Glycoconjugate Journal 15, 835–840 (1998)

Glycation changes the charge distribution of type I collagen fibrils

Julia C. Hadley*, Keith M. Meek and Nageena S. Malik

Oxford Research Unit, The Open University, Foxcombe Hall, Boars Hill, Oxford, OX1 5HR UK

In aging and diabetes, glycation of collagen molecules leads to the formation of cross-links that could alter the surface charge on collagen fibrils, and hence affect the properties and correct functioning of a number of tissues. The electron-optical stain phosphotungstic acid (PTA) binds to positively charged amino acid side-chains and leads to the characteristic banding pattern of collagen seen in the electron microscope; any change in the charge on these side-chains brought about by glycation will affect the uptake of PTA. We found that, upon glycation, a decrease in stain uptake was observed at up to five regions along the collagen D-period; the greatest decrease in stain uptake was apparent at the c1 band. This reduction in PTA uptake indicates that the binding of fructose leads to an alteration in the surface charge at several sites along the D-period. Not all lysine and arginine residues are involved; there appear to be specific residues that suffer a loss of positive charge.

***Keywords*: charge distribution, glycation, sclera, collagen, aging, diabetes**

Introduction

Glycation is the reaction of a reducing sugar with the amino acid groups of proteins, which leads ultimately to the formation of cross-links between neighboring molecules [1, 2]. It is a non-enzymatic reaction and occurs at faster rates in diabetics than other individuals. The advanced products of glycation have now been established to play a role in the evolution of age-related physical changes and diabetic complications – retinopathy, neuropathy, renal failure and atherosclerosis. We have previously shown that glycation of human corneal and scleral collagen changes the molecular packing [3], and that these changes are also associated with aging [4]. Similar changes occur in tendon collagen [5], and it has been suggested that they result from the formation of advanced glycation end products (AGEs), some of which are covalent cross-links that can alter the behavior of proteins catastrophically. Pentosidine is a well characterized AGE that cross-links proteins via arginine and lysine residues [6]. This may account for arg and lys losses observed in electrophoretic profiles of *in vitro* glycated protein [7]. However, at present there are no data as to the extent to which arg and lys charges are lost when intact collagen fibrils are glycated.

*To whom correspondence should be addressed. Julia C. Hadley, Biophysics Group, Oxford Research Unit, The Open University, Foxcombe Hall, Boars Hill, Oxford OX1 5HR UK. Tel: (01865) 487677; Fax: (01865) 326322; E-mail: j.c.hadley@open.uc.uk.

The positive charges carried by arg and lys are an important factor with respect to the interactions and properties of proteins. For example, electrostatic interactions between adjacent collagen molecules are implicated in the self-assembly and stability of the resultant collagen fibrils [8]. Consequently, any sugar-induced reactions leading to the formation of AGEs that alter the charge distribution, may influence the quaternary structure as well as the interaction of the glycated protein with other proteins. Carboxymethyl lysine, for example, is an AGE that replaces the positive charge on lysine with a negative charge [9, 10].

Heavy metal salts are widely used in electron microscopy to enhance image contrast. In positive staining, unreacted stain molecules are washed away, leaving only molecules which have reacted specifically with the tissue proteins. Collagen fibrils have long been known to exhibit a characteristic positive staining pattern that consists of a series of cross striations with an axial periodicity D, where $D \sim 67$ nm (Figure 1) [11, 12]. There are up to twelve bands within a D-period, designated a1–a4, b1, b2, c1–c3, d and e1, e2, although c3 is rarely discerned and the a and e bands are often difficult to resolve. Using phosphotungstic acid (PTA) as the positive stain, the striations have been shown to be due to the uptake of PTA anions on the positively charged side-chains of the amino acids – lysine, hydroxylysine, arginine and, to a lesser extent, histidine [13, 14].

