Glucosepane is associated with changes to structural and physical properties of collagen fibrils

Anthony Nash [1], Maria Notou [2], Andrea F. Lopez-Clavijo [3], Laurent Bozec [4], Nora H. de Leeuw [5] and Helen L. Birch

University College London, *Institute of Orthopaedics and Musculoskeletal Science, Stanmore Campus, Royal National Orthopaedic Hospital, Stanmore HA7 4LP, UK*

Correspondence to Helen L. Birch: *anthony.nash@ndcn.ox.ac.uk, m.notou@ucl.ac.uk, l.bozec@utoronto.ca, deLeeuwN@cardiff.ac.uk, h.birch@ucl.ac.uk*
https://doi.org/10.1016/j.mbplus.2019.100013

Abstract

Collagen glycation, and in particular the formation of advanced glycation end-product (AGE) crosslinks, plays a central role in the ageing process and in many of the long-term complications of diabetes. Glucosepane, the most abundant and relevant AGE crosslink, has been suggested to increase the stiffness of tissue and reduce its solubility, although no evidence is available concerning the mechanisms. We have used a combination of computational and experimental techniques to study a collagen-rich tissue with a relatively simple organisation to further our understanding of the impact of glucosepane on the structural and physical properties of collagen fibrils. Our work shows that glucosepane levels increase dramatically in aged tendon tissue and are associated with the reduced density of collagen packing and increased porosity to water molecules. Our studies provide the basis to understand many of the tissue dysfunctions associated with ageing and diabetes across a range of different tissues types.

Introduction

Glucosepane has emerged as the most abundant and clinically relevant advanced glycation end-product (AGE) crosslink in type I collagen-rich tissues [1] and it is associated with many of the long-term complications of diabetes, including nephropathy, neuropathy and retinopathy [2]. More recently, increases in glucosepane have been correlated with the progression of the age-related degenerative disease osteoarthritis [3]. Despite the association of glucosepane with severe tissue dysfunction, the mechanism by which these changes are imposed on the tissue remains largely unexplored.

Like all AGE crosslinks, glucosepane is formed following a series of spontaneous adventitious modifications to the amino group on the side-chain of lysine and arginine residues in protein polypeptide chains [4]. Although the formation of AGE crosslinks has been known about for almost forty years [5], the AGE crosslink glucosepane was only identified relatively recently [6]. In the case of glucosepane, its formation is initiated by the amino group of a lysine residue undergoing a nucleophilic attack by a reactive carbonyl group of glucose in a reaction commonly known as the Maillard reaction. The product of this reaction undergoes a rearrangement to form the more stable Amadori product. This is followed by a series of reactions, where the carbonyl group is passed down the backbone of the glucose molecule before cyclisation takes place; further reaction with the amino group on an arginine residue side-chain then forms the glucosepane crosslink [6,7].

Glucosepane is generating much interest, particularly with respect to the mechanisms of ageing, although the complex chemistry of glucosepane has hindered progress in assessing its impact on function [8]. Glucosepane levels were first measured

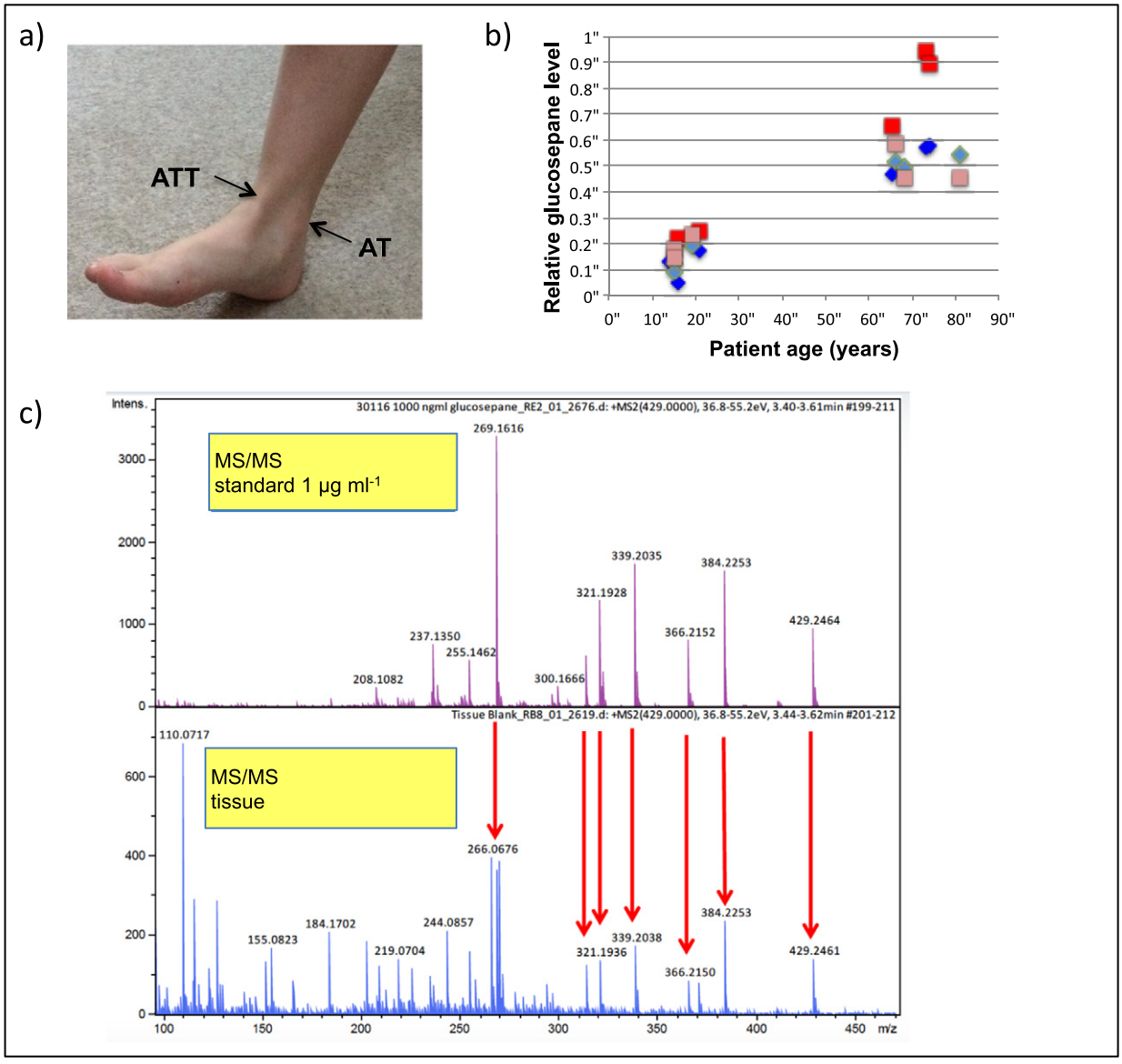

Fig. 1. Human tendon tissue was subjected to sequential enzymatic digestion to release glucosepane, which was separated and quantified using LC Q-TOF MS (a). Relative levels of glucosepane in Achilles and anterior tibialis tendon tissue against patient age. Red squares represent anterior tibialis tendon tissue (dark red = male, light red = female), blue diamonds represent Achilles tendon tissue (dark blue = male, light blue = female) (b). The fragmentation pattern of glucosepane in the tissue sample matched that of the glucosepane standard preparation (c).

in human dermal collagen and glomerular basement membrane in diabetic and non-diabetic subjects [9]. Levels of glucosepane in dermal collagen were found to increase significantly with increasing age in non-diabetic samples and to be significantly higher in diabetic skin samples [9]. To our knowledge, glucosepane levels have not been quantified in other collagen-rich human tissues such as tendon, ligament, bone, cartilage and blood vessels, although it has been shown robustly that glucosepane levels in dermal collagen are associated with micro-vascular disease in other organs [9] and they are an accurate predictor of future microvascular complications in diabetic patients [10]. Furthermore, plasma levels of glucosepane were found to be increased sixfold in patients with advanced osteoarthritis relative to healthy controls [3]. Disease progression is thought to be manifested by an AGE crosslink-induced stiffening of the tissue and reduced susceptibility to proteolytic degradation; both factors are also associated with advancing chronological age.

