences on glycation reactions, and increase glycation product diversity. It is interesting to note that very little of the acetic and formic acids that might have been expected as a result of the model systems are observed in the supernatant.

Collagen glycation by R5P

Under biomimetic reaction conditions, changes in collagen due to unlabelled R5P glycation were hard to detect by ssNMR beyond a general reduction in 13C linewidths. Thus in order to observe glycation products more easily, the reaction was facilitated through slightly more basic (pH 9) conditions and sonicating the collagen to disperse it and ensure as many reactive (lysyl, arginyl) residues as possible were accessible for reaction.
The results of collagen glycation by R5P are shown in Figure 5. The slight sharpening of collagen signals is attributed to an increase in molecular hydration and molecular mobility after sonication and glycation. Areas where there appears to be new AGE signals are asterisked. The 12, 68 and 192 ppm signals are also signals are asterisked. The 12, 68 and 192 ppm signals are also signals are asterisked. The 12, 68 and 192 ppm signals are also.

### DISCUSSION

NMR in both the liquid and solid phases is well known as a technique giving detailed chemical information on the composition of biological fluids and tissues in a near-native state. Characterization generally does not require separation of components, prior assumptions about possible composition, or that molecules of interest fortuitously possess distinctive spectroscopic ‘marker’ properties such as fluorescence. These factors have facilitated identification of a new glycation product arising from ribose and RSP glycation for which we propose the common name norpronyl lysine, and some previously unknown and unexpected aspects of biological glycation processes and products, such as amine catalysed sugar epimerizations under mild, approximately biological conditions.

Ribose [35], and especially RSP and ADPR [36–38], are vigorous glycat chemical implications. Ribose and RSP are ubiquitous and key intermediates in several biochemical pathways including nucleotide biosynthesis and degradation and the pentose phosphate pathway, and can act as important carbon and energy sources. RSP attains significant levels greater than approximately millimolar in some tissues [39], certainly high enough to raise the possibility of its acting as a reactive glycat agent. Type 2 diabetes is characterized by activation of multiple inflammatory processes, mediators and pathways [40,41]; vascular damage leading to atherosclerosis is a typical pathological consequence [42] partly mediated in fact by specific RAGE (receptor of AGE) which initiate further cascades of inflammation [43].

ADPR and PAR (poly-ADP-ribose) are long recognized major components in crucial cell signalling, control and repair pathways [44], and there are intensive efforts to discover pharmacological modulators [45]. Potential benefits include down-regulation of multiple parallel pathways of inflammation and tissue injury in, for instance, diabetic complications. Important components of these pathways include enzymes such as PARG (PAR glycohydrolase) [46] and ADP-ribose protein lyase, which rapidly [44] generate free ADPR from PAR. This hydrolysis exposes the reactive ribosyl 1'-anomeric carbons, which are protected by glycosidic conjugation in PAR. The reactive species, normally intracellular, will flood out of necrotizing cells and be able to react with extracellular matrix, which they will do much faster than the comparatively stable glucose.

In addition to the real possibility of these pentose sugars and their derivatives acting as significant glycat agents in their own right, the reactions they undergo with proteins reproduce mechanistically many if not all of the glycat actions of generally more abundant but much more sluggish reactants such as glucose and other hexoses and hexose phosphates. This not only makes the pentoses attractive reagents for understanding the nature and effects of the more abundant sugars, it raises the possibility of performing in vitro whole tissue or even in vivo experiments with...
labelled reactive glycating agents with subsequent ex vivo NMR and other, e.g. MS, analyses greatly facilitated by the resultant increase in sensitivity and discriminating power.

**AUTHOR CONTRIBUTION**

Peter Bullock performed experiments, analysis and made a contribution to preparing the paper. David Reid conceived the study with Melinda Duer, and contributed to experiment design, implementation and interpretation, and preparing the paper. W. Ying Chow set up and optimized solid-state NMR methodology. Wendy Lau performed and analysed some experiments. Melinda Duer contributed to study conception and preparation of the paper.

**ACKNOWLEDGEMENTS**

We thank Dr Peter Grice, Duncan Howe and Andrew Mason for acquisition of solution state NMR data.

