Published in final edited form as:
J Proteome Res. 2008 July ; 7(7): 2756–2768. doi:10.1021/pr700874a.

Identification of Glucose-Derived Cross-Linking Sites in Ribonuclease A

Zhenyu Dai[†], **Benlian Wang**[‡], **Gang Sun**[§], **Xingjun Fan**[||], **Vernon E. Anderson**[†], and **Vincent M. Monnier**[*,†,||]

Department of Biochemistry, Case Western Reserve University, Cleveland, Ohio 44106, Center for Proteomics and Mass Spectrometry, Case Western Reserve University, Cleveland, Ohio 44106, Department of Internal Medicine, Washington University, St. Louis, Missouri 63110, and Department of Pathology, Case Western Reserve University, Cleveland, Ohio 44106

Abstract

The accumulation of glycation derived cross-links has been widely implicated in extracellular matrix damage in aging and diabetes, yet little information is available on the cross-linking sites in proteins and the intra- versus intermolecular character of cross-linking. Recently, glucosepane, a 7-membered heterocycle formed between lysine and arginine residues, has been found to be the single major cross-link known so far to accumulate during aging. As an approach toward identification of glucose derived cross-linking sites, we have preglycated ribonuclease A first for for 14 days with 500 mM glucose, followed by a 4-week incubation in absence of glucose. MALDI-TOF analysis of tryptic digests revealed the presence of Amadori products ($\Delta m/z$ = 162) at K1, K7, K37 and K41, in accordance with previous studies. In addition, K66, K98 and K104 were also modified by Amadori products. Intramolecular glucosepane cross-links were observed at K41-R39 and K98-R85. Surprisingly, the only intermolecular cross-link observed was the 3-deoxyglucosone-derived DODIC at K1-R39. The identity of cross-linked peptides was confirmed by sequencing with tandem mass spectrometry. Recombinant ribonuclease A mutants R39A, R85A, and K91A were produced, purified, and glycated to further confirm the importance of these sites on protein cross-linking. These data provide the first documentation that both intramolecular and intermolecular cross-links form in glucose-incubated proteins.

Keywords

glycation; ribonuclease A; cross-linking; glycation sites; glucosepane; DODIC; AGEs

Introduction

The Maillard Reaction is initiated by reaction of a reducing sugar with a primary amine to form an unstable Schiff base, followed by its rearrangement to relatively stable Amadori products. These can undergo fragmentation, oxidation and rearrangements resulting in a plethora of compounds named advanced glycation end products (AGEs). In the past 30 years, the quantitative relationship between accumulation of AGEs in tissues and the extent of tissue pathology has been extensively studied in animals as well as in humans. AGE-like fluorescence

* To whom correspondence should be addressed: Vincent M. Monnier, Department of Pathology, Case Western Reserve University, Cleveland, OH 44120. Tel: (216) 368-6238. E-mail: vmm3@cwru.edu.
[†]Department of Biochemistry, Case Western Reserve University.
[‡]Center for Proteomics and Mass Spectrometry, Case Western Reserve University.
[§]Department of Internal Medicine, Washington University.
[||]Department of Pathology, Case Western Reserve University.

(excitation at 370 nm, emission at 440 nm) has been shown to increase in collagen from aged and diabetic patients.[1] In retinal vessels of diabetic rats, AGE-specific fluorescence increased 2.6-fold after 26 weeks of diabetes.[2] Similar increases in AGE fluorescence have been found in proteins from diabetic lens[3] and renal cortex.[4] AGE accumulation was accompanied by histological evidence of diabetic tissue damage. The senescent extracellular matrix exhibits characteristics including decreased solubility,[5] decreased proteolytic digestibility,[6,7] and accumulation of yellow and fluorescent material,[8] all of which are commonly associated with diabetes. These age-related and diabetes-enhanced changes are thought to result in part from AGE-derived cross-links.

The accumulation of AGEs on proteins depends on reducing sugar concentration and the protein turnover rate. The low turnover rate of collagen, lens protein and neural proteins predisposes them to AGE accumulation. Changes in collagen-rich tissues, such as arteries, lungs, and joints, have been correlated with hypertension, emphysema, and decreased joint mobility.[9]

In collagen, cross-links derived from lysyl oxidase activity have been unequivocally proven to be involved in determining the mechanical character of immature collagen. [10] In contrast, the glycation derived cross-links have been hypothesized, but not proven, to be responsible for the collagen stiffening during aging and in diabetes. The type of cross-links formed, that is, inter- versus intrafibrillar cross-links, could have a very different mechanical impact on the tissue properties.

Few studies have been done on the glycation-derived cross-linking sites of proteins in spite of their potential importance on the mechanical character. One of the reasons for the slow progress in this field may be linked to the low level of individual AGEs found *in vivo*. Biemel et al. provided convincing evidence that glucosepane is the dominant cross-link in both human serum albumin and lens protein[11] (Figure 1). Our laboratory provided additional data supporting the premise that glucosepane is the single major glycation derived cross-link known to date in aging human skin and glomerular basement collagen.[12] In addition, its levels are dramatically elevated in diabetes,[13] a condition that is associated with increased stiffness of all collagen-rich tissues. Glucosepane levels increased up to ~2 nmol/mg collagen in old nondiabetic controls, and further increased to ~4.3 nmol/mg in diabetic patients.[12] These numbers can be translated into one cross-link for every 2–5 triple helical collagen molecules in diabetic and nondiabetic aged controls, respectively. This level of modification might contribute to the accumulation of collagen matrix due to impaired proteolytic digestion in diabetes and aging, and perhaps even increased matrix stiffening, particularly if the cross-links are intermolecular or interfibrillar.

Because of the highly repetitive structure of the helical domain of the type I collagen molecule, and the anticipated difficulty in making unequivocal structural assignments, we opted to first approach the problem of site specificity using ribonuclease A, a protein that has been widely used in glycation studies.

Experimental Procedures

Reagents

RNase A from bovine pancreas (Type XII-A, ≥90% SDS-PAGE purity) and D-($^{13}C_6$)glucose were from Sigma (St. Louis, MO). Chelex 100 Resin was from Bio-Rad (Hercules, CA). Deionized water (18.2 MΩ cm^{-1}) was used for all experiments. Sequencing grade trypsin was purchased from Promega (Madison, WI). Chymotrypsin and endoproteinase Asp-N were from Roche Applied Science (Indianapolis, IN). All other reagents were obtained in the highest quality available from Sigma (St Louis, MO), unless indicated otherwise.

Preparation of Glycated RNase A

Bovine RNase A at a concentration of 50 mg/mL was incubated with a 1:1 mixture of 250 mM D-glucose and 250 mM D-($^{13}C_6$)glucose in 100 mM deaerated, Chelex-treated sodium phosphate buffer containing 1 mM diethylenetriamine pentaacetic acid (DTPA) at 37 °C for 2 weeks.[14] All *in vitro* incubations were in these "anaerobic" conditions. Glucose was removed by dialysis against 4 L of the same buffer for 2 days with one buffer change. Spectra/Por 7 dialysis tubing (Spectrumlabs, Rancho Dominguez, CA, 2 kDa MWCO) was used. Upon determination of protein concentration, this "preglycated" preparation[15] was further incubated with an equal amount of freshly added native RNase A for 4 weeks at 37 °C. Chloroform and toluene (0.15% each, v/v) were added to prevent bacterial growth during the incubation.

Enzymatic Digestion for Peptide Mapping

Prior to enzymatic digestion, protein samples (20 μL) were diluted to 500 μL with water and concentrated to 20 μL using an Ultrafree–0.5 centrifugal filter unit (Millipore, Billerica, MA, 5 kDa MWCO). Per manufacturer's information, three cycles of concentration should remove ~99% of the initial salt content. The resulting protein sample was subjected to in-solution digestion by trypsin in the presence of RapiGest SF (Waters, Milford, MA). The ratio of enzyme to protein substrates was 1:100 and the digestion was carried out in a 50 mM NH_4HCO_3 buffer (pH 8.0) at 37 °C for 3 h. The ultracentrifuged protein was also subjected to digestion by chymotrypsin and endoproteinase Asp-N. Chymotryptic digestion was carried out in 100 mM Tris-HCl, 10 mM $CaCl_2$, pH 7.8, whereas the Asp-N digestion buffer was 50 mM Tris-HCl, pH 8.0. In both digestions, the ratio of enzyme to protein substrates was around 1:100 and the digestion was carried out at 37 °C (endoproteinase Asp-N) or at 25 °C (chymotrypsin) overnight. Before digestion, protein samples were reduced by 5 mM dithiothreitol (60 °C, 30 min) and alkylated with 15 mM iodoacetamide (30 min in the dark).

Matrix-Assisted Laser Desorption and Ionization Mass Spectrometry (MALDI)

The digested peptides were analyzed by using a prOTOF 2000 MALDI O-TOF mass spectrometer (PerkinElmer, Inc.). The prOTOF mass spectrometer typically has a mass accuracy of better than 10 ppm. Samples were mixed with an equal volume of α-cyano-4-hydroxycinnamic acid (Fluka, St. Louis, MO) matrix solution (10 mg/mL in 50% acetonitrile, 0.1% TFA). Typically, 1 μL of the mixture was applied onto the laser target probe and was air-dried before being introduced into the mass spectrometer. For identification of possible glycation modified peptides, a database including all possible *m*/*z* values of Amadori product, glucosepane, and DODIC modified peptides was generated and used to compare with the isotopic doublet or triplet peaks found in the MALDI experiment. Laser-desorbed positive ions were analyzed after acceleration by 19 kV in the reflectron mode for the peptide digest. Other instrument settings were the following: laser energy at 65%; laser rate, 100.0 Hz; declustering potential, 25.0 V; cooling flow, 200 mL/min; mass range, 600.0–5000.0 Da; focus flow, 200.0 mL/min.

Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) Analysis

LC-MS/MS analyses of the proteolytic digests were performed using a quadrupole ion trap mass spectrometer (model LTQ) from Thermo-Finnigan (San Jose, CA) coupled with an Ettan MDLC system (GE Healthcare, Piscataway, NJ), chromatographed with a gradient of 0–60% acetonitrile/0.1% formic acid for 30 min. The spectra were acquired by data-dependent methods, consisting of a full scan and then MS/MS on the six most abundant precursor ions at a collision energy of 35%. The previously selected precursor ions were repeated twice during 45 s and then were excluded for 180 s. The data were submitted to a BioWorks Rev. 3.3 database search. Further interpretation of the tandem mass spectra of the modified peptides was assisted

by an Excel spreadsheet that generated predicted fragment ions from glycated peptides available from V. Anderson.

Mutagenesis of RNase A

A synthetic gene for bovine RNase A was a gift from the Genex Corp. This gene was placed in M13mp18 (United States Biochemical).[16] RNase A in M13mp18 was subcloned into pET22b(+) (Novagen) between the *Msc*I and *Hind*III restriction enzyme sites. Mutant plasmids for the expression of mutant RNase A, R39A, R85A, and K91A were constructed with a Quickchange Site-Directed Mutagenesis Kit (Stratagene Cloning Systems). The sequence of the primers designated to replace the Arg39 codon with Ala codon was 5′-CT AGA AAC TTG ACC AAG GAC GCA (for Ala) TGT AAG CCA GTT AAC ACA T-3′. The sequence of the primers designated to replace the Arg85 codon with Ala codon was 5′-G TCC ATC ACT GAC TGT GCT (for Ala) GAG ACA GGC TCG AGT-3′. The sequence of the primers designated to replace the Lys91 codon with Ala codon was 5′-CGT GAG ACA GGC TCG AGT GCG (for Ala) TAT CCT AAT TGT GCT TAC-3′. Mutagenesis was confirmed by DNA sequencing at Biotic Solutions, Inc. The mutagenized plasmids were transformed into XL1-Blue super-competent cells. The plasmids containing RNase A mutants R39A, R85A, K91A were transformed into the *Escherichia coli* cell line BL21-Gold(DE3)pLysS (Stratagene) which was stored at −70 °C.

Production and Purification of Wild-Type and Mutant RNase A

The frozen cells were plated on a 50 μg/mL ampicillin and 34 μg/mL chloramphenicol agar plate, and the cells were grown overnight at 37 °C. A single colony of BL21-Gold(DE3)pLysS containing an alanine mutant in pET22b(+) from the plate was transferred to 50 mL of 2× YT LB medium with 50 μg/mL ampicillin and 34 μg/mL chloramphenicol at 37 °C and agitated with a gyratory shaker overnight at 225 rpm. This culture was then diluted to 1 L of the same medium and shaken under the same conditions until the optical density at 600 nm reached 0.8. At this time, protein expression was induced by addition of 2 mL of 0.5 M IPTG to 1 L of cells. After induction, the cells were shaken at 37 °C for another 3 h. The cells were then centrifuged at 6000 rpm and the pellet was collected and stored at −80 °C. The pellet from 1 L cell culture was suspended in 140 mL of 100 mM NaCl and incubated on ice for 15 min. The cells were sonicated for 5 min with 2 s pulses to lyse the cells and break up the genomic DNA. The lysed cells were centrifuged and resuspended in 140 mL of 100 mM Tris-HCl (pH 8.0). The suspension was centrifuged and the pellet was suspended in 10 mL of solubilization buffer (20 mM Tris-HCl, pH 8.0; 7 M guanidine-HCl; 10 mM DTT; 10 mM EDTA; 1 mM PMSF). This solution was stirred under nitrogen for 2 h to solubilize the RNase A mutants. The soluble portion was diluted to 100 mL of 20 mM acetic acid solution and dialyzed against 4 L of 20 mM acetic acid solution. The resulting solution was further diluted to 600 mL of refolding buffer (0.1 M Tris-acetic acid buffer, pH 8.0; 0.1 M NaCl; 3.0 mM reduced glutathione; 0.6 mM oxidized glutathione). This solution was kept at room temperature overnight for refolding. After refolding, the volume of this solution was reduced using an ultrafiltration cell (Amicon) with an ultrafiltration membrane (MWCO 3000).

R39A, R85A, and K91A were purified separately on a SP Sepharose fast flow cation exchange column (GE Healthcare). The protein solution was applied to this cation exchange column, washed with a buffer of 25 mM Hepes and 1 mM EDTA at pH 7, and eluted with a linear gradient from 5 to 500 mM NaCl. The flow was monitored by UV absorption at 280 nm. Major UV active peaks were collected and characterized by SDS-PAGE. The fractions containing mutant RNase A were combined and dialyzed against the equilibration buffer of 5 mM potassium phosphate and 40 mg/L $CaCl_2$ at pH 6.0. The preparation was concentrated and applied to a Macro-Prep Ceramic Hydroxyapatite column (Bio-Rad Laboratories). Proteins were washed with the equilibration buffer and eluted with an 80 min linear gradient (5–500

mM) of potassium phosphate buffer. The fractions with highest purity RNase A (based on SDS-PAGE) were collected.

Results

Throughout the studies described below, we have used incubation of D-glucose and $^{13}C_6$ D-glucose in an equimolar ratio in order to obtain a characteristic isotopic signature of modified peptides by mass spectrometric analysis of proteolytic digests. Doublet peaks separated by 6 Da were obtained if modification was on a single residue, and triplet peaks (1:2:1) were obtained if the peptide incorporated two modifications. This strategy helped us unequivocally identify the glucose-modified peptide signals. Furthermore, by using preglycation under strict anaerobic conditions for 2 weeks to prevent glycoxidation, followed by dialysis to remove free glucose and oxoaldehydes, we hoped to minimize modifications by reactive intermediates in free form such as glyoxal, methylglyoxal, and 3-deoxyglucosone.

Sequence Coverage of RNase A from MALDI Analysis

The sequence coverage of RNase A following tryptic digestion from MALDI experiments was at least 85%. The high sequence coverage of the enzymatic digestions decreased the chance of missing modified peptides whose corresponding native peptides were not sensitive in mass spectrometry. In comparison, chymotrypsin digestion only covered 50% of the sequence, while Asp-N digestion covered almost 100% of the sequence. The sequence coverage is apparently not directly related to the number of cutting sites by different enzymes because chymotrypsin yields the smallest average peptide size while having the lowest sequence coverage. The coverage difference between different enzymes may be linked to the charge distribution on the individual peptides after enzymatic cutting. For example, chymotryptic cutting may lead to more peptides which have no positive charged residues. These noncharged peptides appear with significantly reduced intensities in MALDI peptide maps.

Amadori- and Glucosepane-Modified Peptides Are Major Modifications in Ribonuclease A Incubated with Glucose

It is well-established that oxidative conditions during incubation of proteins with glucose favor the formation of certain AGEs over others, in particular the glycoxidation products CML and pentosidine.[17] Instead, glucosepane does not require oxidation for its formation according to the proposed mechanism of formation.[18] Thus, in order to enhance glucosepane formation and diminish glycoxidation, we have used metal-depleted phosphate buffer and nonoxidative conditions. The results of the tryptic digest of glycated RNase A confirmed the above hypothesis. Twelve doublet peaks, $\Delta m/z = 6$, each corresponding to a native peptide with a modification by a single glucose-derived product, and four triplet peaks corresponding to native peptides with two glucose-derived products were observed. Most of these peaks could be assigned to Amadori intermediates and/or glucosepane-modified peptides (Table 1). Five peptides were modified with Amadori intermediates only, except for one peptide with *m*/*z* 2677.3 corresponding to ^{40}C*KPVNTFVHESLADVQAVC*SQK^{61} with Lysine 41 modified with $\Delta m/z = 160$, as unequivocally confirmed by tandem mass spectrometry (Figure 2). The modification by $\Delta m/z = 160$ instead of 162 (Amadori product) is quite unexpected as it suggests a previously unreported and poorly understood oxidation product of an Amadori product. The *m*/*z* 3251.6 peak also could be assigned to the peptide ^{40}C*KPVNTFVHESLADVQAVC*SQKNVAC*K^{66} with Amadori modification at K41. The *m*/*z* 1312.7 peak corresponds to peptide ^{1}KETAAAKFER^{10} with Amadori modification at K1 or K7. The other glycation hot spots of Amadori modifications at K37 are comprised in the *m*/*z* 3514.7 and 3630.75 peaks, corresponding to ^{34}NLTKDRC*KPVNTFVHESLADVQAVC*SQK^{61} and ^{38}DRC*KPVNTFVHESLADVQAVC*SQKNVAC*K^{66}. These peptides have in

common the sequence 38–61, and in addition to the Amadori product (Δ*m*/*z* = 162), an *m*/*z* increase of 108, consistent with the presence of an intramolecular glucosepane. Thus, these peptides have a characteristic isotopic triplet peak ratio of 1:2:1. Besides the glycation hot spots at K1, K7, K37 and K41, other lysine residues, such as K66, K104, and K98 were also shown to be modified by Amadori products.