Tendon tissue is one of the most collagen-rich tissues in the body. Tendons are rope-like structures, connecting muscle to bone, which are

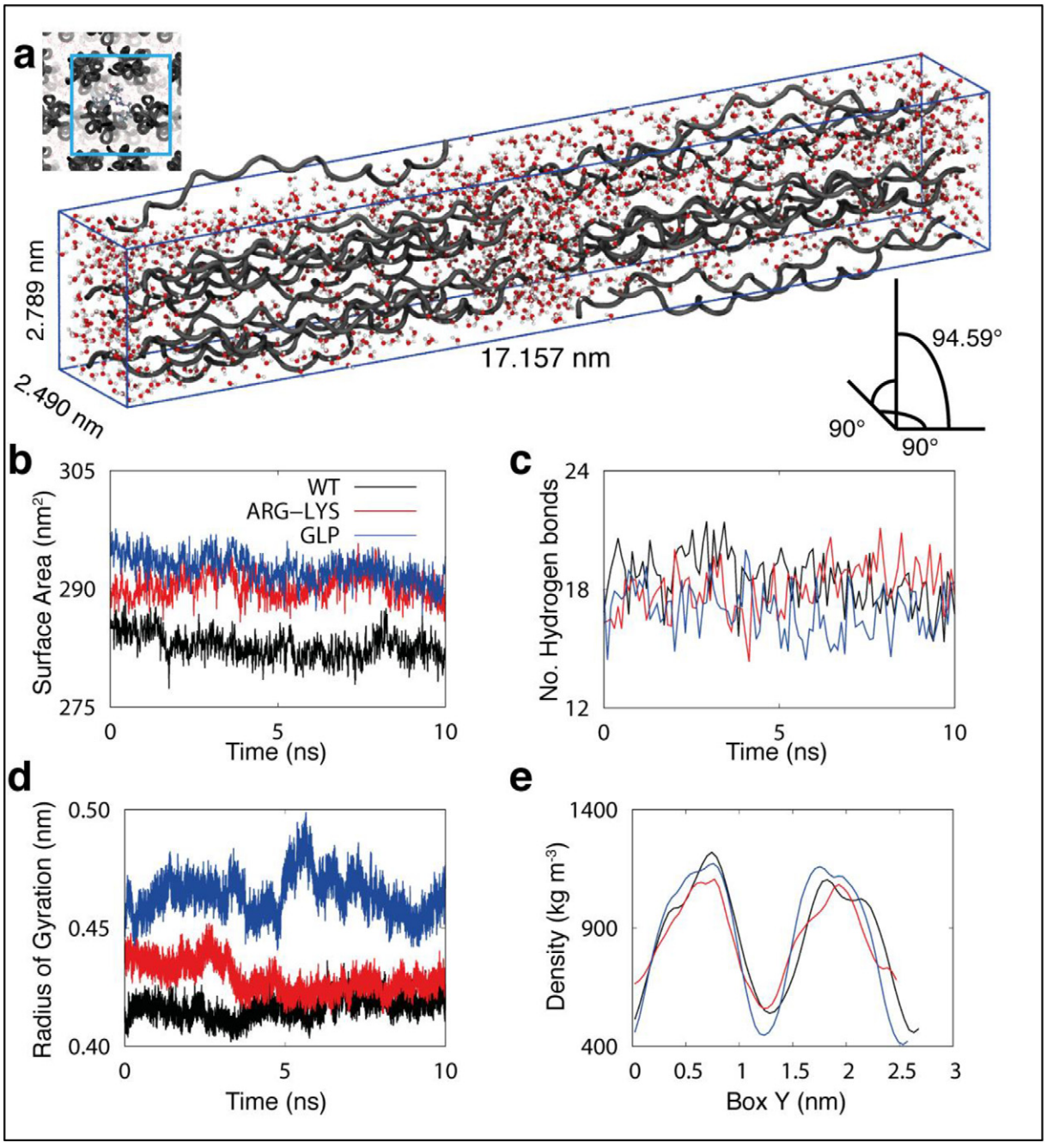

Fig. 2. Crystallographic dimensions of lengths and angles (a) accompanied by a cross-section representation of the unit cell periodically replicated (insert). Single collagen protein surface area (b), the number of inter-polypeptide hydrogen bonds found across a single collagen protein (c), the radius of gyration of a single collagen protein (d) and a systemic protein density plot (e).

responsible for transmitting tensile forces generated by muscle contraction to the bony skeleton to bring about movement around a joint. The high tensile strength required for their function is provided by type I collagen, which forms the majority (about 80%) of the dry weight of tendon [11]. In addition to high tensile strength, the stiffness and elasticity of tendon are highly tuned and tendon-specific for efficient function [12,13]. These properties stem from the precise arrangement of collagen throughout the hierarchical structure of the tissue from collagen molecules to fibrils, fibres, fascicles and finally the whole tendon (for a review of tendon structure see Thorpe and Screen (2016) [14]).

Collagen in tendon has a long half-life, estimated to be around 200 years in equine flexor tendon

tissue [11], with almost no turnover in human Achilles tendon [15]; as a result, the collagen is vulnerable to glycation and crosslinking. Indeed, other more easily quantified measures of AGE formation, such as collagen-linked fluorescence and the AGE crosslink pentosidine [11], show a significant positive correlation with donor age, providing evidence of an accumulation of age-related post-translational modifications to collagen and suggesting that glucosepane is highly likely to accumulate in tendon tissue. Interestingly, the incidence of tendon injury is well known to increase with older age in human [16–18] and equine [19–21] species, suggesting a decline in the mechanical competence of tendon with ageing.

Given the high content of type I collagen molecules in tendon and its relatively simple parallel arrangement into fibrillar structures, tendon has the potential to provide an ideal model to study the impact of increasing glucosepane levels on collagen structural and molecular properties. In this study, we have used a combination of computer simulations and experimental work on *ex vivo* human tissue to determine the impact of glucosepane crosslinks on the material properties of tendon tissue at a nanoscale level.

Results and discussion

Glucosepane levels in tendon tissue

In our studies, we have used human Achilles and anterior tibialis tendon tissue from young (16.7 ± 2.7 years, n = 6) and old (71.2 ± 6.0 years, n = 6) groups of donors to quantify the levels of glucosepane and determine changes in levels with age (Fig. 1a & b). A glucosepane standard was prepared by incubating lysine, arginine and glucose followed by liquid chromatography (LC) separation and purification. High resolution-high accuracy mass spectrometry analysis confirmed the expected monoisotopic mass of 429.2461 Da with an error of −1.18 ppm (Fig. 1c). Human tendon tissue was digested using sequential enzymatic digestion [1], and glucosepane was separated and relative levels quantified using LC mass spectrometry. Our results show a highly significant ($p < 0.001$) four-fold increase in glucosepane levels in the old group of donors compared to the young group for the Achilles and anterior tibialis tendon. The fragmentation pattern of glucosepane in the tissue samples was identical to that of the standard preparation.