**FUNDING**

This work was supported by the U.K. Medical Research Council (MRC) [grant number MR/J007692/1 (to D.G.R.)], and Engineering and Physical Sciences Research Council (EPSRC) [general DTA fund (to W.Y.C.).]

**REFERENCES**


Received 11 December 2013/15 January 2014; accepted 29 January 2014

Published as Immediate Publication 12 February 2014, doi 10.1042/BSR20130135

© 2014 The Author(s) This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC-BY) (http://creativecommons.org/licenses/by/3.0/) which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.
SUPPLEMENTARY DATA

A new glycation product ‘norpronyl-lysine,’ and direct characterization of cross linking and other glycation adducts: NMR of model compounds and collagen

Peter T. B. BULLOCK*, David G. REID*, W. YING CHOW*, Wendy P. W. LAU* and Melinda J. DUER*

*Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, U.K.

AGE formation

An ever growing number of AGEs have been identified and characterised in the last 20 years [1–6]. The many possible reaction pathways the initial Amadori product can explore is reflected in the structural variety of AGEs discovered so far (Figure S1).

Put broadly, the possible reaction routes fall into three groups, anticipated as early as 1953 [7].

The first category, that of products formed via formation of a reactive α-dicarbonyl species, is exemplified by pentosidine, a reaction product of pentose sugars [8]. This sort of α-dicarbonyl mechanism also leads to the formation of 4-hydroxyl-5-methyl-3(2H)-furanone (norfuraneol) from ribose as shown in Scheme S2 [9].

The second category involves fragmentation of the sugar carbon chain by retro-aldol reaction, hydrolytic cleavage or oxidative cleavage. The reactive sugar degradation products glyoxal and methylglyoxal in particular play a role in the formation of many AGEs. The two pathways to formation of CML [10,11] are a good example (Scheme S3), with oxidative cleavage of the Amadori product [3] competing with reaction of lysine with glyoxal [12].

Recently a third category of AGE has been discovered in vitro and in vivo: that of carboxylic amides, derived from β-dicarbonyl cleavage (Scheme S4) [13]. Of these Nε-acetyl lysine and Nε-formyl lysine are present in human plasma at comparable levels to CML [14].

1 To whom correspondence should be addressed (email mjd13@cam.ac.uk).

© 2014 The Author(s) This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC-BY) (http://creativecommons.org/licenses/by/3.0/) which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.
Scheme S2
Mechanisms of formation of pentosidine, pentosidine and norfuraneol from ribuloselysine, the Amadori product of ribose and lysine
Scheme S3  Mechanisms of formation of CML
The pathway which dominates has been shown to depend on phosphate concentration and sugar concentration, with the glyoxal route dominant except at low phosphate and high sugar [10,11].

Scheme S4  Mechanism of formation of N\textsuperscript{ε}-acetyl lysine and N\textsuperscript{ε}-formyl lysine from ribose
The shown starting material is a degradation product of ribulosesylsine (Scheme S2). N\textsuperscript{ε}-lactoyl lysine and N\textsuperscript{ε}-glycolyl lysine can also be formed if the lysine residue attacks the other carbonyl group in the β-dicarbonyl intermediate.
Glucosepane is considered the most important known crosslink [10], being derived from glucose. Pentosidine and its precursor pentosinane are formed in the same manner as glucosepane from pentose sugars (which are formed by the autoxidation of glucose). DOGDIC (N6-[(2S,3R)-2,3,4-trihydroxybutyl]-3,5-dihydro-4H-imidazol-4-ylidene]-L-lysinate) and DOPDIC are hydrolysed forms of glucosepane and pentosinane, respectively. GOLD and Nε-CML are the products formed by reaction of lysine with glyoxal (a product from the oxidation of sugars). MOLD (methylglyoxal lysine dimer) and Nε-CEL are the analogous products formed from methylglyoxal (a product from the fragmentation of sugars). Norfuranol is formed from the amine-catalysed rearrangement and dehydration of pentose sugars.

