Characterization of Glycated Proteins by $^{13}$C NMR Spectroscopy

Carolyn I. Neglia
Helga J. Cohen
Albert R. Garber
Suzanne R. Thorpe
John W. Baynes

University of South Carolina - Columbia, john.baynes@sc.edu

Follow this and additional works at: http://scholarcommons.sc.edu/chem_facpub

Part of the Chemistry Commons

Publication Info

This Article is brought to you for free and open access by the Chemistry and Biochemistry, Department of at Scholar Commons. It has been accepted for inclusion in Faculty Publications by an authorized administrator of Scholar Commons. For more information, please contact SCHOLARC@mailbox.sc.edu.
Characterization of Glycated Proteins by $^{13}$C NMR Spectroscopy

IDENTIFICATION OF SPECIFIC SITES OF PROTEIN MODIFICATION BY GLUCOSE

(Received for publication, September 13, 1984)

Carolyn I. Neglia‡, Helga J. Cohen‡, Albert R. Garber‡, Suzanne R. Thorpe‡, and
John W. Baynes§§

From the ‡Department of Chemistry and §§School of Medicine, University of South Carolina, Columbia, South Carolina 29208

$^{13}$C NMR spectroscopy has been used to characterize Amadori (ketoamine) adducts formed by reaction of [2-$^{13}$C]glucose with free amino groups of protein. The spectra of glycated proteins were acquired in phosphate buffer at pH 7.4 and were interpreted by reference to the spectra of model compounds, $N^\alpha$-formyl-$N^\beta$-fructose-lysine and glycated poly-L-lysine (GlcPLL). The anomic carbon region of the spectrum (approximately 90–105 ppm) of glycated cytochrome $c$ was superimposable on that of $N^\alpha$-formyl-$N^\beta$-fructose-lysine, and contained three peaks characteristic of the $\alpha$- and $\beta$-furanose and $\beta$-pyranose anomers of Amadori adducts to peripheral lysine residues on protein (pK$_a$ ~ 10.5). The spectrum of GlcPLL yielded six anomic carbon resonances; the second set of three was displaced about 2 ppm to lower shielding of the first and was assigned to the Amadori adduct at the $\alpha$-amino terminus (pK$_a$ ~ 7.5). The spectrum of glycated RNase was similar to that of GlcPLL, but contained a third set of three signals attributable to modification of active site lysine 41 (pK$_a$ ~ 8.8). The assignments for RNase were confirmed by analysis of spectra taken at pH 4 and under denaturing conditions. The spectrum of glycated hemoglobin was comparable to that of GlcPLL, and distinct resonances could be assigned to Amadori adducts at amino-terminal valine and intra-chain $N^\beta$-lysine residues. Chemical analyses were performed to measure the relative extent of $\alpha$- and $\gamma$-amino group modification in the glycated macromolecules, and the results were compared with estimates based on integration of the NMR spectra.

Nonenzymatic glycosylation or glycation is a stable post-translational modification of protein which occurs in vivo by direct chemical reaction between glucose and the primary amino groups of protein (1, 2). The initial product is a labile Schiff base adduct (Fig. 1), which undergoes a slow Amadori rearrangement to a stable ketoamine derivative of the protein. Glycation is considered to be the first step in a complex series of browning or Maillard reactions of protein which occur in vivo.

In the present study we have increased the resolution and the signal/noise ratio in the anomic region by incorporating [2-$^{13}$C]glucose (99 atom % enriched) into protein, specifically enhancing the intensities of the anomic carbons of the Amadori rearrangement products. The complex spectrum obtained with [2-$^{13}$C]GlcRNase has been interpreted by reference to the spectra of other glycated macromolecules (cytochrome $c$ and poly-L-lysine), as well as by evaluation of changes in chemical shift of the anomic carbons as a function of buffer, pH, and denaturation of the protein. The information obtained has been used to interpret the spectrum of glycated hemoglobin (GlcHb), a protein which exists in a glycated form in vivo (1, 2, 5, 6).