As shown in Table 1, the modified peak with *m*/*z* of 3233.58 cannot be readily assigned to an Amadori product or glucosepane-modified peptide. However, it is consistent with loss ofH_2OfromAmadorimodified ^{40}C*KPVNTFVHESLADVQAVC*SQKNVAC*K^{66}. A lysine-bound 1,4-dideoxy-5,6-glucosone, which has been suggested as the precursor of glucosepane,[18] is one possible modification consistent with this mass increment. The detection of this dideoxy-glucosone modified peptide strengthens previous data suggesting that it is an important glucosepane precursor.

The glucose-derived modifications are evidenced by the characteristic doublet or triplet peaks (for examples, see Figures 5 and 6b). These peaks can be partly explained by Amadori-modified peptides, while the non-Amadori product peaks in Table 1 are consistent with the presence of intra- or intermolecular glucosepane modifications.

The most prominent glucosepane containing peptide was ^{38}DRC*KPVNTFVHESLADVQAVC*SQK^{61}, corresponding to a doublet peak with the *m*/*z* 2896.4 and 2902.4, respectively. The LC/MS/MS fragmentation data shown in Figure 3 support the existence of an intramolecular glucosepane cross-link between K41 and R39. The appearance of a b_4 ion that contains the mass increment localizes the modification to the first four residues of the peptide. The very intense y^{23} ion, containing both R39 and K41 but not D38, further limits the location of this adduct. The presence of y^4–y^{17} ions consistent with the unmodified sequence confirms the identity of the peptide. The absence of b_2, b_3, y_{21} and y_{22} ions is predicted for an intramolecular cross-link.

Besides this peptide, several other modified peptides can be assigned to peptides with intramolecular glucosepane between R39 and K41, including

1. ^{34}NLTKD**R**C*__K__PVNTFVHESLADVQAVC*SQK^{61},
2. ^{34}NLT**K**D**R**C*__K__PVNTFVHESLADVQAVC*SQK^{61},
3. ^{38}D**R**C*__K__PVNTFVHESLADVQAVC*SQKNVAC*K^{66}.

Peptide 1 and 2 actually are same peptides with different modifications: peptide 1 only has an intramolecular glucosepane modification, whereas peptide 2 has both an intramolecular glucosepane and an Amadori product, presumably at K37, resulting in the missed tryptic cleavage. Of relevance in the above analyses and throughout these studies is that the peptide containing the glucosepane modification at K41 and R39 was a major signal, strongly suggesting that this site is a highly favored target for glucosepane formation.

The *m*/*z* 2914.4 cannot be easily explained by either Amadori or glucosepane modifications. Instead, it is consistent with the peptide ^{34}DRC*KPVNTFVH ESLADVQAVC*SQK^{61} with, however, a Δ*m*/*z* of 126. This mass increment is 18 Da greater than that of an intramolecular glucosepane modification, that is, potentially due to addition of H_2O. Another possible AGE cross-link, DODIC (Figure 1), also has a six carbon backbone derivable from glucose that would yield M and M + 6 doublets with Δ*m*/*z* = 126 from the unmodified peptide. The LC/MS/MS spectrum of this peptide strengthened this tentative assignment. As shown in Figure 4a, the existence of the predicted Δ*m*/*z* = 126 b_4 ion again limits the modification to the fragment ^{34}DRC*K^{37}, suggesting the modification could be DODIC. Interestingly, when we tried to find the triply charged peptide with *m*/*z* 972.1 (M_r = 2914.4 Da) in the LC/MS spectrum,

we found that there are two major species with different elution times at 22.47 and 24.38 min (Figure 4b). The second 2914.4 peak coeluted with the peak with *m*/*z* 2896.4, leading us to speculate that the first 2914.4 peptide may be modified by, for example, DODIC, whereas the second 2914.4 peptide might be modified by glucosepane (2896.4) plus a water molecule added in the mass spectrometer. Coelution may suggest in-source fragmentation, that is, loss of H_2O.

A major question this work hoped to address is the extent to which glucose participates in intermolecular cross-linking. The peptide with *m*/*z* of 4541.0 matches the theoretical mass of the peptide ^{67}NGQTNC*YQSYSTMSITDC*RETGSSK^{91} cross-linked to ^{92}YPNC*AYKTTQANK^{104}, by glucosepane. Glucosepane would necessarily cross-link R85 and K98, as K104 and K91 must be unmodified to allow for tryptic digestion. But with the existence of the nearby doublet peak with *m*/*z* of 4523.0, we need to reconsider the origin of the peak at 4541.0 (Figure 5). This hypothesized cross-linked peptide, ^{67}NGQTNC*YQSYSTMSITDC*RETGSSK$^{91—92}$YPNC*AYKTTQANK^{104}, is comprised of sequential tryptic peptides. As indicated by the 3-D crystal structure of RNase A (Figure 7), residues K98 and R85 are very close to each other, making formation of an intramolecular glucosepane probable and it would be indistinguishable by mass spectrometry from an intermolecular cross-link. With all these factors considered, the peak at *m*/*z* 4541.0 may actually originate by hydrolysis of the peak with *m*/*z* 4523.0. In a glycated protein containing an intramolecular cross-link, the appearance of an M + 18 peak may be due to addition of water upon enzymatic hydrolysis of an internal peptide bond. In this case of RNase A with glucosepane cross-linking K98 and R85, either partial tryptic digestion at K91 or chymotryptic digestion at Y97 will result in two mass spectral doublets separated by 18 Da. The ratio of original peptide and the M + 18 peptide derived from enzymatic hydrolysis will depend on the efficiency of enzymatic digestion (on K91–Y92 in tryptic digestion and on Y97–K98 in chymotryptic digestion).

The search for intermolecular cross-linking sites revealed unexpected results. The anticipated intermolecular glucosepane cross-link has proven elusive. However, after searching for all glucosepane-based theoretical *m*/*z* values compatible with cross-linked peptides, we could not identify candidate glucosepane cross-linked tryptic peptides consistent either with *m*/*z* of 4064.01 (doublet peak) or 4226.1 (1:2:1 triplet peak). Interestingly, both of these *m*/*z* values matched theoretical values of intermolecular cross-links assuming the cross-link had a molecular mass 18 Da greater than glucosepane. These data prompted a search for other intermolecular cross-links which contain all 6 carbons of a single glucose. Again, the DODIC cross-link is consistent with these mass spectral data. The tandem mass spectrum of the 4064.01 peptide supported the presence of an intermolecular cross-link with Δ*m*/*z* of 126 between K1 and R39 (Figure 6a). The b ion, y ion, b′ ion and y′ ion series establish the presence of both peptides, ^{1}KETAAAKFER^{10} and ^{38}DRC*KPVNTFVHESLADVQAVC*S QK^{61}. The appearance of b′4 and b′5 also suggests that the cross-linking site contains K1 instead of K7. In the same manner, the peptide with *m*/*z* value of 4226.1 also fits the cross-linking of peptide ^{1}KETAAAKFER^{10} and ^{38}DRC*KPVNTFVHESLADVQAVC*S QK^{61} by DODIC, with additional Amadori modification at K7 or K41 (Figure 6b).

In order to strengthen the above observations, digestions with the enzymes Asp-N and chymotrypsin were also carried out to verify the presence of the modification by Amadori, glucosepane and DODIC (Table 2). These two enzymes yield relatively low amounts of modified peptides and some of them cannot be assigned readily so that less information was obtained than with tryptic digestion.

In the Asp-N digestion map, peptides with *m*/*z* of 1480.7 and 1880.9 confirmed the existence of a major intramolecular glucosepane between K41 and R39. A relative weak doublet peak

($\Delta m/z = 6$), with m/z 1898.9 shows that an intramolecular DODIC cross-link could be present between K41 and R39. The doublet peak with m/z 2830.3 and 2836.3 corresponds to the peptide ^{14}DSSTSAASSSNYC*NQMMKSRNLTK^{37}, with an Amadori modification at K31 or K37. On the basis of the data with trypsin and the known glycation hot spots, modification at K37 is more likely than at K31. In the same manner, the peptide with m/z 3604.6 is assigned to ^{53}DVQAVC*SQKNVAC*KNGQTNC*YQSYS TMSIT^{82}, with Amadori modification at K61 or K66. Combined with the data from tryptic digestion, K66 would be the more likely site. Both a weak doublet peak at 4142.8 and a strong double peak at 4574.2 suggest an Amadori modification at K91, K98 or K104. The two sets of doublet peaks with m/z 2444.1 and 3734.7 did not correspond to modification of predicted Asp-N peptides. But it is possible that peak 2444.1 comes from an intramolecular glucosepane-modified peptide, whereas peak 3734.7 comes from an Amadori-modified peptide due to the appearance of their prominent corresponding native peptides (m/z = 2336.0 and 3572.5, respectively).