Figure S1  Some important AGES
Glucosepane is considered the most important known crosslink [10], being derived from glucose. Pentosidine and its precursor pentosinane are formed in the same manner as glucosepane from pentose sugars (which are formed by the autoxidation of glucose). DOGDIC (N6-[(2S,3R)-2,3,4-trihydroxybutyl]-3,5-dihydro-4H-imidazol-4-ylidene]-L-lysinate) and DOPDIC are hydrolysed forms of glucosepane and pentosinane, respectively. GOLD and Nε-CML are the products formed by reaction of lysine with glyoxal (a product from the oxidation of sugars). MOLD (methylglyoxal lysine dimer) and Nε-CEL are the analogous products formed from methylglyoxal (a product from the fragmentation of sugars). Norfuranol is formed from the amine-catalysed rearrangement and dehydration of pentose sugars.
Figure S2  $^{13}$C solution-state NMR spectra of incubations of ribose with (a) spermidine after 30 days (green), (b) AcLys after 9 days (purple), (c) PLL after 11 days (red) and (d) no amine present (blue)

In each case the peaks at δ 69.2, 71.5 (multiple), 74.8, 75.4, 95.5 and 99.7 ppm which are not present in a control incubation of ribose in the absence of amine are diagnostic of formation of arabinose, the C2-epimer of ribose, with the assignments given.
Figure S3  $^{13}$C solution-state NMR spectra of the reaction between [U-$^{13}$C]ribose and PLL with data taken over a period of 23 days

The broad signal at 209 ppm, assigned to the carbonyl carbon of ribuloselysine, the Amadori product, reaches a maximum after 3 days before slowly decaying, although it is still present after 23 days suggesting the glycation process had not yet gone to completion. The broadness of the signal is presumably attributable to the polymeric nature of PLL: each ribuloselysine entity is in a slightly different chemical environment. The multiplet (doublet of doublets of doublet) at 202 ppm couples to 136 ppm and is from norfuraneol. The multiplet around 215 ppm is from relatively stable ketone intermediates; a corresponding feature is also present in reactions of R5P.
Figure S4  $^{31}$P solution-state NMR of ADPR before (blue) and after (red) reaction with PLL.

The ADPR signal loses intensity and is replaced with two doublets (asterisked) at $-7.3$ and $-10.2$ ppm ($J = 20$ Hz) which correspond to the $\beta$- and $\alpha$-phosphate groups of free ADP released from ADPR, respectively.
Figure S5  $^{13}$C solution-state NMR of ADPR before (blue) and after (red) reaction with PLL. ADPR assignment is given. The chemical shifts of C$_1$–C$_5$ of the ribosyl phosphate moiety are dependent on the configuration of the anomeric carbon, C$_1$, and accordingly signals from the $\alpha$-anomer (hydroxyl group ‘down’) and the $\beta$-anomer (hydroxyl group ‘up’) are differentiated. After the generation of a precipitate (see text for details), the signals from the free ribose, asterisked, are significantly reduced, whereas the signals from ADP are unchanged in intensity. Taken with Figure S4, these results confirm that the ADP is released during the reaction that leads to crosslinking and precipitation. The same signal depletion over time was observed for glycation of PLL by R5P.
Figure S6 $^{31}$P ssNMR of R5P before (red) and after (blue) reaction with PLL

The $^{31}$P chemical shift changes from the phosphomonoester shift of 6.1–2.3 ppm of free phosphate ion. This further demonstrates that phosphate acts as a leaving group in the Maillard reactions of R5P.
Figure S7  $^{13}\text{C}$ solution-state NMR spectrum of the supernatant phosphate buffer from [U-$^{13}\text{C}$]ribose glycated collagen (blue) and that of unreacted [U-$^{13}\text{C}$]ribose (red)

The dominant species in the supernatant is unreacted ribose, but significant quantities of glycolic and oxalic acids were also observed (signals indicated, see Section 5.1.2), along with a multitude of weaker signals from minor AGEs which require further assignment.
Novel AGE, and protein crosslink characterization, by NMR

Figure S8  Comparison of the $^{13}$C NMR spectrum of the supernatant from [U-$^{13}$C]ribose glycated collagen (blue) and the model system: [U-$^{13}$C]ribose glycated PLL (red)

Several weak signals are observed in both spectra, such as the signal due to CML at 52 ppm. However, many of the supernatant AGE signals are not observed in the model system, showing that PLL is an incomplete model for collagen glycation. The lack of norpronyl-lysine in the supernatant is expected as it is attached to the solid collagen. The absence of norfuraneol and acetic acid is surprising: either they are not generated in observable quantities, or bind strongly to collagen, or had degraded by the time the spectrum was obtained.