EXPERIMENTAL PROCEDURES

Materials—RNase A (Type XII-B, $M_r = 13,680$), poly-L-lysine ($M_r$ ~ 14,000), horse heart cytochrome $c$ (Type VI, $M_r = 13,384$), and Na$_3$H$_6$L$_2$CN were purchased from Sigma. Human hemoglobin was prepared from fresh citrated blood by the procedure of Drabkin (9). D-$[6-^{2}$H]Glucose and $[^{3}$H]sodium borohydride were obtained from New England Nuclear. D-$[2-^{13}$C]Glucose (99 atom % enriched) was

* This work was supported by United States Public Health Service Research Grants AM-19171 and 25373. The Nuclear Magnetic Resonance Center at the University of South Carolina is supported by National Science Foundation Grant CHE-82-07445. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† Recipient of United States Public Health Service Research Career Development Award AM-00931. To whom reprint requests should be addressed: Department of Chemistry, University of South Carolina, Columbia, SC 29208.

‡ The abbreviations used are: RNase A, bovine pancreatic ribonuclease A; fPLL, $N^\alpha$-formyl-$N^\beta$-fructose-lysine; GlcRNase, glycated ribonuclease; GlcPLL, glycated poly-L-lysine; GlcCytc, glycated cytochrome $c$; GlcHb, glycated hemoglobin.
were recorded for the GlcRNase spectrum and 250,000 transients for to suppress the nuclear Overhauser effect. A total of 28,000 transients

were reduced with 100-fold molar excesses of NaB$_3$H$_4$ (specific activity = 333 Ci/mol) in phosphate buffer, pH 8.0–8.5, using the procedure described previously (10). Excess borohydride was discharged by addition of acetic acid, except for the GlcHb reduction. In this case, borohydride was destroyed and heme extracted simultaneously by addition of 20 volumes of 0.25 M HCl in acetic acid with vortexing; the globin was isolated by centrifugation (11). Proteins were then hydrolyzed in 6 N HCl at 95 °C for 18 h (12). The hydrolysates were chromatographed on a Waters high pressure liquid chromatography amino acid analyzer by cation exchange chromatography using a continuous gradient sodium citrate buffer system. Glycated amino acids were detected by scintillation counting using a Radiomatic FLO-ONE radioactive flow detector with FLO-SCINT III scintillation fluid (10:1, v/v).

RESULTS AND DISCUSSION

NMR Spectrum of GlcRNase—The anomeric regions of the spectra of the Amadori compound, fFL, and of [U-13C]GlcRNase are shown in Fig. 2, A and B (adapted from Ref. 8). Distinct resonances were observed for the anomeric carbon of each of the four sugar conformations of fFL: α-furanose (101.7

Fig. 1. General reaction scheme for glycation of protein.

from Omicron Biochemicals. Reagents for assay of glucose by the glucose oxidase-peroxidase assay were purchased from Amresco.

N'-Formyl-N'-fructosyl-lysine (fFL) was prepared from N'-formyl-
lysine and glucose, as previously described (8). N'- and N''-glucitol-
lysine, N''- and N'-mannitol-lysine, and glucitol-valine were prepared by reacting the appropriate t-butoxycarbonyl-lysine derivatives or valine with sugars in the presence of sodium cyanoborohydride. The t-butoxycarbonyl group was removed by hydrolysis in 2 N HCl for 1 h at 95 °C, and the products were purified by cation exchange chromatography (8).

Glycation of Proteins—Glycated proteins were prepared essentially as described previously (8). Briefly, except for Hb, proteins (140 mg/ml) were incubated under sterile conditions with 0.25 mM [2-13C]glucose in 0.2 M phosphate buffer, pH 7.4, at 37 °C. Hb was saturated with carbon monoxide and dialyzed against 0.2 M phosphate buffer, pH 7.4. The solution was concentrated to 200 mg/ml by ultrafiltration (Amicon, YM-10 membrane), then re-gassed with carbon monoxide, and incubated at 200 mg/ml with 0.5 mM [2-13C]glucose under a carbon monoxide atmosphere. The extent of glycation of the proteins was estimated from parallel incubations containing radioactive glucose purified by high pressure liquid chromatography. Incubation times required for preparation of the various compounds were: GlcRNase (1.2 mol/mol), 6 days; GlcPLL (6 mol/mol), 6 days; GlcCyt (1.3 mol/mol), 5 days; and GlcHb (0.96 mol/mol), 3 days. After incubation, protein solutions were dialyzed overnight against 0.2 M phosphate buffer at 4 °C and concentrated approximately 20% by rotary evaporation; the GlcHb solution was concentrated by ultrafiltration (Amicon YM-10 membrane). D$_2$O (20% v/v) was then added prior to NMR spectroscopy.