Glucosepane impact on the backbone conformation of collagen

We have used molecular dynamics (MD) simulations to investigate the impact of glucosepane on collagen molecule packing within the fibril. MD simulations were carried out using GROMACS 5.1.2 [22] with a collagen-like peptide in a unit cell repeated *via* periodic boundary conditions in all three dimensions (Fig. 2a). Three models were created for study; a wild-type, which is effectively an extension to the unit-cell of the 4Z1R crystal structure of a collagen-like peptide; the same structure but modified with a lysine and arginine side chain substituted between two polypeptide chains at residue positions 17 and 49, respectively (pre-cross-link formation); and the structure modified with a glucosepane crosslink between two polypeptide chains of the same collagen molecule, at positions 17 and 49. In the pre-cross-linked and cross-linked models, only one collagen molecule from the eight was modified in the final extended unit cell. Each system was subjected to a series of equilibrium simulation steps, as described in the methodology. After equilibration, each system was simulated for 20 ns using the NPT ensemble. The initial 10 ns of equilibration were discarded and all analyses were performed over the remaining 10 ns trajectories.

Structural changes to the collagen molecule were investigated first by using a rolling sphere with a radius of 1.2 nm to calculate the total solvent-accessible surface area (SASA) [23] of the peptide backbone atoms for each of the three models. Each set of surface area values were shown to remain consistent (Fig. 2b) providing evidence suggesting a well-equilibrated system. The addition of an arginine and lysine sidechain contributed to a 7.5 nm^2 increase in surface area compared to the wild-type, whereas the surface area increased by 9.8 nm^2 as a result of the inclusion of a glucosepane crosslink. The increase in SASA is indicative of backbone access increasing in the presence of glucosepane, due to an increase in distance between collagen polypeptide chains.

Second, we investigated whether the change in SASA was reflected by a change in the number of hydrogen bonds within the single collagen molecule. A hydrogen bond in this analysis was defined as having the frequently used distance cut-off of 0.35 nm and angular cut-off of 30 degrees. Carbonyl oxygen atoms and nitrogen atoms were regarded as hydrogen acceptors, whilst hydroxyl group hydrogen atoms and amine hydrogen atoms were regarded as hydrogen atom donors. The results show that the increase in the SASA is associated with a decrease in the number of hydrogen bonds. The number of hydrogen bonds was consistently between 14 and 22 (Fig. 2c), with an average of 18.7 bonds for the wild-type, 18.0 bonds due to the inclusion of lysine and arginine residues, and a drop to 16.9 bonds in the presence of the glucosepane crosslink, due to a parting of the crosslinked backbone.

Third, the spatial distribution of the peptide atoms about the peptide principal axis was calculated over

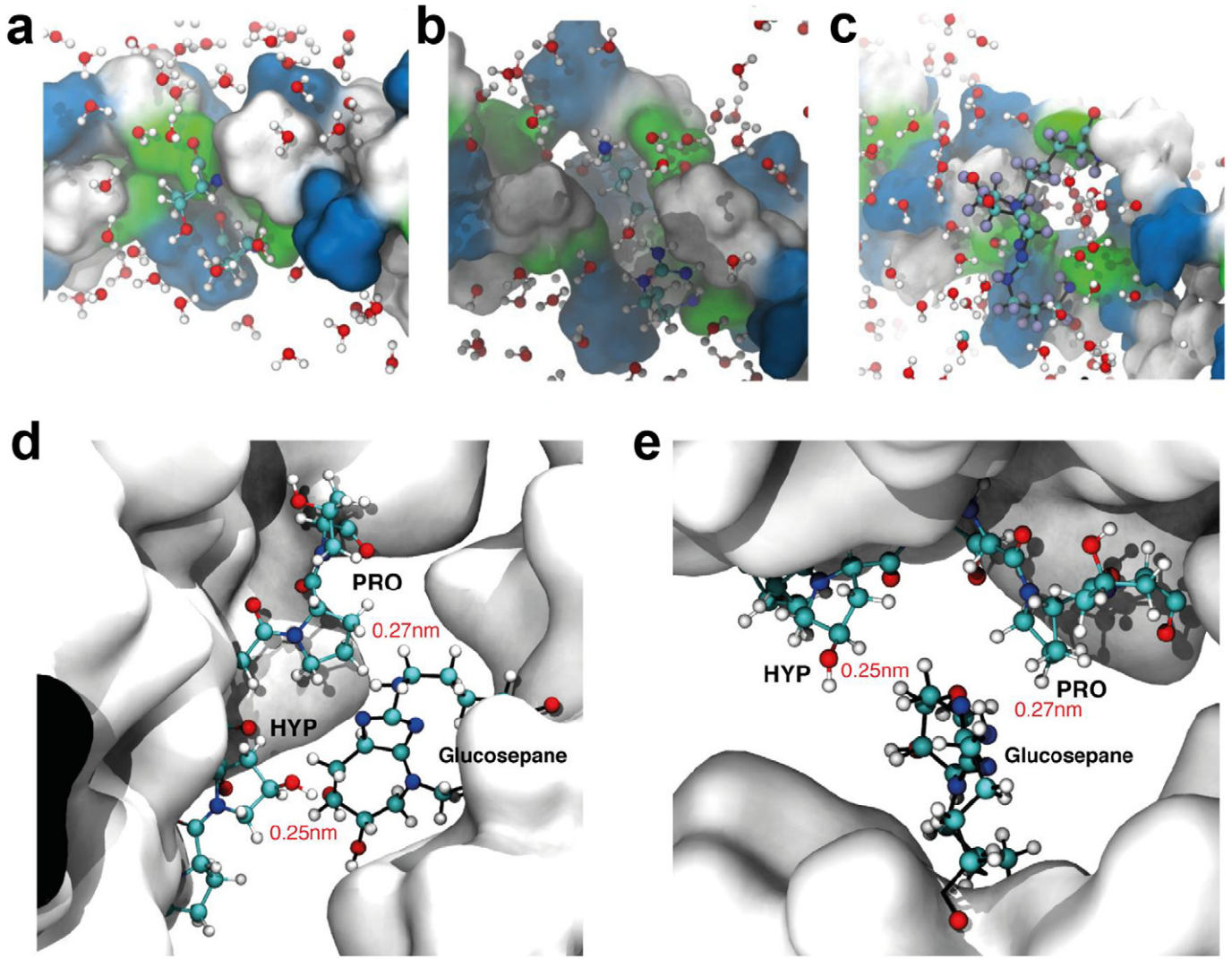

Fig. 3. The diffusion of interstitial water as seen in the wildtype (a), with lysine and arginine amino acids (b) and with glucosepane (c): blue - basic, green - polar, and white - nonpolar. A representative structure of the position of glucosepane relative to the neighbouring collagen (d) and rotated 90 degrees (e).

the duration of each production simulation. Calculation of the radius of gyration remained consistent within the time frame (Fig. 2d) and the error was insignificant (in the magnitude of 10^{-5}). Given the flexibility and length of the arginine and lysine side chains, the radius of gyration increased by 0.0123 nm in the pre-cross-linked model compared to the wild-type collagen. The presence of glucosepane increased the radius of gyration by as much as 0.0473 nm compared to the wild-type indicating a greater degree of conformational change to the collagen protein.

Finally, the density profile, measured along the unit cell width for the complete peptide content for each system was calculated. Each system was centred to factor in fluctuations to the unit cell volume and the calculation was partitioned into 60 bins and data points were collected at 20 ps step intervals. The peptide density plots for each of the three models show two peptide peaks (Fig. 2e) as expected due to the alignment of the collagen molecules within the fibril and bound by the unit cell (see Fig. 2a insert). The wild-type and the pre-crosslinked model yield similar profiles; of particular interest was the similarity in a drop of collagen density at approximately 1.4 nm along the Y-axis of the unit cell; this drop was significantly greater in the presence of glucosepane, presumably due to the bulky size of glucosepane that prohibits close interactions between neighbouring collagen molecules.