The positive staining pattern of collagen fibrils is attributable to the architecture of the collagen molecules within the

fibrils. Each collagen molecule consists of three triple-helical α-chains, two α1 chains and one α2 chain, each chain axially staggered by one amino acid residue with respect to the next, in the order α1, α2, α1. These molecules assemble in a parallel array with a staggered association of D or multiples of D [15] (see Figure 1). The D-periodic nature of the fibril arises from the D-stagger. Each D-period is made up of two broad zones; the gap and the overlap. For every five molecules present in the overlap four will be present in the gap zone. The juxtaposition of the constituent positively charged amino acid side chains within the molecule gives rise to the twelve observable bands in PTA stained type I collagen [12].

Figure 2a represents the molecular packing in a portion of the fibril D-period and shows amino acid sequence data.

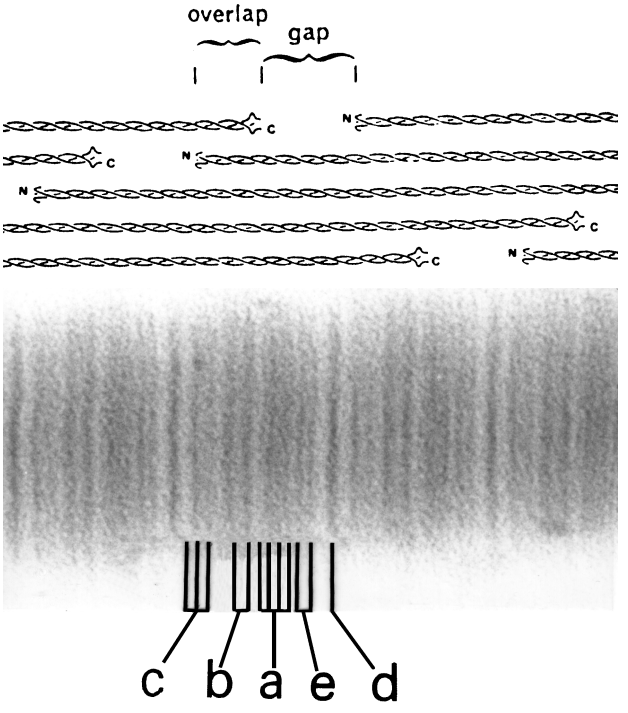

Figure 1. The top of the figure shows the axial arrangement of the triple-stranded collagen molecules within a fibril (after Hodge and Petruská 22). The N- and C- termini of the individual molecules are indicated as is the non-triple-helical conformation of the terminal peptides. Molecules are mutually staggered by integral multiples of D (234 amino acid residues), and this gives rise to an axial periodicity, referred to as the D-period. Each D-period is divided into a 'gap' and an 'overlap' region. One D-period is about 67 nm long in scleral collagen.

The micrograph shows a portion (just under 4 D-periods) of a scleral collagen fibril positively stained with phosphotungstic acid (PTA). The staining bands within one D-period are indicated using the notation of Hodge and Schmitt [19] – a strong d-band, two e-bands (e1 and e2), four a-bands (a1 to a4), two strong b-bands (b1 and b2), three c-bands (c1 to c3). The D-periodicity and alignment of the micrograph matches the molecular diagram above (i.e., bands c3, d, e, a1 and a2 fall in the gap region while bands a4, b and c1 are in the overlap and c2 and a3 are at the gap/overlap junctions.)

The five adjacent molecular segments found in this part of the fibril are presented as triple stranded amino acid sequences (only the positively charged residues are named, all others being represented by dots). The diagram thus illustrates which charged amino acid residues contribute to stain uptake in this region of the fibril. Figure 2b is the