13C NMR Spectroscopy—Spectra were obtained at the National Science Foundation Southeastern Regional NMR Center located at the University of South Carolina, using a Bruker WH-400 NMR spectrometer operating at 100.6 MHz, equipped with quadrature phase detection and interfaced with an Aspect 2000A computer. Experimental conditions are described in the figure legends and, as noted, acquisition parameters and treatment of data were as previously reported (8). Chemical shifts were referenced to external dioxane which resonated 66.5 ppm to lower shielding from tetramethylsilane. Quantitative data to determine the relative extent of modification of α- and ε-amino groups of proteins (described in Table II) were obtained with a 30° radiofrequency pulse and 0.5-sec acquisition time. A 0.7- or 1-s recycle time was used with gated proton decoupling to suppress the nuclear Overhauser effect. A total of 28,000 transients were recorded for the GlcRNase spectrum and 250,000 transients for the GlcHb spectrum.

Reduction and Amino Acid Analysis of Proteins—Glycated proteins

were recorded for the GlcRNase spectrum and 250,000 transients for the GlcHb spectrum.
The spectrum of GlcPLL (Fig. 2B) was similar to that of fFL but not as well resolved and contained an unidentified signal at 98.1 ppm which was not present in the spectrum of fFL. In contrast to these results obtained with [U-13C]glucose (14.5 atom % enriched), 9 distinct resonance signals were clearly resolved in the anomeric region of the spectrum of [2-13C]GlcRNase (99 atom % enriched) (Fig. 2C). The complexity of this spectrum suggested the possibility that distinct signals, or sets of signals, were being generated by glucose adducts to nonequivalent amino groups in RNase, for example, peripheral ε-amino groups (pK,a ~10.5), the α-amino group (pK,a ~7.5), and active site lysine 41 (pK,a ~8.8) (13, 14). To address this question we prepared two additional model compounds, glycated poly-L-lysine and cytochrome c. Poly-L-lysine was selected because it exists as a random coil in solution at pH 7.4 and contains only two types of nonequivalent primary amino groups, the α-amino group and ε-amino groups. Cytochrome c was chosen because it has a blocked, acetylated α-amino-terminal glycine residue which is unavailable for glycation (15). The spectrum of the anomeric region of GlcPLL, shown in Fig. 2D, was similar to, but less complex than, the spectrum of GlcRNase, while the spectrum of GlcCyt (Fig. 2E) was almost identical to that of the model compound, fFL.

The complex spectrum of GlcRNase can now be interpreted as the sum of three sets of three resonances, each set assignable to modification of different types of amino groups in RNase. The α- and β-furanose and β- pyranose signals are clearly resolved in each set, although the α-pyranose resonance is obscured in most of the spectra by the broad β-pyranose signal. The first set of signals, identified by the letter ε in the spectrum of GlcCyt, is attributed to the anomeric carbon of Amadori adducts to peripheral lysine residues on protein (pK,a ~10.5). In GlcRNase this set may include 9 of its 11 amino groups, but excludes at least the α-amino terminus and the low pK,a ε-amino group of lysine 41. The additional set of peaks identified by the letter α in the spectrum of GlcPLL, but not seen in GlcCyt, is assigned to the anomeric carbons of adducts to the α-amino terminus. The third set of signals, marked with an asterisk in the spectrum of GlcRNase, resonates between the corresponding anomeric signals of the other sets and results from modification of lysine 41 in the active site of RNase. The chemical shift of these anomeric carbons is clearly dependent on the pK,a, i.e., extent of protonation, of the amino group modified in the protein. It can be seen, for example, that the location of the β-pyranose signals in each set (in order, right to left from higher to lower shielding: peripheral ε-amino, active site ε-amino, α-amino) correlates with their respective pK,a (highest to lowest) in RNase. To exclude the possibility that the chemical shift of the Amadori adduct to lysine 41 was influenced by phosphate bound noncovalently in the active site of the protein, GlcRNase was dialyzed exhaustively against Tris buffer, pH 7.4. The spectrum of GlcRNase in Tris buffer, shown in Fig. 3A, was essentially the same as that obtained in phosphate buffer. Thus, it is the pK,a of lysine 41 rather than the presence of phosphate in the active site which affects the chemical shift of the anomeric carbon.