Overall these simulations provide evidence that the presence of intra-molecular glucosepane disrupts the local structure of the collagen molecule by opening the triple helix of the collagen molecule and potentially disrupting routine packing between collagen molecules.

Influence of glucosepane on the diffusion of water within the collagen structure

The collagen structure includes a significant number of structural water molecules within its framework. Structural water molecules are hydrogen-bonded within the collagen triple helix [24] and form part of the primary hydration or bound water, along with water molecules forming hydrogen bonded chains lying on the surface of the collagen molecule in the cleft of the collagen triple helix. Secondary or free water molecules reside within and between collagen fibrils. Given the impact of glucosepane on the collagen structure, the diffusion of water within and around the collagen molecule was investigated. Calculations of water diffusion were made using the Einstein relation and least squares fitting of a straight line through the mean squares displacement of all water molecules over the duration of a simulation. The results showed an increase in water diffusion by $0.0661 \times 10^{-5} \pm 0.0172 \times 10^{-5}$ cm^2 s^{-1} in the lysine/arginine model and an increase of $0.0956 \times 10^{-5} \pm 0.0267 \times 10^{-5}$ cm^2 s^{-1} in the glucosepane model compared to the wild-type collagen.

Detailed observation of the simulation trajectories revealed that the water dynamics were unique to each model. In the wild-type model, interstitial water was present and the signature triple helical fold was consistent throughout the polypeptide chains (Fig. 3a). Upon substituting arginine and lysine side chains, there was a noticeable relaxation in the triple helical fold, with water gaining access to the backbone atoms at the substituted sites (Fig. 3b). Finally, the inclusion of a glucosepane cross-link resulted in further separation in the triple helical fold at the immediate site by increasing the distance between two of the three-polypeptide chains, and as a result, water was seen to diffuse through the collagen molecule (Fig. 3c). As expected, water formed transitory short-range interactions with the polar hydroxyl group of the seven-membered ring of glucosepane. The bulk of the crosslink extended into the interstitial collagen space (Fig. 3d & e) and remained positioned within the reach of short-range interactions with the three neighbouring collagen molecules. The aliphatic side chains of the cross-link remained within close proximity of the pyrrolidine of neighbouring proline residues (Fig. 3d & e), whilst the polar seven-membered ring remained consistently lodged between collagen molecules, and neighbouring hydroxyproline and proline backbone atoms and water molecules rapidly exchanged with the polar groups of glucosepane.

These simulations suggest that water molecules readily diffuse between collagen polypeptide chains and that glucosepane positions itself to maximise polar contacts with neighbouring polar side chains.

Experimental determination of tendon tissue water content

To test the predictions from the computational studies, the water content in the young and old groups of tendon tissue was investigated, first by drying under vacuum, and second using the thermo-analytical technique of differential scanning calorimetry (DSC). DSC is able to detect physical transformations in the tissue as specific heat capacity changes; using the latent heat of fusion, melting and evaporation allows calculation of tissue water content [25]. In these experiments tendon tissue was cooled to –40 °C and then heated to 160 °C in a perforated pan. The exothermic peak corresponding to water freezing and the endothermic peak due to ice melting were used to calculate free water content and the endothermic peak resulting from water evaporation was used to calculate total water content (free and bound) (Fig. 4a). The results showed that the total water content of tendon

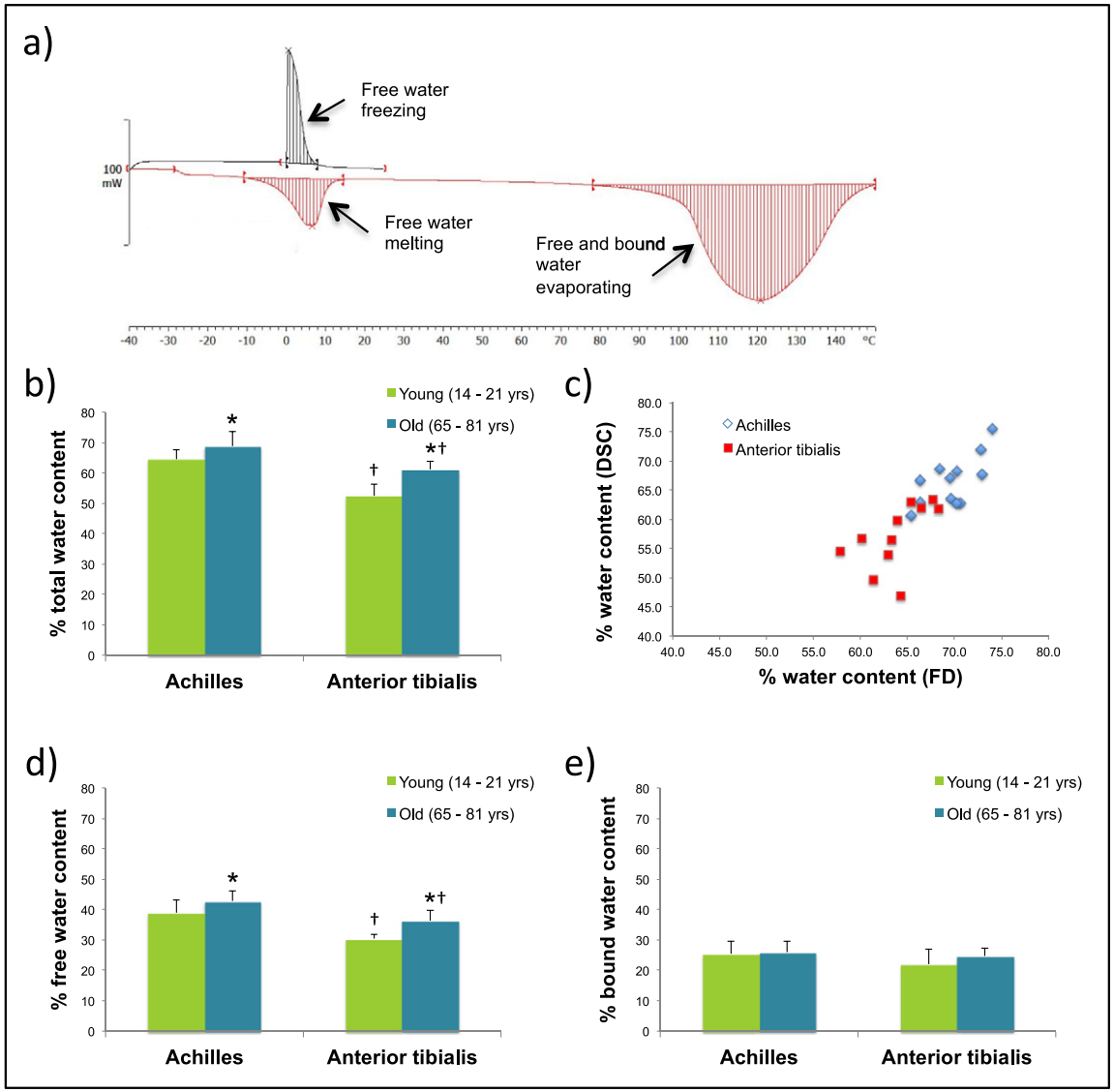

Fig. 4. DSC thermogram of human tendon tissue showing physical transformations of water within the tissue (a) and total water content calculated from the evaporation peak (b). Total water content showed a good correlation (r^2 = 0.66) between the two different methods (freeze-drying and DSC) of quantification (c). Free water content (d) and bound water content (e) in young and old groups of Achilles and anterior tibialis tendon tissue. Data are shown as mean ± SD. * represent a significant difference relative to the young group of tendons and † a significant difference relative to Achilles tendon.