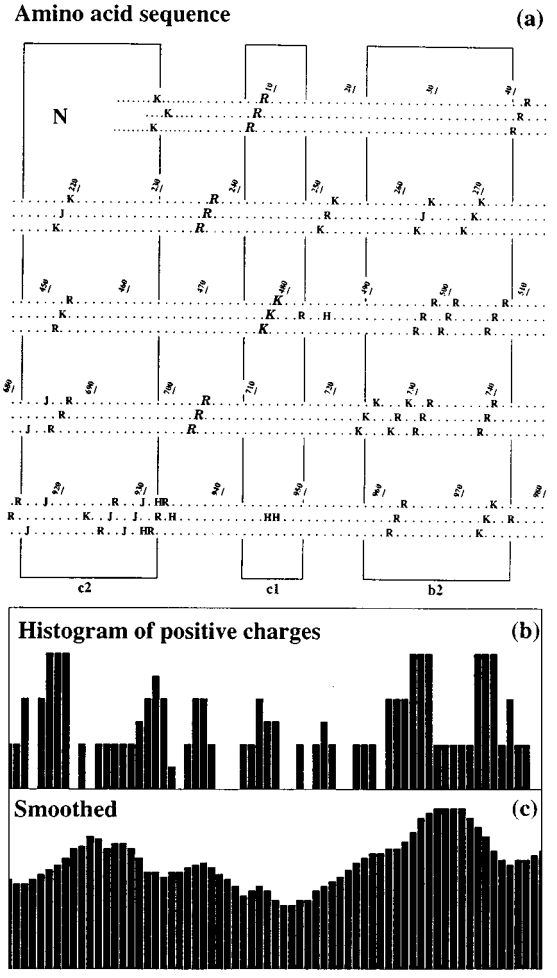

Figure 2. The formation of the histogram of positive charges for the region c2, c1 and b2 of the collagen sequence. (a) Shows the sequence data for c2, c1 and b2. Each molecular segment contributing to the fibril is represented by three amino acid sequences from the three constituent alpha chains, each sequence staggered by one amino acid residue in the order *a*1, *a*2, *a*1 to simulate the arrangement of the alpha chains within the collagen triple helix. Only positively charged residues are included, all others being represented by dots. The N-terminus of a collagen molecule is indicated. Residues are numbered from the N-terminus of the molecule. The diagram shows five such molecular sequences, each axially staggered by 234 residues to simulate the packing of molecules within a collagen fibril. The rectangles indicate how charged side-chains in different molecules tend to line up across the fibril to give rise to staining bands. By adding up the number of positively charged amino acid residues vertically across all molecules a histogram can be formed. Residues shown in italics are those that appear to be unstained in the glycated sample (see Results). (b) Shows the histogram of positive charges. (c) Shows the same histogram smoothed.

corresponding histogram of positive charges, obtained by summing across the constituent molecules and Figure 2(c) shows the histogram of positive charges smoothed by convolution with a top-hat function, to simulate the level of resolution seen in the electron microscope. The sequence data come from the peptide sequencing of calf skin collagen [16, 17]. DNA sequencing of the human procollagen gene for type I collagen has produced data which show good agreement with the calf skin sequence [18, 19]. Any interspecies differences are highly conservative and do not result in alteration of the distribution of charged residues along the molecule.

Reactions that alter the charge distribution along the collagen lead to changes in stain uptake. Comparison between sequence-generated charge histograms and the observed stain distribution have allowed the stagger between adjacent collagen molecules to be determined to less than one residue spacing (i.e., less than 2.9 Å) [20]. Small changes in stain uptake can be detected and the method has provided considerable insight into the reaction mechanisms of a range of different stains [21, 22] and fixatives [14-Review, 23]. The purpose of this study, therefore, is to use collagen as a model system to investigate the extent to which *the surface charge on intact fibrils* is modified by glycation *in vitro* and to elucidate the regions along the fibril that are involved.