To confirm our assignments for the spectrum of GlcRNase, we also studied the effects of denaturation and titration of the protein. The spectrum of GlcRNase in 8 M urea and 0.1 M dithiothreitol is shown in Fig. 3B. Denaturation was apparent from the full spectrum of the protein (data not shown), in which the carbon signals of the amino acids were significantly sharpened. Because denaturation of the enzyme and unfolding of its active site, the signals assigned to the low pK,a lysine 41 disappeared from the spectrum. The resultant spectrum contains only the resonances assigned to α-amino and peripheral ε-amino groups and is superimposable on that of GlcPLL (Fig. 2D). The spectrum shown in Fig. 3C is obtained on titration of this same GlcRNase solution to pH 4.0. This spectrum is identical to the spectra of fFL and GlcCyt (Fig. 2, A and E), reflecting the fact that all of the amino groups of RNase are almost completely protonated at pH 4.

The information obtained from the above studies with model proteins was then applied for the analysis of the spec-
**13C NMR Studies of Glycated Proteins**

**Fig. 4.** Proton-decoupled 13C NMR spectrum of GlcHb. Hb, liganded with carbon monoxide, was glycated in vitro to 1 mol of [2-13C]glucose/mol of Hb (2.5 mM in tetramer). Inset shows the anomeric region of the spectrum, with identification of the signals from adducts at Nα-valine and Nε-lysine residues.

**Table I**

Chemical shifts of anomeric carbons of glycated proteins and model compounds

<table>
<thead>
<tr>
<th>Compound</th>
<th>Conformation</th>
<th>α-Fructo</th>
<th>β-Fructo</th>
<th>α-Fructo</th>
<th>β-Fructo</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>furanose</td>
<td>furanose</td>
<td>pyranose</td>
<td>pyranose</td>
</tr>
<tr>
<td>Nα-amino adducts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(pKα = 7.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nα-GlcPLL</td>
<td>103.7</td>
<td>101.4</td>
<td>—*</td>
<td>97.8</td>
<td></td>
</tr>
<tr>
<td>Nα-GlcRNase</td>
<td>103.9</td>
<td>101.4</td>
<td>—*</td>
<td>97.7</td>
<td></td>
</tr>
<tr>
<td>Nα-GlcHb</td>
<td></td>
<td>100.6</td>
<td>—*</td>
<td>97.0</td>
<td></td>
</tr>
<tr>
<td>Nα-amino adducts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(pKα = 10.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fFL</td>
<td>101.7</td>
<td>98.7</td>
<td>96.2</td>
<td>95.4</td>
<td></td>
</tr>
<tr>
<td>Nε-GlcCyt</td>
<td>102.1</td>
<td>99.1</td>
<td>96.3*</td>
<td>95.7</td>
<td></td>
</tr>
<tr>
<td>Nε-GlcPLL</td>
<td>101.9</td>
<td>98.8</td>
<td>96.3*</td>
<td>95.5</td>
<td></td>
</tr>
<tr>
<td>Nε-GlcRNase</td>
<td>101.9</td>
<td>98.9</td>
<td>95.8*</td>
<td>95.6</td>
<td></td>
</tr>
<tr>
<td>Nε-GlcHb</td>
<td>101.5</td>
<td>98.6</td>
<td>95.7*</td>
<td>95.2</td>
<td></td>
</tr>
<tr>
<td>Nε-GlcRNase (pKε = 8.8)</td>
<td>102.8</td>
<td>99.9</td>
<td>—*</td>
<td>96.1</td>
<td></td>
</tr>
</tbody>
</table>

* Discrete signal not observed with 1 Hz line broadening.
* Peak not assignable because of signal/noise limitations.
* Assignments based on processing data with 1 Hz line broadening.
* Assignment obtained from denatured sample (see Fig. 3B).