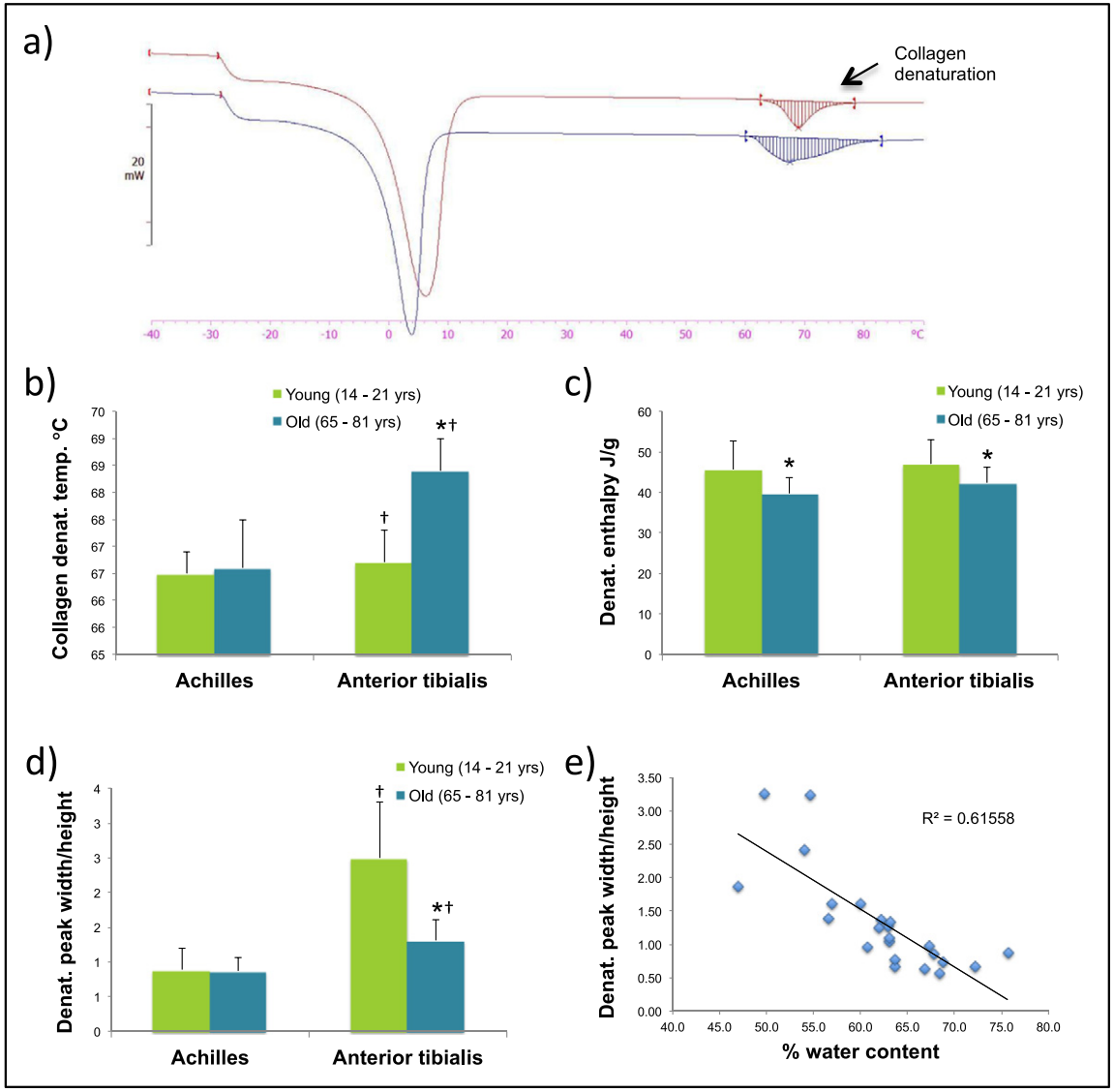

Fig. 5. Typical DSC thermogram of human anterior tibialis tendon tissue showing denaturation of collagen within the tissue from two representative samples (red trace = old tissue, blue trace = young tissue) (a), peak temperature of collagen denaturation (b), enthalpy of collagen denaturation (c) and the collagen denaturation peak width/height ratio (d). Data are shown as mean ± SD. * represent a significant difference relative to the young group of tendons and † a significant difference relative to Achilles tendon. Total water content showed a significant negative correlation (r^2 = 0.62) with the shape of the collagen denaturation peak (e).

tissue was significantly ($p < 0.001$) higher in the Achilles tendon than the anterior tibialis tendon and was significantly ($p < 0.001$) higher in the old groups for both tendon types (Fig. 4b). Water content calculated by weight loss following drying under vacuum showed close agreement ($r^2 = 0.66$) with that obtained using DSC (Fig. 4c). Free and bound water pertains to the location of water molecules within the collagen supramolecular assembly [26] and free water molecules are more easily removed. The higher water content detected in our samples appeared to be due to significantly higher free water content ($p < 0.001$ between tendon types, $p = 0.002$ between age groups) (Fig. 4d) rather than bound water (Fig. 4e), suggesting that the collagen molecules might be less densely packed within the fibril in older tissue. In other skeletal tissues such as articular cartilage, the glycosaminoglycan side-chains of proteoglycans play an important role in attracting water into the matrix. In our study however, the GAG content did not differ significantly between old and young groups of tendons (Supplementary fig. 1 and Supplementary note 1). Previous work has demonstrated that glucosepane is very hydrophilic [27] and may therefore be responsible for the higher water content detected in the tendon tissue from old donors.

Thermo-kinetic properties of collagen in tendon tissue from young and old donors

Collagen molecules are composed of three polypeptide chains (α chains) each with a left-handed helix formation and wrapped around each other to form a right-handed super-helix [28]. The collagen molecule is stabilised by the water-bridged hydrogen bonds within the helix (as described above) and, following fibrillogenesis, further stabilised by enzyme-mediated intermolecular crosslinks between collagen molecules. When heated above physiological temperatures the water-bridged hydrogen bonds become less stable and eventually break thereby allowing the triple helix to unwind. The denaturation temperature increases substantially as a result of collagen crosslinking and depends, in part, on the number of crosslinks present [29]. We have used DSC to measure the thermokinetic properties of collagen in the young and old groups of Achilles and anterior tibialis tendons (Fig. 5a). In these experiments, the tendon tissue was heated in a sealed pan to 90 °C prior to the water content measurements described above.

The results show a collagen denaturation temperature of between 65 and 69 °C. The denaturation temperature was significantly ($p = 0.002$) higher in the anterior tibialis tissue compared to the Achilles tendon tissue and significantly ($p = 0.006$) higher in the old group of donors compared to the young group for the anterior tibialis tendon (Fig. 5b). A higher peak temperature for collagen denaturation suggests greater cross-linking between collagen molecules. Quantification of lysyl oxidase-mediated pyridinoline crosslinks showed a significant decrease in total pyridinoline crosslinks in old groups of tendons compared to young (Supplementary fig. 2a, b & c and Supplementary note 2), implicating AGE crosslinking as the likely cause for increased denaturation temperature. The unexpected relatively low denaturation temperature for the old group of Achilles tendons compared to the old group of anterior tibialis tendons, suggest there may be fragments of collagen retained in the matrix, as has been previously observed during ageing in the functionally equivalent and frequently injured equine superficial digital flexor tendon [11].