Finally, in the chymotryptic digestion map, the peptide with m/z 1473.7 confirms the major intramolecular glucosepane between K41 and R39. Furthermore, the peptide with m/z 2271.0 indicates that the peptide ^{80}SITDC*RETGSSKYPNC*AY^{97} is modified by an Amadori product at K91, a site whose modification was not identified in the tryptic digest. The peptides with m/z 4259.8 and 4277.8 have been discussed before and could represent the same intramolecular glucosepane-modified peptide (vide supra) and after enzymatic hydrolysis between the cross-linked residues.

The tryptic mapping of wild-type and RNase A mutants R39A, R85A, and K91A incubated with glucose is summarized in Table 3. As predicted, when R39 is changed to A39, the assigned R39-K41 intramolecular cross-links and K1-R39 intermolecular cross-links disappeared. Instead, new peptides with K1 or K41 modified with Amadori product were shown as glucose-modified peaks, such as peptide doublet peaks with m/z of 1312.7, 3321.6, 3419.6, 3437.5, 3481.6 (triplet). In contrast, peptides with cross-linking sites different from R39, such as K98-R85 and K1–R85, were not affected by the R39A mutation. Similarly, the mutation R85A only affects R85 containing cross-linking, such as K98–R85 and K1–R85. The newly identified peptide with m/z of 4491.5 is assigned to peptide^{67} NGQTNC*YQSYSTMSITDC*RETGSSKYPNC*AYKTTQANK^{104}, with possible Amadori modification at K98, due to lack of R85. All the peptides using R39 as the proposed cross-linking site remain unaffected.

Discussion

As Biemel et al.[19] proposed, both glucosepane and DODIC are derived from Amadori products. Consequently, clarifying the glycation hot spots for the Amadori product was a necessary step prior to investigating the sites of glucosepane or DODIC formation. Two studies by Watkins et al. and Brock et al. agreed on the glycation hot spots of RNase A with the order of K41, K7, K37 and K1[20,21] with slight differences. Watkins group showed that K1 is more glycated than K37 based on radioactivity measurement of [^{14}C]glucose, which should be more reliable than the semiquantitative methods of mass spectrometry. In their study, 38% of K41, 29% of K7, 24% of K1 and 9% of K37 were glycated by Amadori products. One can assume that because the Amadori product is the precursor of glucosepane and DODIC, the sites that are preferentially modified by Amadori product accumulation will have a greater potential for glucosepane formation, particularly if there is a proximal arginine residue.

Our observation confirmed the four glycation hot spots listed above. Additional glycation sites also were indicated by enzymatic peptide maps as K66, K91, K98, and K104. The identification of additional sites of modification may be due to the different incubation conditions and digestion procedures. For example, we have used 500 mM glucose, while previous studies by

Watkins et al. and Brock et al. have used 400 mM glucose. We used nonoxidative incubation conditions, while previous studies did not exclude O_2. As expected, the major cross-link identified under our nonoxidative conditions is glucosepane as evidenced by the dominant tryptic peptide with *m*/*z* 2896.4 (residues 38–61) in the digestion of glycated RNase. In this aspect, our *in vitro* nonoxidative incubation results conform well with the predominant level of glucosepane observed during analysis of *in vivo* human tissue modifications.[11,12]

Most established protocols and routine methods used to identify and confirm cross-links *in vivo* utilize complete acid hydrolysis or exhaustive enzymatic digestion of proteins to the constituent amino acids, and amino acids with modifications.[11,22] The cross-links identified by these protocols will include both intra- and intermolecular cross-links. Thus, discrimination of the cross-links as being primarily intra- or intermolecular is important for understanding of the *in vivo* implication of cross-linking.

Whether intra- or intermolecular glucosepane cross-links are formed should depend on the availability of nearby arginine residue(s) at the major sites of Amadori product accumulation. Thus, we calculated the distance between several pairs of nearby lysine and arginine residues from the crystal structure 1RBX deposited in the protein data bank (Dunbar, J.; Yennawar, H. P.; Banerjee, S.; Farber, G. K., unpublished results) using PyMOL (DeLano, W.L., The PyMOL Molecular Graphics System (2002) on World Wide Web http://www.pymol.org). These distances are highlighted in Figure 7.

From the values of the paired lysine/arginine distances, the most likely intramolecular cross-links are K98–R85 and K41–R39. In the MALDI tryptic peptide map, the intensity of the peptide with intramolecular glucosepane between K41 and R39 (*m*/*z* 2896.4) is much higher than the peptide with intramolecular glucosepane between K98 and R85 (*m*/*z* 4523.0). K41 has been unequivocally established as the most prevalent Amadori accumulation site, whereas K98 is not an established glycation hot spot, although our data proved the presence of the Amadori product on K98. Other factors may have minor effects on the level of intramolecular glucosepane formation at different sites, such as the relative surface accessibility of lysine and arginine. As shown in Table 4, the relative surface accessibility (RSA) values of lysines and arginines are predicted by the Structural Analysis of Residue Interaction Graphs (SARIG) server provided by Weizmann Institute of Science (bioinfo2.weizmann.ac.il/pietro/SARIG/). The more the arginine is exposed, the greater propensity for intramolecular glucosepane formation. The lowest accessibility of K41 will not inhibit the intramolecular glucosepane formation as long as it has a local arginine “mate”. The highest accessibility of R39 makes it not only a suitable local “mate” for K41, but also a good candidate for intermolecular cross-linking.

The existence of both modifications with Δ*m*/*z* 108 and 126 is very interesting. These paired modifications suggest that both glucosepane and DODIC modifications are occurring at the same sites. Initially, DODIC modifications were not anticipated because of low levels observed *in vivo* as compared to glucosepane. Glucosepane represents the dominant cross-link in collagen[12] and in lens protein (132.3–241.7 pmol glucosepane/mg of protein compared to 1.3–8.0 pmol DODIC/mg of protein).[11] However, the appearance of DODIC at the same sites as intramolecular glucosepane is quite reasonable because glucosepane and DODIC share the same precursor, the aminoketose. An interesting part of this work is the relative peak intensity of intramolecular glucosepane and intramolecular DODIC modifications at different intracross-linking sites. At the K41 and R39 cross-linking site, the intramolecular glucosepane containing peptide ion (*m*/*z* 2896.4) is much more intense than the intramolecular DODIC containing peptide ion (*m*/*z* 2914.4). While at the K98 and R85 cross-linking site, instead, the intramolecular DODIC containing peptide ion (4541.0) is more intense than the intramolecular glucosepane containing peptide ion (4523.0). The peptide ion at *m*/*z* 4541.0 may be derived

from two sources, that is, the peptide ^{67}NGQTNC*YQSYS TMSITDC*RETGSSKYPNC*AYKTTQANK^{104} with an intramolecular DODIC modification, or the peptide ^{67}NGQTNC*YQSYSTMSITDC*RETGSSKYPNC*AYKTTQANK^{104} with intramolecular glucosepane modification and an additional tryptic cleavage at K91.

An important aspect of this study was to identify the sites of glucose-derived intermolecular cross-links. As noted above, the peaks at *m*/*z* 4064.0 and 4226.1 fit the theoretical masses for intermolecular glucosepane containing peptides. Furthermore, the LC/MS/MS fragmentation data of the peptide with *m*/*z* 4064.0 not only confirmed the presence of both peptides, but also suggested that K1 and R39 were the cross-linked residues. As shown in Table 2, the *m*/*z* value of 4226.1 was consistent with the same peptides being cross-linked with an additional Amadori modification at K7. As shown in Table 4, the relative surface accessibilities of K1 and R39 are the highest among lysines and arginines of RNase A, respectively. Although the accumulation of Amadori product at K1 is not the dominant site in RNase A, K1 is one of the glycation hot sites. With all these factors taken into consideration, the major intercross-link found between K1 and R39 is not so surprising.

However, the observation of DODIC as the only intercross-link observed is quite surprising. In *in vitro* incubations of bovine serum albumin with glucose,[19] Biemel et al. reported about 3-fold higher yields of glucosepane than DODIC. Instead, glucosepane is present at 10–40 times more than DODIC in both human serum albumin and human lens protein samples.[11] The authors attribute the dominance of glucosepane *in vivo* to that of its precursor, N^6-(2,3-dihydroxy-5,6-dioxo-hexyl)-L-lysinate, which is irreversibly linked to the protein and inaccessible to the detoxifying reductases and thus becomes a persistent glycation agent. This character is quite different from other important intermediates including 3-DG, MGO, and GO, which exist in equilibrium between free and loosely protein-associated forms.

Although Lederer et al. initially proposed a mechanism of glucosepane formation through DODIC as a precursor,[23] the presence of DODIC and the absence of glucosepane in the incubation of BSA with a 3-deoxyosone mixture exclude DODIC as a precursor of glucosepane.[19] Using *o*-phenylenediamine as a trapping agent for protein bound dicarbonyl groups, Biemel et al. suggest that N^6-(2,3-dihydroxy-5,6-dioxohexyl)-L-lysinate is the only prominent lysine linked α-diketo compound formed from the incubation of glucose with lysozyme.[24] Later they proposed a mechanism for the formation of glucosepane from Amadori products via N^6-(2,3-dihydroxy-5,6-dioxohexyl)-L-lysinate and the seven member ring intermediate, azepanone. Thus, the formation of glucosepane and DODIC only shares the step from lysine to Amadori product. The formation of glucosepane through N^6-(2,3-dihydroxy-5,6-dioxohexyl)-L-lysinate requires several enolization steps along the sugar backbone to effect the intramolecular oxidation–reduction reactions, whereas these steps are not required in DODIC formation. We observed that intramolecular glucosepane is highly favored in the K41–R39 cross-link over DODIC. Interestingly, Biemel et al. observed that the three lysine residues K33, K96, and K116 of lysozyme, which all have partner arginine(s) within 5 Å, have the highest rate of transformation from Amadori to N^6-(2,3-dihydroxy-5,6-dioxohexyl)-L-lysinate.[24] This correlation suggested that the guanidino group of arginine catalyzes this transformation.