Table S1  Shifts of C$_1$ and C$_2$ in carboxylic acids are highly pH and concentration dependent, with carboxylate ion C$_1$ typically resonating 4–5 ppm to higher frequency than the corresponding acid [15]

Values in the table are literature values for the acid form [16–18]. Formate resonated at 173.8 ppm, acetate at 184.1 and 26.1 ppm, glycolate at 182.6 and 62.1 ppm, oxalate at 163.0 ppm. Glyoxylate was not observed. These assignments were confirmed by peak multiplicity, J values, and $^{13}$C–$^{13}$C COSY correlations, in isotopically enriched experiments.

<table>
<thead>
<tr>
<th></th>
<th>C$_1$ (ppm)</th>
<th>C$_2$ (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formic acid</td>
<td>167</td>
<td></td>
</tr>
<tr>
<td>Acetic acid</td>
<td>178</td>
<td>22</td>
</tr>
<tr>
<td>Glycolic acid</td>
<td>177</td>
<td>60</td>
</tr>
<tr>
<td>Glyoxylic acid</td>
<td>174</td>
<td>90</td>
</tr>
<tr>
<td>Oxalic acid</td>
<td>161</td>
<td>161</td>
</tr>
</tbody>
</table>
Table S2 A summary of the carboxylic acids generated in the Maillard reactions between model amines and sugars
A tick indicates that the species is appreciably present, a tick in parentheses that it is weak, a cross that there is no evidence of the species.

<table>
<thead>
<tr>
<th>Reactants</th>
<th>Formic acid</th>
<th>Acetic acid</th>
<th>Glycolic acid</th>
<th>Glyoxylic acid</th>
<th>Oxalic acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ribose and PLL</td>
<td>√</td>
<td>×</td>
<td>√</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>[U-13C]-ribose and PLL</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>×</td>
<td>(√)</td>
</tr>
<tr>
<td>R5P and PLL</td>
<td>(√)</td>
<td>×</td>
<td>√</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>[U-13C]-ribose and collagen</td>
<td>(√)</td>
<td>(√)</td>
<td>√</td>
<td>×</td>
<td>(√)</td>
</tr>
<tr>
<td>Ribose and AcLys</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>[U-13C]-ribose and AcLys</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>×</td>
<td>(√)</td>
</tr>
<tr>
<td>R5P and AcLys</td>
<td>√</td>
<td>×</td>
<td>×</td>
<td>×</td>
<td>×</td>
</tr>
<tr>
<td>Ribose and spermidine</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>×</td>
<td>×</td>
</tr>
</tbody>
</table>

Table S3 Full $^{13}$C solution-state NMR assignment of $^{13}$C$_5$-norfuraneol as determined from its observation as a product of the reaction between [U-13C]ribose and PLL
The assignment of C$_5$ was aided by the detection of norfuraneol as a product of reaction between ribose and spermidine, where its NMR signal did not overlap with others.

<table>
<thead>
<tr>
<th>Carbon atom</th>
<th>δ (ppm)</th>
<th>Multiplicity</th>
<th>J (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15.9</td>
<td>D</td>
<td>49</td>
</tr>
<tr>
<td>2</td>
<td>183.5</td>
<td>Ddd</td>
<td>88, 49, 14</td>
</tr>
<tr>
<td>3</td>
<td>136.1</td>
<td>Dddd</td>
<td>88, 65, 20, 3</td>
</tr>
<tr>
<td>4</td>
<td>201.5</td>
<td>Ddd</td>
<td>65, 43, 14</td>
</tr>
<tr>
<td>5</td>
<td>76.5</td>
<td>Dd</td>
<td>43, 20</td>
</tr>
</tbody>
</table>
REFERENCES


Received 11 December 2013/15 January 2014; accepted 29 January 2014
Published as Immediate Publication 12 February 2014, doi 10.1042/BSR20130135