The enthalpy of denaturation of collagen is provided by the area of the collagen denaturation peak, which has been shown to relate to the density of packing within the collagen fibril as described by the “polymer in a box” theory [30]. The enthalpy of denaturation was significantly lower ($P = 0.037$) for the old groups of tendons, suggesting that the collagen molecules are less confined within the fibril structure (Fig. 5c). These results are in keeping with a higher free water content, thereby reducing the density of collagen within the fibril. The shape of the denaturation peak, when the scanning rate is constant, is related to the activation energy required for triple helix unfolding [31]. A higher activation energy is the consequence of a greater number of hydrogen bonds, either through hydroxyproline or water-bridged bonding between the three polypeptide chains and is observed by a narrower denaturation peak. The peak width to height ratio shows a significantly sharper peak ($p < 0.001$) in Achilles tendon tissue compared to anterior tibialis tendon tissue, and a significantly sharper peak in the old tissue ($p = 0.002$), particularly for the anterior tibialis tissue (Fig. 5d) showing a higher activation energy for collagen denaturation. As the collagen turnover is essentially zero in human tendon tissue [15], it is likely that the increase in activation energy results from an increase in water content rather than a change in the hydroxyproline level. Indeed the activation energy, measured by peak width to height ratio, shows a significant negative correlation ($p < 0.001$) with tendon water content (Fig. 5e).

Conclusions

Our study is the first to show that glucosepane levels accumulate in human Achilles and anterior tibialis tendon with increasing chronological age. Direct evidence of the impact of glucosepane on collagen properties is not possible at present due to the difficulties in creating a physiologically relevant model of a non-enzymatic process that occurs over

many years. We have demonstrated however that ageing is associated with changes to the physical organisation of collagen molecules within the fibril. Molecular simulations suggest that the presence of intra-molecular glucosepane increases hydration around the collagen molecules and results in a less tightly held helical structure with increased porosity to water. These findings are reinforced by experimental work showing a corresponding increase in the free water content in tendons from older aged individuals. In addition, the changes detected in the physical properties of collagen molecules can be explained by a more highly cross-linked collagen scaffold, but less confinement of molecules within the collagen fibril. Despite a more open structure to the collagen molecule, susceptibility to degradation is likely to be reduced in the presence of a higher density of crosslinks.

Alongside advances in understanding the chemistry of formation and in developing pharmacological agents to break AGE crosslinks, an understanding of the impact of glucosepane on tissue properties is essential. Our study demonstrates impact at the molecular level, which may at first seem counter-intuitive but can explain age-related changes at a tissue level. The findings of this study are key to understanding age-related changes in healthy individuals and the multi-organ pathologies in diabetes and are essential to developing effective therapies.

Methods

Human tendon collection

The Achilles tendon (AT) and anterior tibialis tendon (ATT) were harvested from limbs amputated in tumour resection operations at the Royal National Orthopaedic Hospital, Stanmore, UK. Only tendons where there were no visual signs of injury or disease and no previous surgeries affecting the tendons were selected. Tendons were collected into gender-balanced young groups (age range 14–21 years, mean ± SD 16.7 ± 2.7 years, n = 6) and old groups (age range 65–81 years, mean ± SD 71.2 ± 6.0 years, n = 6). The ATT was unavailable for one male donor in the young age group due to technical reasons. Tissue was retrieved from the limb within 24 h of limb amputation, snap frozen in hexane on dry ice and stored at –80 °C.

Ethical approval

Tissue collection was carried out initially through the Stanmore Musculoskeletal BioBank (REC reference 09/H0304/78) which subsequently operates as a satellite to UCL/UCLH Biobank for Studying Health and Disease (HTA license number 12055). All patients gave consent for their tissue to be used for musculoskeletal related research. Local R&D approval for this project was given by the UCL/UCLH/RF Joint Research Office (Ref: 11/0464).

Water content

Outer paratendinous and epitenon tissue was trimmed from the tendon tissue blocks whilst semi-frozen and discarded, the remaining tendon tissue was finely diced. Approximately 100 mg wet weight tissue was reserved for DSC analysis (Mettler-Toledo, Leicester, U.K.) and the remaining tissue was weighed accurately (±0.01 mg) on a digital balance (Mettler-Toledo, Leicester, U.K.) prior to lyophilisation using a freeze drier (Alpha1-2 LDplus, Christ, Germany). When a constant weight was reached the tissue was re-weighed and the water content calculated as a percentage of the wet weight.

Glucosepane quantification

A sample of freeze dried tissue (approximately 10 mg) was weighed accurately and subjected to a sequential enzymatic digestion with 500 µl collagenase from *Clostridium histolyticum* (24 mg/ml or 3000 CDU/ml) at 37 °C for 24 h, 450 µl pronase from *Streptomyces griseus* (5 mg/ml or 35 units/ml) at 37 °C for 24 h and 50 µl aminopeptidase from *Aeromonas proteolytica* (4 units/ml) at 37 °C for 16 h, respectively. All enzymes were purchased from Sigma, U.K. and prepared in 0.1 M Tris (hydroxymethyl)methylamino]propanesulfonic acid buffer (TAPS buffer, pH 8.5) containing 0.1 M $CaCl_2$. After each digestion step, the corresponding enzyme was deactivated by immersion into a water bath thermostated at 70 °C for 10 min.

The relative levels of glucosepane were measured using LC-MS on a Bruker Impact-series Q(q)-TOF geometry mass spectrometer (Bruker-Daltonics, Coventry, U.K.), interfaced with a Dionex Ultimate 3000 UHPLC chromatograph (Thermo Scientific, U.K.). Chromatography was performed on a 150 mm × 2.1 mm, 3 µm, Waters Atlantis C18 column (Waters Corporation, Wilmslow, U.K.) using a binary solvent system under gradient conditions and a flow rate of 0.20 ml/min (Table 1). Mobile Phases A and B were 5 mM ammonium formate buffer (pH 5.8) containing 0.1% formic acid in water and methanol, respectively. The mass spectrometer was set up to acquire the range m/z 30 to 650 at 2 Hz, optimised for transmission at m/z 429 in positive ion. Analyses were externally mass calibrated using sodium formate clusters with addition of imidazole. As some interference issues were seen in two samples, quantitation was repeated using accurate-mass MS/MS for added selectivity in addition to the orthogonal

Table 1. Gradient elution program for the UHPLC separation and quantification of glucosepane in tendon samples.

Time (min)	Mobile Phase A (%)	Mobile Phase B (%)
0.0	99	1
1.0	80	20
3.5	80	20
4.0	20	80
6.0	20	80
7.0	80	20
9.0	80	20
10.0	99	1
15.0	99	1

Mobile Phase A: 5 mM Ammonium formate buffer (pH 5.8) in water containing 0.1% formic acid.
Mobile Phase B: 5 mM Ammonium formate buffer (pH 5.8) in methanol containing 0.1% formic acid.

separation offered by the liquid chromatography technique. Collision induced dissociation was used as tandem mass spectrometry technique. All analyses were performed in triplicate. Glucosepane was prepared using the method of Biemel et al. (2002) [6] and standards diluted in Solvent A in the range of 0.01 to 10 µg ml^{-1}. Processing of data was performed in Bruker QuantAnalysis software (Bruker-Daltonics, Coventry, U.K.).

Differential scanning calorimetry (DSC)

Calorimetric measurements were made using a computer controlled DSC 1 (Mettler Toledo, Leicester, U.K.) fitted with an intracooler and purged with nitrogen gas. The manufacturers software (STARe version 13.0) was used to run the program and analyse the data. Samples of tendon tissue (approximately 20 mg) were weighed accurately on a precision balance (0.01 mg resolution) into a 40 µl aluminum crucible and the pan sealed. The DSC was run from 25 °C to –40 °C and – 40 °C to 90 °C at a heating rate of 5 °C per min. Free water content was calculated from the exothermic freezing peak and the collagen denaturation peak used to determine the characteristics of collagen denaturation. Following the first run, the pan was removed from the DSC and perforated using a sharp pin. A second run was performed from 25 °C to 160 °C in the perforated pan to determine the total water content of the sample using the endothermic peak for water evaporation. Following DSC measurements the pans were opened and the tissue removed and weighed and placed in a small glass hydrolysis tube. Samples were hydrolysed in 1 ml of 6 M HCl at 110 °C for 24 h, dried under vacuum and re-dissolved in H_2O prior to hydroxyproline assay as previously described [32] to determine collagen content.