By analogy, the lysine sites with nearby arginine residues in RNase A could also favor the formation of glucosepane, whereas those lysine residues without nearby arginine(s) will be more likely to form DODIC. In our case, the highly glycated lysine K41 is more prone to form intramolecular glucosepane due to close proximity of R39. Because of extremely low surface accessibility, K41 can form neither intermolecular glucosepane nor intermolecular DODIC cross-links. Among all other glycation hot spots, K1 has high surface accessibility and

relatively high Amadori accumulation, which makes it a good candidate site to form an intermolecular cross-link. Without nearby arginine-catalyzed enolization, DODIC formation is favored at this site. This process is presented in Figure 8.

In summary, this work demonstrates that the Amadori products, glucosepane, and DODIC are major glucose-derived modifications accumulating in RNase A under nonoxidative glycation conditions. Two intramolecular glucosepane cross-links were found at sites R39–K41 and R85–K98. At the site R39–K41, an intramolecular DODIC cross-link was also observed. The only intermolecular cross-link was also observed at K1–R39, but surprisingly found to be a DODIC cross-link.

Acknowledgements

This work was supported by a grant from the National Institute on Aging AG-18426 and in part by NEI grant EY-07099.

Abbreviations

3-DG
: 3-deoxyglucosone

AGE(s)
: advanced glycation end product(s)

DODIC
: 3-deoxyglucosone-derived imidazolium cross-link

DTPA
: diethylenetriamine pentaacetic acid

LC/MS
: liquid chromatography–mass spectrometry

MALDI
: matrix-assisted laser desorption and ionization

RNase A
: ribonuclease A

ROS
: reactive oxygen species

RP-HPLC
: reversed phase-high performance liquid chromatography

TFA
: trifluoroacetic acid

References

1. Monnier VM, Vishwanath V, Frank KE, Elmets CA, Dauchot P, Kohn RR. Relation between complications of type I diabetes mellitus and collagen-linked fluorescence. N Engl J Med 1986;314 (7):403–8. [PubMed: 3945267]
2. Hammes HP, Martin S, Federlin K, Geisen K, Brownlee M. Aminoguanidine treatment inhibits the development of experimental diabetic retinopathy. Proc Natl Acad Sci USA 1991;88(24):11555–8. [PubMed: 1763069]

3. Nakayama H, Mitsuhashi T, Kuwajima S, Aoki S, Kuroda Y, Itoh T, Nakagawa S. Immunochemical detection of advanced glycation end products in lens crystallins from streptozocin-induced diabetic rat. Diabetes 1993;42(2):345–50. [PubMed: 8425672]
4. Mitsuhashi T, Nakayama H, Itoh T, Kuwajima S, Aoki S, Atsumi T, Koike T. Immunochemical detection of advanced glycation end products in renal cortex from STZ-induced diabetic rat. Diabetes 1993;42(6):826–32. [PubMed: 8495806]
5. Schnider SL, Kohn RR. Effects of age and diabetes mellitus on the solubility and nonenzymatic glucosylation of human skin collagen. J Clin Invest 1981;67(6):1630–5. [PubMed: 6787079]
6. Hamlin CR, Luschin JH, Kohn RR. Partial characterization of the age-related stabilizing factor of post-mature human collagen--II. By the use of trypsin. Exp Gerontol 1978;13(6):415–23. [PubMed: 33060]
7. Hamlin CR, Luschin JH, Kohn RR. Partial characterization of the age-related stabilizing factor of post-mature human collagen--I. By the use of bacterial collagenase. Exp Gerontol 1978;13(6):403–14. [PubMed: 33059]
8. LaBella FS, Paul G. Structure of collagen from human tendon as influenced by age and sex. J Gerontol 1965;20:54–9. [PubMed: 14246531]
9. Paul RG, Bailey AJ. Glycation of collagen: the basis of its central role in the late complications of ageing and diabetes. Int J Biochem Cell Biol 1996;28(12):1297–310. [PubMed: 9022289]
10. Kivirikko KI, Peltonen L. Abnormalities in copper metabolism and disturbances in the synthesis of collagen and elastin. Med Biol 1982;60(2):45–8. [PubMed: 6124662]
11. Biemel KM, Friedl DA, Lederer MO. Identification and quantification of major maillard cross-links in human serum albumin and lens protein. Evidence for glucosepane as the dominant compound. J Biol Chem 2002;277(28):24907–15. [PubMed: 11978796]
12. Sell DR, Biemel KM, Reihl O, Lederer MO, Strauch CM, Monnier VM. Glucosepane is a major protein cross-link of the senescent human extracellular matrix. Relationship with diabetes. J Biol Chem 2005;280(13):12310–5. [PubMed: 15677467]
13. Monnier VM, Mustata GT, Biemel KL, Reihl O, Lederer MO, Zhenyu D, Sell DR. Cross-linking of the extracellular matrix by the maillard reaction in aging and diabetes: an update on ”a puzzle nearing resolution. Ann NY Acad Sci 2005;1043:533–44. [PubMed: 16037276]
14. Price DL, Rhett PM, Thorpe SR, Baynes JW. Chelating activity of advanced glycation end-product inhibitors. J Biol Chem 2001;276(52):48967–72. [PubMed: 11677237]
15. Khalifah RG, Todd P, Booth AA, Yang SX, Mott JD, Hudson BG. Kinetics of nonenzymatic glycation of ribonuclease A leading to advanced glycation end products. Paradoxical inhibition by ribose leads to facile isolation of protein intermediate for rapid post-Amadori studies. Biochemistry 1996;35(15):4645–54. [PubMed: 8664253]
16. Laity JH, Shimotakahara S, Scheraga HA. Expression of wild-type and mutant bovine pancreatic ribonuclease A in Escherichia coli. Proc Natl Acad Sci USA 1993;90(2):615–9. [PubMed: 8421696]
17. Miyata T, Wada Y, Cai Z, Iida Y, Horie K, Yasuda Y, Maeda K, Kurokawa K, van Ypersele de Strihou C. Implication of an increased oxidative stress in the formation of advanced glycation end products in patients with end-stage renal failure. Kidney Int 1997;51(4):1170–81. [PubMed: 9083283]
18. Biemel KM, Conrad J, Lederer MO. Unexpected carbonyl mobility in aminoketoses: the key to major Maillard crosslinks. Angew Chem, Int Ed Engl 2002;41(5):801–4. [PubMed: 12491341]
19. Biemel KM, Reihl O, Conrad J, Lederer MO. Formation pathways for lysine-arginine cross-links derived from hexoses and pentoses by Maillard processes: unraveling the structure of a pentosidine precursor. J Biol Chem 2001;276(26):23405–12. [PubMed: 11279247]
20. Watkins NG, Thorpe SR, Baynes JW. Glycation of amino groups in protein. Studies on the specificity of modification of RNase by glucose. J Biol Chem 1985;260(19):10629–36. [PubMed: 4030761]
21. Brock JW, Hinton DJ, Cotham WE, Metz TO, Thorpe SR, Baynes JW, Ames JM. Proteomic analysis of the site specificity of glycation and carboxymethylation of ribonuclease. J Proteome Res 2003;2(5):506–13. [PubMed: 14582647]
22. Sell DR, Monnier VM. Structure elucidation of a senescence cross-link from human extracellular matrix. Implication of pentoses in the aging process. J Biol Chem 1989;264(36):21597–602. [PubMed: 2513322]

23. Lederer MO, Klaiber RG. Cross-linking of proteins by Maillard processes: characterization and detection of lysine-arginine cross-links derived from glyoxal and methylglyoxal. Bioorg Med Chem 1999;7(11):2499–507. [PubMed: 10632059]

24. Biemel KM, Lederer MO. Site-specific quantitative evaluation of the protein glycation product N6-(2,3-dihydroxy-5,6-dioxo-hexyl)-L-lysinate by LC-(ESI)MS peptide mapping: evidence for its key role in AGE formation. Bioconjugate Chem 2003;14(3):619–28.

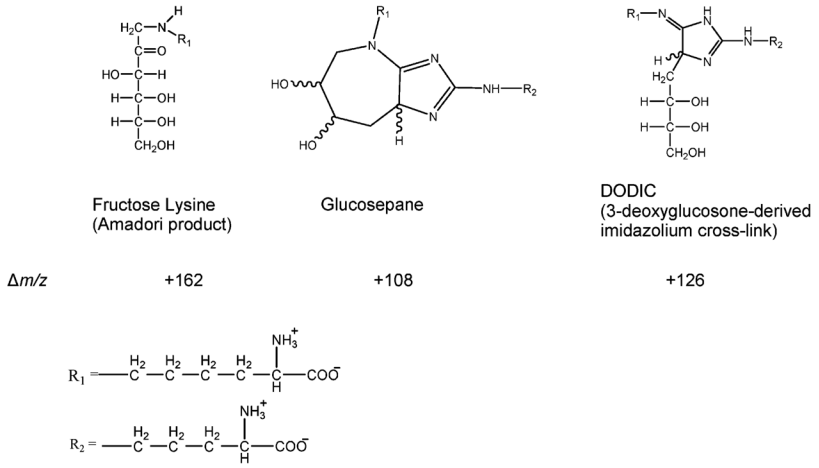

Figure 1.
Structure of Amadori product, glucosepane, and DODIC with corresponding Δ*m*/*z* values.