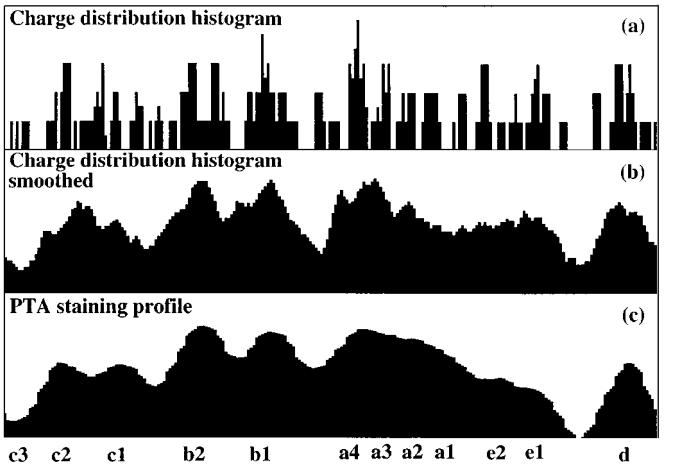

Figure 3. The histogram of positive charges for one D-period and how this compares with a PTA staining profile. (a) Histogram of positive charges for 1 D-period; (b) histogram smoothed; (c) PTA staining profile for the control sample.

Materials and methods

Human scleral collagen was obtained from two donors aged 57 years (donor A) and 65 years (donor B). Corneoscleral samples were obtained from the United Kingdom Transplant Service Eye Bank (Bristol); samples comprise the cornea plus a 5 mm scleral rim. The scleral rim from each donor was split into three pieces, one was kept frozen; this sample was the control. The remaining two pieces of sclera were glycated with 0.5 M fructose in 0.05 M phosphate buffer (PB), at 37 °C. One was glycated for 5 days the last piece was incubated in fructose for 11 days. The collagen was mechanically dispersed using a small amount of liquid nitrogen to aid fracture of the scleral tissue. The collagen fibrils were briefly suspended in a small amount of dH_2O to allow them to be captured on carbon coated grids. They were stained for 1 h, at room temperature, with 2% aqueous PTA, pH 3.2, washed with distilled water to remove unbound stain ions, then allowed to dry at room temperature. The grids were examined using a Philips 301 transmission electron microscope, calibrated using a diffraction grating replica with 2160 lines per mm, and micrographs were taken at high magnification (× 72,000).

Fibrils showing a clear banding pattern were selected for analyses. A number of densitometric traces (each one D-period long) were made from micrographs by scanning along the banding patterns of isolated collagen fibrils in the direction of the fibril axis using a laser densitometer (LKB Ultrascan XL, line beam = 50 μm × 800 μm). To improve the signal-to-noise ratio, 20 such traces were taken and averaged for each sample. The computer averaging has been described previously [24]. Briefly, it involved correcting each trace for any varying background intensity and interpolating it into 234 divisions (so that each division corresponded to one amino acid residue along the fibril axis.) The second trace in a set was then translated stepwise past the first trace in the set (with wrap around), and the similarity between the two traces was quantified by linear regression analysis. In this way, the best axial superposition of the traces could be found, together with the optimum scaling factors. This process was repeated for each subsequent trace in the set. When all the traces had been scaled and shifted axially to match the first, the whole set was averaged. In this way, averaged traces were produced from glycated samples and controls. The averaged traces from glycated samples were subtracted from the control trace to generate difference spectra and these spectra were compared. Averaged traces were also compared to computer-generated histograms derived from the amino acid sequence (see [25] and Figure 2). To illustrate this, Figure 3 shows an averaged PTA staining profile for one D-period (c), compared to the charge distribution histogram of type I collagen. (a) Shows the histogram of charges; (b) shows the histogram smoothed by convolution with a top-hat function.

Results

Figure 4 shows the PTA staining profiles for (a) the control sample; and (b) the sample glycated for 11 days (T_{11}) (donor B). Figure 4(c) shows the difference spectrum (control-T_{11}), on an expanded scale, where a peak signifies a decrease in