Statistical analysis

Statistical significance was evaluated using a General Linear Model in SPSS Statistics Software, version 24, (IBM, New York, USA). Age group and tendon type were used as fixed factors and interaction between factors were noted. The level of significance was set at $p \leq 0.05$. Results are given as mean ± standard deviation.

Electronic structure calculations

Electronic structure optimisation and frequency calculations were performed using Gaussian G09 [33]. The Hartree-Fock (HF) method was employed for the derivation of force constants of glucosepane and a basis set of 6-31G(p) was used. The optimisation threshold was set to "tight" and checked for imaginary frequencies using frequency analysis calculations. *Ab initio* calculations were performed in a vacuum.

Restrained electrostatic partial charge derivation

The optimised structures were submitted to the pyRED server [34]. Gaussian geometry optimisation was switched off and the capped groups were restrained for RESP charge fitting to the central fragment only, resulting in a net charge of zero. Partial atomic charges compatible with the Amberff99SB force field were assigned to the glucosepane structure. The bonded force constants and equilibrium values were derived using ForceGen [35]. Validation of the force constants and the link to the topology files can be found in the Supplementary information (note 3) documentation.

Molecular dynamics simulation details

All MD simulations were carried out using GROMACS 5.1.2 [22], utilising a time step of 0.002 ps. Short-range Van der Waals interactions were defined using the LJ-6-12 expression with a 1.2 nm cut-off and electrostatic interactions were defined using the Particle Mesh Ewald summation method [36] with a 1.2 nm cut-off. All bonded interactions were holonomically constrained using the LINCS constraint algorithm [37]. The Verlet cut-off scheme was employed to keep energy drift constant and the minimum update frequency of the nearest neighbour search was set to 5 steps. The temperature and pressure were held at 310 K and 1 atm, respectively, using the steps outlined below. The unit cell was accompanied by periodic boundary conditions in all three dimensions. Energies, velocities and coordinates were collected every 2 ps.

Preparation of model collagen system

Three models were created for study; a wild-type, which is an extension to the unit-cell of the 4Z1R crystal structure of a collagen-like peptide [38], a model with a lysine and arginine side chain substituted between two polypeptide chains at residue positions 17 and 49 respectively, and a glucosepane cross-link between two polypeptide chains of the same collagen molecule at positions 17 and 49. Only one collagen molecule from the eight available was modified (with the exception to the wild-type) in the final extended unit cell.

The 4Z1R crystal structure was downloaded as an asymmetric unit cell. The water molecules were removed and any missing side-chains were recovered using the Protein Preparation Wizard in Schrodinger Maestro. The crystal structure unit cell was reproduced by replicating the asymmetric unit cell on the y-axis, yielding to sets of two neighbouring molecules. To prevent the short-range interaction distance from being longer than half the unit cell dimension, the four molecules and the crystallographic dimensions were replicated normal to the long axis, yielding two repeats of four collagen molecules. All adjustments were made using the CCP4 software package [39].

The atom types were encoded using the Amber99FFSB force field [40], incorporating the bonded and angular force constants, equilibrium constants and LJ-potential terms for the glucosepane molecule (see Supplementary information, note 3, for parameter set). Eight times the quantity of water from the asymmetric unit cell was added using the TIP3P water molecule and the system was neutralised with counter ions. The completed model with accompanying extended unit cell dimensions can be seen in Fig. 2a.

System equilibration

The three models were equilibrated using the following steps. Firstly, to resolve steric clashes between solute and solvent and the discrepancy in bond lengths and bond angles from manually inserting glucosepane and the two charged side chains, each model was subjected to a steepest descent energy minimisation with a force threshold of 1000 kJ mol^{-1} nm^{-1}. Further minimisation was performed using the conjugate gradient energy descent with a force threshold of 100 kJ mol^{-1} nm^{-1}. The cut-off scheme, electrostatics and Van der Waals properties were as described. Secondly, each model adopted the NVT ensemble (fixed number of particles, volume and temperature) by setting a velocity scaling thermostat to 310 K, and the first simulation was performed for 100 ps. Coordinates and velocities were preserved and each model equilibrated for a further 100 ps whilst coupling an isotropic Berendsen pressure scheme. Compressibility was set to $4.5e^{-5}$ in all three spatial dimensions. Finally, the pressure coupling was replaced with the Parrinello-Rahman pressure coupling for a truer representation of an NPT ensemble (fixed number of particles, pressure and temperature) and 10 ns equilibrium simulation was performed.

Acknowledgements

The authors are grateful to Advanced Research Computing at Cardiff University, for the use of Raven and Dr. George McLeod From Bruker U.K. Ltd. for assistance with mass spectrometry techniques.

Funding sources

This work was supported by the Biotechnology and Biological Sciences Research Council (BBSRC), UK (Grant no: BB/K007785).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mbplus.2019.100013.

Received 1 April 2019;
Received in revised form 22 August 2019;
Accepted 23 August 2019
Available online 31 August 2019

Keywords:
Collagen;
Ageing;
Glycation;
Glucosepane;
Hydration;
Collagen fibril

†University of Oxford, Nuffield Department of Clinical Neurosciences, Oxford, OX1 3QX, UK.
‡Pharmaron, Hertford Road, Hoddesdon, Hertfordshire, EN11 9BU, UK.
§Babraham Institute, Babraham Research Campus, Cambridge, CB22 3AT, UK.
||Faculty of Dentistry, University of Toronto, Toronto, Canada, ON M5G 1G6.
¶Cardiff University, School of Chemistry, Cardiff, CF10 3AT, UK.

Abbreviations used:
AGE, advanced glycation end-product; AT, Achilles tendon; ATT, anterior tibialis tendon; DSC, differential scanning calorimetry; LC, liquid chromatography; MD, molecular dynamics; SASA, solvent-accessible surface area.

References

[1] D.R. Sell, K.M. Biemel, O. Reihl, M.O. Lederer, C.M. Strauch, V.M. Monnier, Glucosepane is a major protein cross-link of the senescent human extracellular matrix. Relationship with diabetes, J. Biol. Chem. 280 (13) (2005), 12310-5.

[2] V.M. Monnier, S. Genuth, D.R. Sell, The pecking order of skin Advanced Glycation Endproducts (AGEs) as long-term markers of glycemic damage and risk factors for micro- and subclinical macrovascular disease progression in Type 1 diabetes, Glycoconj. J. 33 (4) (2016) 569–579.

[3] C. Legrand, U. Ahmed, A. Anwar, K. Rajpoot, S. Pasha, C. Lambert, R.K. Davidson, I.M. Clark, P.J. Thornalley, Y. Henrotin, N. Rabbani, Glycation marker glucosepane increases with the progression of osteoarthritis and correlates with morphological and functional changes of cartilage in vivo, Arthritis Res. Ther. 20 (1) (2018) 131.

[4] J.S. Sjoberg, S. Bulterijs, Characteristics, formation, and pathophysiology of glucosepane: a major protein cross-link, Rejuvenation Res. 12 (2) (2009) 137–148.

[5] V.M. Monnier, A. Cerami, Nonenzymatic browning in vivo: possible process for aging of long-lived proteins, Science 211 (4481) (1981) 491–493.

[6] K.M. Biemel, D.A. Friedl, M.O. Lederer, 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. 277 (28) (2002), 24907-15.

[7] K.M. Biemel, O. Reihl, J. Conrad, M.O. Lederer, 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. 276 (26) (2001) 23405–23412.