**Table II**

Comparison of sites of glycation of polypeptides: results of NMR versus chemical measurements

<table>
<thead>
<tr>
<th>Compound</th>
<th>Ratio α:ε glycation</th>
<th>NMR spectroscopy*</th>
<th>Chemical analysis*</th>
</tr>
</thead>
<tbody>
<tr>
<td>GlcPLL</td>
<td>1.3:6</td>
<td>1.3:8</td>
<td></td>
</tr>
<tr>
<td>GlcRNase</td>
<td>1.3:3*</td>
<td>1.6:0*</td>
<td></td>
</tr>
<tr>
<td>GlcHb</td>
<td>1.3:5</td>
<td>1.3:1</td>
<td></td>
</tr>
</tbody>
</table>

* NMR data were acquired as described under "Experimental Procedures." Ratios of α:ε lysine substitution were determined by weighing β-pyranose peaks cut from expanded spectra.
* Ratios calculated from dpm in peak fractions of radiochromatograms shown in Fig. 5.

The NMR spectra can be integrated to obtain an estimate of the relative extent of modification of the various types of amino groups in proteins or polypeptides. Nuclear Overhauser effect suppression did not significantly affect the relative intensities of the signals from the anomeric carbons of GlcRNase, but did have a slight effect on the GlcHb spectrum (Fig. 4). In neither case were relative signal intensities affected by increasing the relaxation delay. The results of integration...
of the NMR spectra of GlcRNase, GlcPLL, and GlcHb are summarized in Table II. These data are compared to the results of radiochemical measurements of the heptitol-amino acids obtained on hydrolysis of NaB³H₄-reduced proteins. Typical radiochromatograms, shown in Fig. 5, yield base-line separation of the α- and ε-hexitol-lysines and partial resolution between their glucitol and mannitol isomers. Glucitol- and mannitol-valine are not resolved and elute as a single heptitol-valine peak at about 9 min into the gradient. When the integrated peak areas were corrected for quench to obtain product ratios, the results of chemical and NMR integrations were in good agreement, except for GlcRNase (Table II). For GlcHb the results are also in good agreement with chemical analyses of Shapiro et al. (19) who reported a 3:1 ratio of lysine to valine modification for Hb glycated in vitro. The difference between the NMR and radiochemical analyses of GlcRNase was reproducible with separate preparations of GlcRNase. The reason for this discrepancy is not clear; however, data from the analysis of individual peptides isolated from GlcRNase reveals a strong signal for glycated W-lysine 41, and both N"-lysine 1 and lysine 41 appear to be preferred sites for glycation of RNase.

Overall, it appears that analysis of the 13C NMR spectra of glycated proteins should be useful for detecting the glycation of unique types of amino groups in protein and for assessing the relative reactivity of the different types of primary amino groups in protein with glucose. Thus, resonances for glucose adducts at the ε-amino terminal of these peptide chains is a preferred site for reaction with glucose. Similarly, with attention to sample pH, it should be possible to detect adducts to α-amino groups with unique pKₐ values, e.g. in the active or allosteric sites of proteins. Thus, the NMR spectrum in the anomeric region of GlcRNase (Fig. 2C) reveals a strong signal for glycated N"-lysine 41, and both N α-lysine 1 and N"-lysine 41 appear to be preferred sites for glycation of RNase.

Peptide mapping of GlcRNase also supports the conclusion that these two residues are preferred sites of reaction with glucose. Further studies are in progress to determine if these sites are also preferentially reactive in subsequent Maillard reactions and to characterize the products formed during the aging of glycated proteins.

REFERENCES