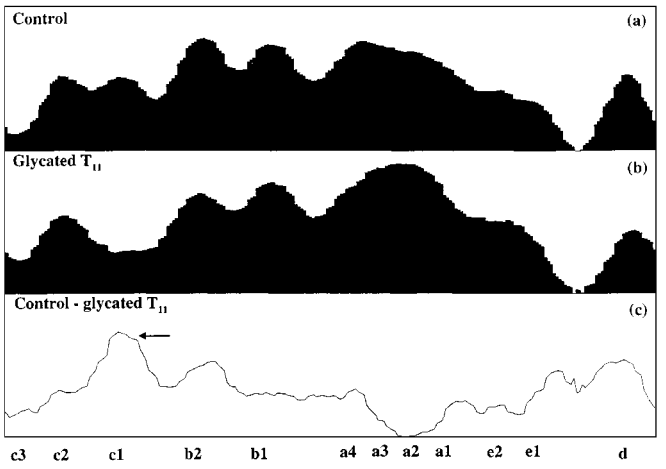

Figure 4. Comparison of the PTA staining profile for the control tissue and the T_{11} tissue (fructated for 11 days). (a) Average of 20 microdensitometric traces for the control tissue. (b) Average of 20 microdensitometric traces for the scleral tissue fructated for 11 days, (c) Difference spectrum, the difference between the control staining profile (shown in (a)) and the staining profile for the glycated tissue, T_{11} (shown in (b). The arrow indicates the major peak (in the c1 band), which shows the area where the major decrease in stain uptake was observed upon glycation. A decrease in stain uptake is evident in three other regions, at b2, d and between e1 and d.

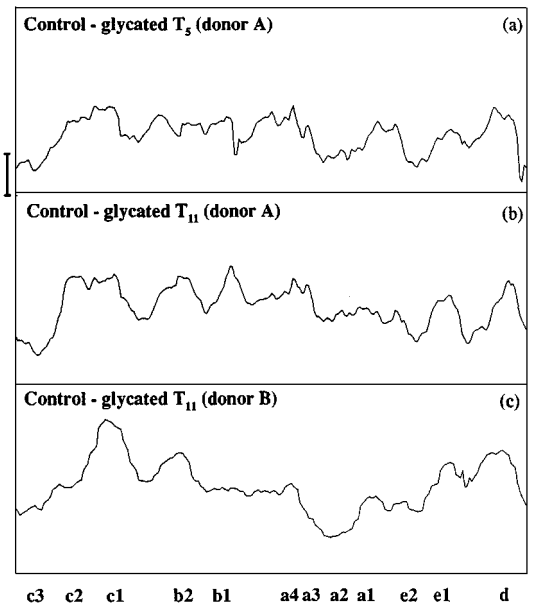

Figure 5. Comparison of the difference spectra for T_5 and two T_{11} samples, plotted on the same scale. The noise level in these traces was quantified by comparing the variances of the noise in the averaged traces (the traces having been corrected for varying background intensities and scaled with respect to each other [31]). (a) control – T_5 (donor A), shows areas of decreased stain uptake. The vertical bar shows the maximum noise level. (b) control – T_{11} (donor A), shows more obvious areas of decrease in stain uptake for example at the c1 band, and smaller peaks at the b2, e1 and d bands. (c) Control – T_{11} (donor B).

stain uptake between the two samples. The arrow shows the major decrease in staining intensity observed at the c1 band. Three other regions of decreased staining intensity are observed after fructation; at b2, d and between e1 and d.

Figure 5(a) and (b) show the difference spectra for T_5 and T_{11} (donor A) compared to the control. Figure 5(c) is the difference spectrum for the second T_{11} sample (donor B) plotted on the same scale. On visual inspection trace (a) shows small peaks at various regions along the D-period, including the c1 region. These peaks are barely greater than the fluctuation due to noise (derived by comparing the variances of the noise levels of the averaged traces). Traces (b) and (c) show clearer peaks above the noise level, indicating that a decrease in staining occurs at certain positions during the first 11 days of glycation. Although Figure 5(b) and (c) show several similar features, they also show some differences, especially in the b1–a4 region. The differences may relate to the original state of glycation of the donor collagen.