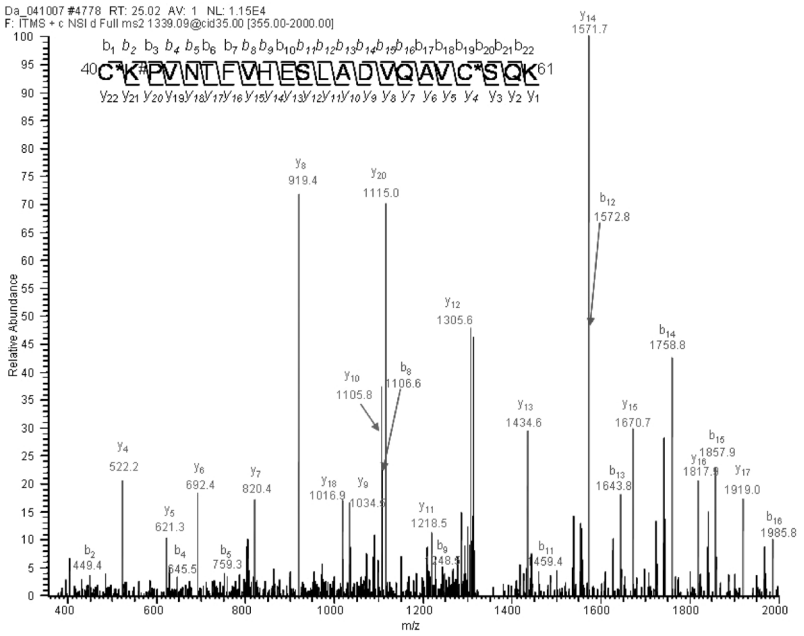

Figure 2.

Tandem mass spectrum of glycated RNase A tryptic peptide with molecular doubly charged ion at *m*/*z* 1339. This signal was assigned to peptide ^{40}C*KPVNTFVHESLADVQAVC*SQK^{61} with modification of K41 by $\Delta m/z = 160$ (oxidized Amadori). Every b ion greater than b2 contained this increment, while every y ion up to y20 was consistent with the unmodified peptide sequence, strongly implicating the modification of K41 by an oxidized Amadori product.

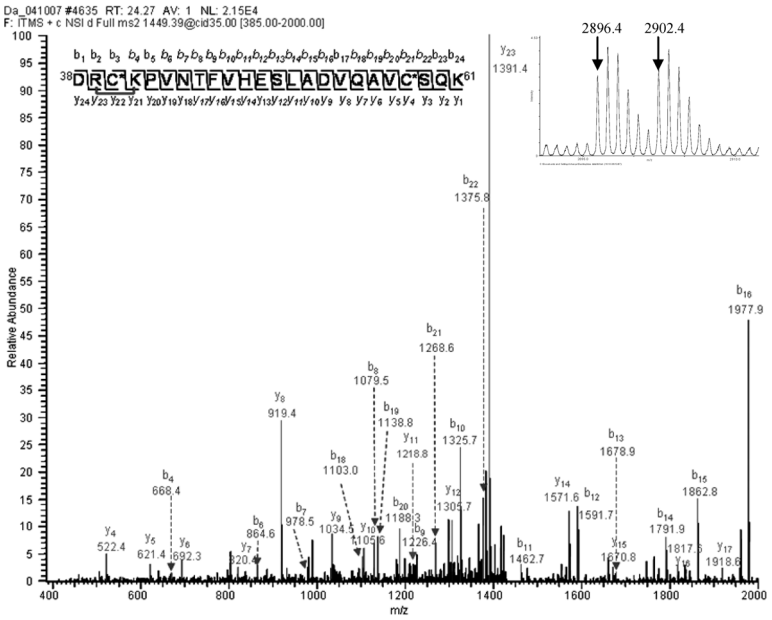

Figure 3.

Tandem mass spectrum of the glycated RNase A tryptic peptide corresponding ^{36}DRC*KPVNTFVHESLADVQAVC*SQK^{61} with precursor *m*/*z* 1449.4 (double charged). The presence of ions b4 and y23, both containing the glucosepane mass increment, suggest the presence of intermolecular glucosepane cross-linking K41 and R39. b18–b22 and y23 ions are doubly charged. The inset shows the corresponding singly charged peptide with *m*/*z* 2896.4 and its *m* + 6 signal at 2902.4.

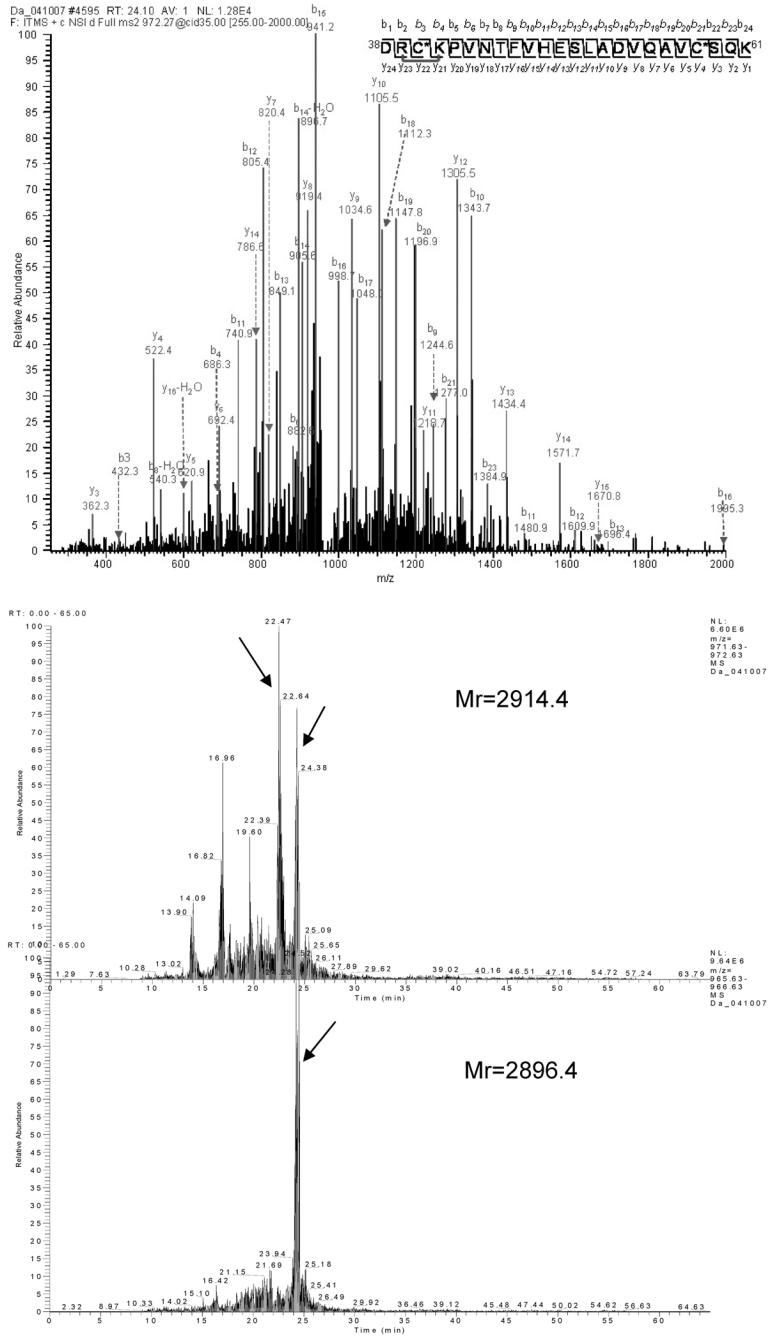

Figure 4.

(a) Tandem mass spectrum of the glycated RNase A tryptic peptide, triply protonated precursor *m*/*z* 972.3, corresponding to ^{36}DRC*KPVNTFVHESLADVQAV C*SQK^{61}, can be interpreted as an intramolecular modification by DODIC cross-linking residues K41 and R39. The b11–b21, b23, and y17 ions are doubly charged. (b) Selected ion chromatograms demonstrating the coelution of peptides with *m*/*z* 2896.4 (presumably glucosepane) and *m*/*z* 2914.4 (presumably DODIC or glucosepane plus 1 mol H2O) suggest that the peptides with *m*/*z* 2914.4 may have different origins.