[8] C. Draghici, T. Wang, D.A. Spiegel, Concise total synthesis of glucosepane, Science 350 (6258) (2015) 294–298.

[9] V.M. Monnier, D.R. Sell, C. Strauch, W. Sun, J.M. Lachin, P. A. Cleary, S. Genuth, The association between skin collagen glucosepane and past progression of microvascular and neuropathic complications in type 1 diabetes, J. Diabetes Complicat. 27 (2) (2013) 141–149.

[10] S. Genuth, W. Sun, P. Cleary, X. Gao, D.R. Sell, J. Lachin, V. M. Monnier, Skin advanced glycation end products glucosepane and methylglyoxal hydroimidazolone are independently associated with long-term microvascular complication progression of type 1 diabetes, Diabetes 64 (1) (2015) 266–278.

[11] C.T. Thorpe, I. Streeter, G.L. Pinchbeck, A.E. Goodship, P.D. Clegg, H.L. Birch, Aspartic acid racemization and collagen degradation markers reveal an accumulation of damage in tendon collagen that is enhanced with aging, J. Biol. Chem. 285 (21) (2010) 15674–15681.

[12] H.L. Birch, Tendon matrix composition and turnover in relation to functional requirements, Int. J. Exp. Pathol. 88 (4) (2007) 241–248.

[13] C.T. Thorpe, C. Klemt, G.P. Riley, H.L. Birch, P.D. Clegg, H. R. Screen, Helical sub-structures in energy-storing tendons provide a possible mechanism for efficient energy storage and return, Acta Biomater. 9 (8) (2013) 7948–7956.

[14] C.T. Thorpe, H.R. Screen, Tendon structure and composition, Adv. Exp. Med. Biol. 920 (2016) 3–10.

[15] K.M. Heinemeier, P. Schjerling, J. Heinemeier, S.P. Magnusson, M. Kjaer, Lack of tissue renewal in human adult Achilles tendon is revealed by nuclear bomb (14)C, FASEB J. 27 (5) (2013) 2074–2079.

[16] R.A. Clayton, C.M. Court-Brown, The epidemiology of musculoskeletal tendinous and ligamentous injuries, Injury 39 (12) (2008) 1338–1344.

[17] G.W. Hess, Achilles tendon rupture: a review of etiology, population, anatomy, risk factors, and injury prevention, Foot Ankle Spec 3 (1) (2010) 29–32.

[18] S.M. Raikin, D.N. Garras, P.V. Krapchev, Achilles tendon injuries in a United States population, Foot Ankle Int. 34 (4) (2013) 475–480.

[19] C.S. Avella, E.R. Ely, K.L. Verheyen, J.S. Price, J.L. Wood, R.K. Smith, Ultrasonographic assessment of the superficial digital flexor tendons of National Hunt racehorses in training over two racing seasons, Equine Vet. J. 41 (5) (2009) 449–454.

[20] E.R. Ely, C.S. Avella, J.S. Price, R.K. Smith, J.L. Wood, K.L. Verheyen, Descriptive epidemiology of fracture, tendon and suspensory ligament injuries in National Hunt racehorses in training, Equine Vet. J. 41 (4) (2009) 372–378.

[21] R.B. Williams, L.S. Harkins, C.J. Hammond, J.L. Wood, Racehorse injuries, clinical problems and fatalities recorded on British racecourses from flat racing and National Hunt racing during 1996, 1997 and 1998, Equine Vet. J. 33 (5) (2001) 478–486.

[22] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindahl, GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX 1-2 (2015) 19–25.

[23] F. Eisenhaber, P. Lijnzaad, P. Argos, C. Sander, M. Scharf, The double cubic lattice method: efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies, J. Comput. Chem. 16 (3) (1995) 273–284.

[24] G.N. Ramachandran, R. Chandrasekharan, Interchain hydrogen bonds via bound water molecules in the collagen triple helix, Biopolymers 6 (11) (1968) 1649–1658.

[25] N. Aktas, Y. Tulek, H.Y. Gokalp, Determination of differences in free and bound water contents of beef muscle by DSC under various freezing combinations, J. Therm. Anal. 50 (4) (1997) 617–624.

[26] G.D. Fullerton, E. Nes, M. Amurao, A. Rahal, L. Krasnosselskaia, I. Cameron, An NMR method to characterize multiple water compartments on mammalian collagen, Cell Biol. Int. 30 (1) (2006) 66–73.

[27] A. Nash, J. Sassmannshausen, L. Bozec, H.L. Birch, N.H. de Leeuw, Computational study of glucosepane-water and hydrogen bond formation: an electron topology and orbital analysis, J. Biomol. Struct. Dyn. (2016) 1–11.

[28] G.N. Ramachandran, G. Kartha, Structure of collagen, Nature 176 (4482) (1955) 593–595.

[29] D.J. Horgan, N.L. King, L.B. Kurth, R. Kuypers, Collagen crosslinks and their relationship to the thermal properties of calf tendons, Arch. Biochem. Biophys. 281 (1) (1990) 21–26.

[30] C.A. Miles, M. Ghelashvili, Polymer-in-a-box mechanism for the thermal stabilization of collagen molecules in fibers, Biophys. J. 76 (6) (1999) 3243–3252.

[31] C.A. Miles, T.V. Burjanadze, A.J. Bailey, The kinetics of the thermal denaturation of collagen in unrestrained rat tail tendon determined by differential scanning calorimetry, J. Mol. Biol. 245 (4) (1995) 437–446.

[32] H.L. Birch, A.J. Bailey, A.E. Goodship, Macroscopic 'degeneration' of equine superficial digital flexor tendon is accompanied by a change in extracellular matrix composition, Equine Vet. J. 30 (6) (1998) 534–539.

[33] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M. A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery, J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, Gaussian 09, Revision E.01, Wallingford CT, 2009.

[34] E. Vanquelef, S. Simon, G. Marquant, E. Garcia, G. Klimerak, J. C. Delepine, P. Cieplak, F.Y. Dupradeau, R.E.D. Server: a web service for deriving RESP and ESP charges and building force field libraries for new molecules and molecular fragments, Nucleic Acids Res. 39 (Web Server issue) (2011), W511-7.

[35] A. Nash, T. Collier, H.L. Birch, N.H. de Leeuw, ForceGen: atomic covalent bond value derivation for Gromacs, J. Mol. Model. 24 (1) (2017) 5.

[36] T. Darden, D. York, L. Pedersen, Particle mesh Ewald: an N·log(N) method for Ewald sums in large systems, J. Chem. Phys. 98 (12) (1993) 10089–10092.

[37] B. Hess, H. Bekker, H.J.C. Berendsen, J. Fraaije, LINCS: a linear constraint solver for molecular simulations, J. Comput. Chem. 18 (12) (1997) 1463–1472.

[38] M.E. Plonska-Brzezinska, D.M. Bobrowska, A. Sharma, P. Rodziewicz, M. Tomczyk, J. Czyrko, K. Brzezinski, Triple helical collagen-like peptide interactions with selected polyphenolic compounds, RSC Adv. 5 (116) (2015) 95443–95453.

[39] M.D. Winn, C.C. Ballard, K.D. Cowtan, E.J. Dodson, P. Emsley, P.R. Evans, R.M. Keegan, E.B. Krissinel, A.G. Leslie, A. McCoy, S.J. McNicholas, G.N. Murshudov, N.S. Pannu, E.A. Potterton, H.R. Powell, R.J. Read, A. Vagin, K.S. Wilson, Overview of the CCP4 suite and current developments, Acta Crystallogr. D Biol. Crystallogr. 67 (Pt 4) (2011) 235–242.

[40] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, Comparison of multiple Amber force fields and development of improved protein backbone parameters, Proteins 65 (3) (2006) 712–725.