An attempt to improve the agreement between the histogram of positive charges and the T_{11} densitometric trace was made by systematically removing various arginine and lysine residues in the c1 region of the sequence data and then correlating the c region of the smoothed histogram with the same region of the T_{11} trace (donor B). This region was chosen because there was a clear change in the staining pattern caused by glycation. An increase in correlation from 0.77 to 0.91 was observed when 12 of the 16 positively charged amino acid residues in this region were omitted. Referring to Figure 2, these residues were arginines 9, 237 and 704; and lysine 479. Removal of other potential glycation sites in this region either did not change, or reduce the correlation. Figure 6 shows the effect on the histogram of positive charges of the removal of these 12 residues. Figure 6a shows the control densitometric trace in this region: (b) shows the histogram of positive charges; (c) shows the effect of removing the contribution from the 12 residues and (d) is the observed T_{11} trace. The similarity between (a) and (b) and now the similarity between (c) and (d) can be seen.

Discussion

Histochemical studies have shown that PTA staining of proteins is a predominantly electrostatic interaction, in which the PTA anions bind to positively charged amino acid side-chains [26–28]. Silverman and Glick [29], found that the stoichiometry of this reaction varies for different proteins, however due to the abundance of arginine and lysine in collagen, PTA stain uptake is relatively high, with all available residues apparently reacting with the stain. For this reason, PTA staining of collagen provides an excellent

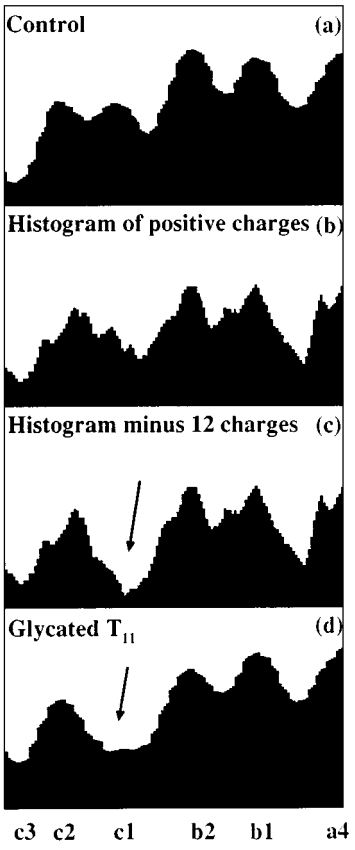

Figure 6. The effect of changing the weighting of the histogram at the c1 region. (a) Shows the control staining profile. (b) Shows the histogram derived from sequence data. Note the similarity with the control profile (a). (c) Shows the modified histogram with 12 positive residues removed from the c1 region. (d) Shows the staining profile for the T_{11} sample. The correlation of the c region of this profile with the c region of the histogram is increased on removal of the 12 residues. This signifies that the altered histogram is more similar to the T_{11} trace than the original sequence-derived histogram.

model system to investigate reactions that alter protein charge. In fact, there is probably no method as sensitive to detect changes in the positive charge on the surface of intact collagen fibrils.

Since we are unaware of any mechanism that could induce the formation of extra positive charges when a protein is glycated, we have assumed that changes in stain uptake reflect the loss of positive charges. Furthermore, a decrease in stain uptake is only identified as such if the peak is relatively large compared to the background noise of the difference spectra.

A decrease in PTA stain uptake at specific regions in samples glycated for 5 or 11 days, when compared to the control, implies that stain inhibition (and probably AGE formation) does not occur at all available sites along the collagen fibril. The sensitivity of this technique allows the regions where fructose binds to be located, but it does not allow the sites to be identified at the level of individual amino acids. Nevertheless, it is interesting that the most marked decrease in stain uptake occurs at the c1 region, and a number of positively charged amino acid residues in this region could be involved in glycation. Correlation analyses of the c region has shown up to 12 residues which could be involved; on removal of these residues from the histogram of positive charges the correlation of the c region of the histogram with the c region of the T_{11} trace increases by nearly 20%. One site in this region has already been shown to be preferentially glycated in rat tail tendon type I collagen [2]. However, it should be noted that the failure of several amino acids in a particular region to stain with PTA does not indicate *per se* that all these residues are glycated or involved in cross-links. It is possible that the presence of sugar molecules or AGEs at one site could prevent stain attachment to nearby sites by steric exclusion.