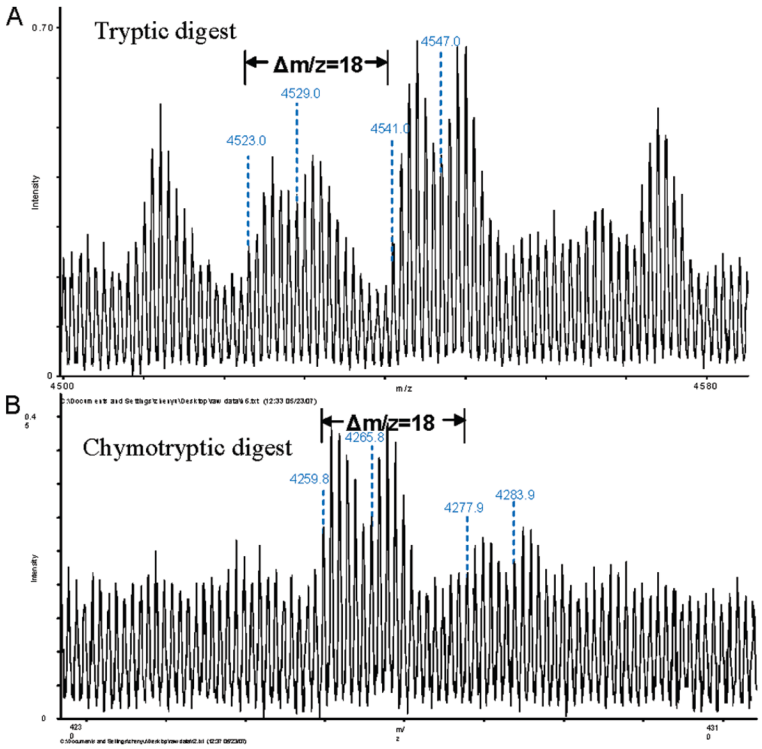

Figure 5.

Both tryptic and chymotryptic peptides indicate the presence of an intramolecular glucosepane cross-link between K98 and R85. Incomplete enzymatic hydrolysis between the cross-linked residues would yield a pseudointermolecular-glucosepane, resulting in an additional *m*/*z* of 18 after digestion. Thus, *m*/*z* 4523.0 corresponds to peptide ^{67}NGQTNC*YQSYSTMSITDC*RETGSSKYPNC*AYKTTQANK^{104} with an intramolecular glucosepane at K98 and R85. The *m*/*z* of 4259.8 corresponds to peptide ^{80}SITDC*RETGSSKYPNC*AYKTTQANKHIIVAC*EGNPY^{115} with an intramolecular glucosepane at R85 and K98/K104.

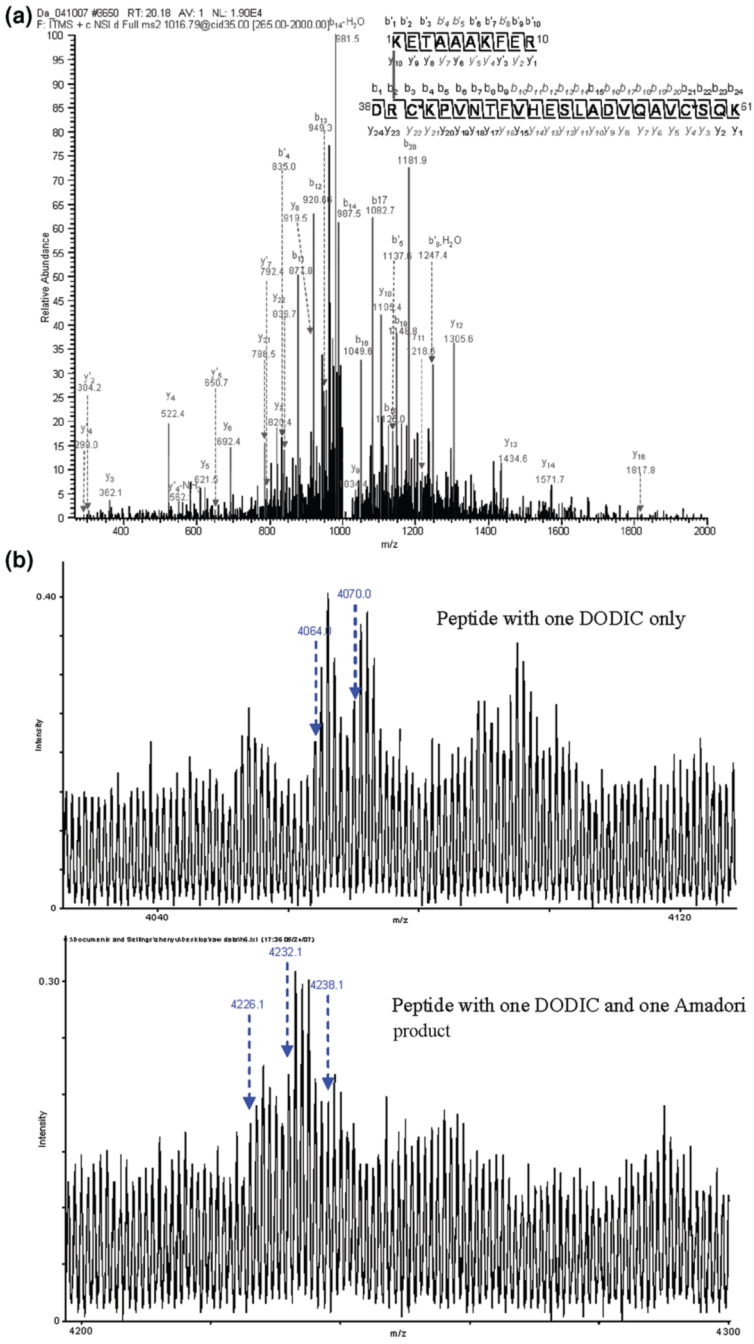

Figure 6.

(a) Tandem mass spectrum of the quadruply charged precursor ion, *m*/*z* 1016.8. y22 ion suggests the cross-linking at R39, while b′4, b′5, and y′4 ion suggests that K1 instead of K7 is the cross-linking site. The b11–b14, b16–b20, b′4, b′5, y21 and y22 ions are triply charged, and the b′8 ion is quadruply charged. The data are compatible with an intermolecular DODIC cross-link. (b) MALDI mass spectrum of the tryptic digestion of glycated RNase A. Two sets of peaks point to intermolecular cross-linking by DODIC between K1 and R39. The two peptides cross-linked are ^{1}KETAAAKFER^{10} and ^{38}DRC*KPVNTFVHESLADVQAVC*SQK^{61}.

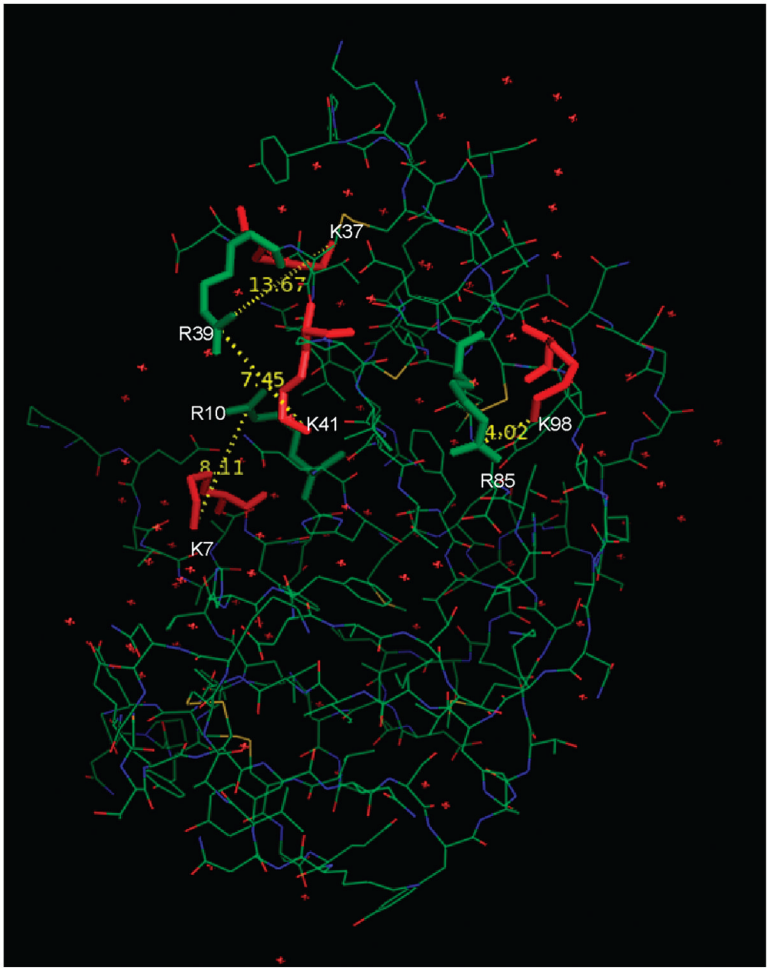

Figure 7.
Distance of all nearby lysine and arginine pairs in RNase A. Calculated lysine and arginine distance: K41 to R39, 7.4 Å; K37 to R39, 13.6 Å; K98 to R85, 4.0 Å; K7 to R10, 8.1 Å.