Longer incubation (T_{11}) probably allows the conversion of some of the fructated residues to advanced products such as AGEs. This might explain why we found a more apparent decrease in stain uptake in specimens incubated for the longer period (T_{11}), compared to the sample incubated for 5 days in fructose. The T_5 sample did show a decrease in stain uptake but not to the same extent as the T_{11} sample.

There is relatively little data in the literature identifying preferential glycation sites in collagen. Brennan [30] found that glycation was not very specific since most cyanogen bromide peptides in rat tail tendon bound some glucose. Tanaka *et al.* [31] partially characterized a CNBr digest fragment of collagen containing a fluorescent crosslink produced by glycation and radiochemical labeling of glycation sites has been used to locate sites in CNBr fragments [32]. Wess *et al.* [33, 34] used neutron diffraction methods to locate glycation sites in normal and diabetic tendons. They found that glycation occurs at different rates within the collagen unit cell but suggested hydroxylysine residues together with lys 434 and lys 855 as possible sites. LePape *et al.* [35] investigated the uptake of radiolabeled glucose and found that glucose uptake occurred to a lesser extent in diabetic collagen. Their hypothesis for this was that lower uptake of glucose in the diabetic tissue is due to the preferential sites along the collagen having already bound sugar. Glycated lysine residues have been identified close to carboxylic acid residues [36] and it is thought that these latter residues may catalyse the Amadori rearrangement. Baynes *et al.* [37] found that the most reactive sites for glycation an RNase and Hb are found in the high affinity binding sites for phosphate ions or organic phosphate. They also found that these regions are relatively basic, containing arginine and histidine residues, and suggested that these residues could be catalyzing the Amadori rearrangement. Reiser *et al.* [2] found a number of lysine residues to be preferentially glycated in type I collagen, for example, lysine 434, lysine 924, lysine 453 and lysine 479. Out of these, lysine 479 is the only one whose position makes it difficult to account for its preferential glycation. The other three lysines

are situated next to acidic residues which may be acting as catalytic agents for glycation.

We have shown that the surface charge of human scleral collagen fibrils can be altered *in vitro* by incubation with fructose. The results also indicate that several sites along the fibril have the potential to be glycated, although the extent to which individual residues react *in vivo* remains to be investigated. Further work is required to identify the factors that predispose specific residues to become glycated, particularly in intact proteins, since the majority of work carried out so far has involved the analysis of protein fragments.

Acknowledgments

The authors thank the United Kingdom Transplant Service Eye Bank (Bristol) for supplying the scleral tissue used in this study. We would like to thank B.T./Research into Ageing (Nageena S. Malik) and the Wellcome Trust (Keith M. Meek) for financial support.