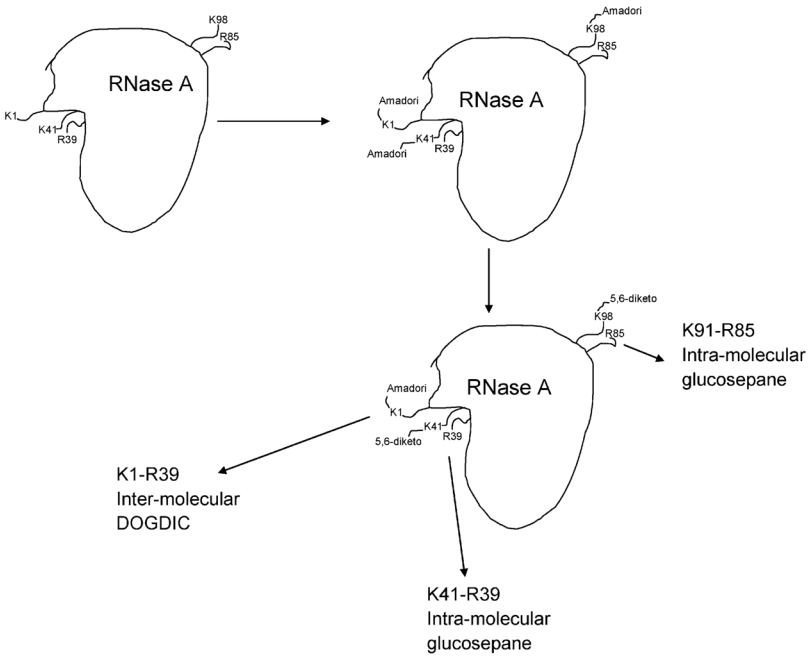

Figure 8.
Schematic representation of glucose mediated cross-link formation in RNase A.

Table 1

Observed Doublet or Triplet Peaks Separated by $\Delta m/z = 6$ in the MALDI Spectrum of the Tryptic Digest of Glycated RNase A[a]

detected *m/z*	theoretical sequence	position of sequence	modified residue	modification type
1312.69	KETAAAKFER (1312.76)	1–10	K1 or K7	Amadori
2677.30	C*KPVNTFVHESLADVQAVC*SQK(2679.3)	40–61	K41	Amadori-2
2896.43	DRC*KPVNTFVHESLADVQAVC*SQK(2896.4)	38–61	R39 to K41	intramolecular glu
2914.43	DRC*KPVNTFVHESLADVQAVC*SQK(2914.4) 2896+ H_2O?	38–61	R39 to K41	
3020.33	NVAC*KNGQTNC*YQSYSTMSITDC*R(3020.3)	62–85	K66	Amadori
3029.48	TTQANKHIIVAC*EGNPYVPVHFDASV(3029.55)	99–124	K104	Amadori
3233.58	3251.6-H_2O			
3251.61	C*KPVNTFVHESLADVQAVCSQKNVAC*K(3251.64)	40–66	K41 or K61	Amadori
3352.7	NLTKDRC*KPVNTFVHESLADVQAVC*SQK(3352.6)	34–61	K41 to R39	Intraglucosepane
3514.75 (1:2:1)	NLTKDRC*KPVNTFVHESLADVQAVC*SQK(3514.85)	34–61	K37(Ama) K41 to R39	Amadori and intramolecular glu
3630.75 (1:2:1)	DRC*KPVNTFVHESLADVQAVC*SQKNVAC*K(3630.85)	38–66	K61(Ama) K41 to R39	Amadori and intramolecular glu
4064.01	KETAAAKFER DRC*KPVNTFVHESLADVQAVC*SQK(4064.07)	1–10 38–61	K1 or K7 R39	intermolecular DC
4088.01 (1:2:1)	YPNC*AYKTTQANKHIIVAC*EGNPYVPVHFDASV (4088.08)	92–124	K98 and K104	2 Amadori
4226.11 (1:2:1)	KETAAAKFER DRC*KPVNTFVHESLADVQAVC*SQK(4226. 1)	1–10 38–61	K1 or K7 R39	Amadori and intermolecular DC
4523.05	NGQTNC*YQSYSTMSITDC*RETGSSKYPNC*AYKTTQANK (4523.0)	67–104	K91 or K98 to R85	Intramolecular glu
4541.05	YPNC*AYKTTQANK NGQTNC*YQSYSTMSITDC*RETGSSK(4541.01)	92–104 67–91	K98 R85	intramolecular glucosepane/pseu intermolecular glu

[a]The data show only the monoisotopic *m/z* value. Asterisks (*) are used to indicate reduced and acetylated cysteine residues.

Table 2

Modified Peptides Mapped by Enzymatic Digestion with (a) Asp-N and (b) Chymotrypsin

detected *m/z*	theoretical sequence	position of sequence	modified residue	modification type
(a) Asp-N				
1394.6				
1898.9	DRC*KPVNTFVHESLA (1898.9)	38–52	R39–K41	Intra-DODIC
1480.7	DRC*KPVNTFVH(1480.7)	38–48	R39–K41	Intraglucosepane
1880.9	DRC*KPVNTFVHESLA (1880.9)	38–52	R39–K41	Intraglucosepane
2444.1	Originated from 2336.0			Intraglucosepane
2830.3?	DSSTSAASSSNYC*NQMMKSRNLTK (2830.3)	14–37	K31 or K37	Amadori
3604.6	DVQAVC*SQKNVAC*KNGQTNC*YQSYSTMSIT (3604.6)	53–82	K61 or K66	Amadori
3734.7	Originated from 3572.5			Amadori
4142.8	ETGSSKYPNC*AYKTTQANKHIIVAC*EGNPYVPVHF (4143.0)	86–120	K91 or K98 or K104	Amadori
4574.2	DC*RETGSSKYPNC*AYKTTQANKHIIVAC*EGNPYVPVHF (4574.2)	83–120	K91 or K98 or K104	Amadori
(b) Chymotrypsin				
811.1				
903.4				
1076.0				
1343.6				
1381.6				
1473.7	TKDRC*KPVNTF(1473.7)	36–46	R39 to K41/K37	Intraglucosepane
2124.8				
2206.1	KTTQANKHIIVAC*EGNPY(2206.1)	98–115	K98 or K104	Amadori
2271.0	SITDC*RETGSSKYPNC*AY (2271.0)	80–97	K91	Amadori
4259.8	SITDC*RETGSSKYPNC*AY	80–97	R85	Interglucosepane
	KTTQANKHIIVAC*EGNPY(4260.0)	98–115	K98 or K104	Or intramolecular gluc
4277.8	SITDC*RETGSSKYPNC*AYKTTQANKHIIVAC*EGNPY(4278.0)	80–115	R85 to K98?	intramolecular glucose

Table 3

The Modifications Affected by Mutations of the Predicted Intra- or Intermolecular Cross-Linking Sites[a]

observed peptides	WT	R39A	R85A	K91A
2677.2 (K41 Amadori-2)	+	3321.6 (K41 Amadori)(36–61)	+	+
2896.4 (K41-R39 Intraglucosepane)	+	3321.6(K41 Amadori)(36–61)	+	+
2914.4 (K41-R39 Intra-DODIC)	+	3321.6(K41 Amadori)(36–61)	+	+
2953.4	+		+	+
3020.3 (K66 Amadori)(62–85)	+	+	4421.8(62–98)	+
3029.4 (K104 Amadori)	+	+	+	+
3233.5 (K41 or K61 Amadori-H_2O)	+	3419.6 (K41 Amadori-H_2O)(38–66)	+	+
3251.5 (K41 or K61 Amadori)	+	3437.5 (K41 Amadori)(38–66)	+	+
3351.6 (K41–R39 Intraglucosepane)	+	3321.6(K41 Amadori)(36–61)	+	+
3380.7	+	+	+	+
3514.7(T) (K41–R39 Intraglucosepane and K37 Amadori)	+	3481.6(K37 and K41 Amadori)(34–61)	+	+
3630.7(T) (K41–R39 Intraglucosepane and K61 Amadori)	+	3481.6(K37 and K41 Amadori)(34–61)		+
4046.0 (K1–R39 Interglucosepane)	+	1312.7 (K1 Amadori) (1–10)	+	+
4064.0 (K1–R39 Inter-DODIC)	+	1312.7 (K1 Amadori) (1–10)	+	+
4082.0 (K1–R39 Inter-DODIC+ H_2O)	+	1312.7 (K1 Amadori) (1–10)	+	+
4226.0(T) (K1–R39 Interglucosepane and K7 Amadori)	+	+	+	+
4523.0 (K98–R85 Intraglucosepane)	+	+	4491.9 (K98 Amadori) (67–104)	4465.9 (K98–R85 intraglucosepane) (67–104)
4541.0 (K98–R85 Intraglucosepane/ pseudo interglucosepane)	+	+	4491.9 (K98 Amadori) (67–104)	4465.9 (K98–R85 intraglucosepane) (67–104)
4558.9 (K98–R85 Intra-DODIC/pseudo inter-DODIC)	+	+	4491.9 (K98 Amadori) (67–104)	4483.9 (K98–R85 intra-DODIC) (67–104)
4581.0 (K1–R85 inter glucosepane-H_2O)	+	+		
4599.0 (K1–R85 inter glucosepane)	+	+		

[a]The plus sign (+) means that the *m*/*z* value (stand for modified peptides) is found in the tryptic digest of glycated RNase A wild-type or mutants.

Table 4

The Relative Surface Accessibility (RSA) Values of Lysines and Arginines Are Predicted by SARIG Server

lysines and arginines on RNase A	RSA
K1	118.0
K7	31.5
K31	46.9
K37	45.1
K41	21.7
K61	44.3
K66	69.2
K91	75.0
K98	39.6
K104	26.8
R10	31.7
R33	23.4
R39	58.2
R85	42.5