References

1 Baynes JW, Monnier VM (1989) *Prog Clin Biol Res* **304**: 1–410.
2 Reiser KM, Amigable M, Last JA (1992) *J Biol Chem* **267**: 24207–16.
3 Malik NS, Meek KM (1994) *Biochem Biophys Res Communications* **199**: 683–6.
4 Malik NS, Moss SJ, Ahmed N, Furth AJ, Wall RS, Meek KM (1992) *Biochim Biophys Acta* **1138**: 222–8.
5 Tanaka S, Augad G, Eikenberg EF, Brodsky B (1988) *J Biol Chem* **263**: 17650–7.
6 Sell D, Monnier VM (1989) *J Biol Chem* **264**: 21597–602.
7 Kato H, Hayase F, Shin DB, Oimoni M, Baba S (1989) *Prog Clin Biol Res* **304**: 69–84.
8 Hulmes DJS, Miller A, Parry DAD, Piez KA, Woodhead-Galloway J (1973) *J Molec Biol* **79**: 137–48.
9 Dyer DG, Blackledge JA, Thorpe SR, Baynes JW (1991) *J Biol Chem* **266**: 11654–60.
10 Ahmed MU, Thorpe SR, Baynes JW (1986) *J Biol Chem* **261**: 8816–21.
11 Wolpers C (1944) *Virchows Arch Path Anat Physiol* **312**: 292–302.
12 Chapman JA (1974) *Connect Tiss Res* **2**: 137–50.
13 Hodge AJ, Schmitt FO (1960) *Proc Natn Acad Sci USA* **46**: 186–97.
14 Chapman JA, Tzaphlidou M, Meek KM, Kadler KE (1990) *Electron Microsc Rev* **3**: 143–82.
15 Hodge AJ, Petruska JA (1963) In *Aspects of Protein Structure* (Ramachandran GN, ed) pp 289–300. New York: Academic Press.
16 Chapman JA, Holmes DF, Meek KM, Rattew CJ (1981) In *Structural Aspects of Recognition and Assembly in Biological Macromolecules* (Balaban M, Sussman JL, Traub W and Yonath A, eds) pp 387–401. Rehovot and Philadelphia; Balaban International Science Services.
17 Weiss JB, Jayson MIV (1982) *Collagen in Health and Disease*. Edinburgh;Churchill Livingstone.
18 Kuivaniemi H, Tromp G, Chu M, Prockup DJ (1988) *Biochem J* **252**: 633–40.
19 Tromp G, Kuivaniemi H, Stacey A, Stukatat H, Balkdwin CT, Jaenisch R, Prickup DJ (1988) *Biochem J* **253**: 919–22.
20 Meek KM, Chapman JA, Hardcastle RA (1979) *J Biol Chem* **254**: 10710–14.
21 Tzaphlidou M, Chapman JA, Meek KMA (1982) *Micron* **13**: 119–31.
22 Tzaphlidou M, Chapman JA, Al-Samman MH (1982) *Micron* **13**: 133–45.
23 Tzaphlidou M, Chapman JA (1984) *Micron and Microscopica Acta* **15**: 69–76.
24 Tzaphlidou M, Hardcastle RA (1984) *Int J Biomed Computg* **15**: 113–20.
25 Chapman JA, Hardcastle RA (1984) *Connect Tiss Res* **2**: 151–9.
26 Quintarelli G, Zito R, Cifonelli JA (1971) *J Histochem Cytochem* **19**: 641–7.
27 Quintarelli G, Cifonelli JA, Zito R (1971) *J Histochem Cytochem* **19**: 648–53.
28 Quintarelli G, Bellocci M, Geremia R (1973) *J Histochem Cytochem* **21**: 155–60.
29 Silverman L, Glick D (1969) *J Cell Biol* **40**: 761–7.
30 Brennan M (1989) *J Biol Chem* **264**: 20953–60.
31 Tanaka S, Avigad G, Brodsky B, Eikenberry EF (1988) *J Biol Chem* **263**: 17650–7.
32 LePape A, Guitton, J-D, Muh J-P (1984) *FEBS Letters* **170**: 23–7.
33 Wess TJ, Miller A, Bradshaw JP (1990) *J Mol Biol* **213**: 1–5.
34 Wess TJ, Wess L, Miller A, Lindsay RM, Baird JD (1993) *J Mol Biol* **230**: 1297–303.
35 LePape A, Guitton JD, Muh JP (1981) *Biochem Biophys Res Commun.* **100**: 1214–21.
36 Shapiro R, McManus MJ, Zalut C, Bunn HF (1980) *J Biol Chem* **255**: 3120–27.
37 Baynes JW, Watkins NG, Fisher CI, Hull CJ, Patrick JS, Ahmed MU, Dunn JA, Thorpe SR (1989) in *The Maillard Reaction In Aging, Diabetes and Nutrition* (Baynes JW, Monnier VM, eds) pp 43–68. New York: Alan R. Liss.

Received 3 October 1997, revised 21 January 1998, accepted 2 February 1998