Electronic Supplementary Material (ESI) for ChemComm.

Electronic Supplementary Information (ESI)

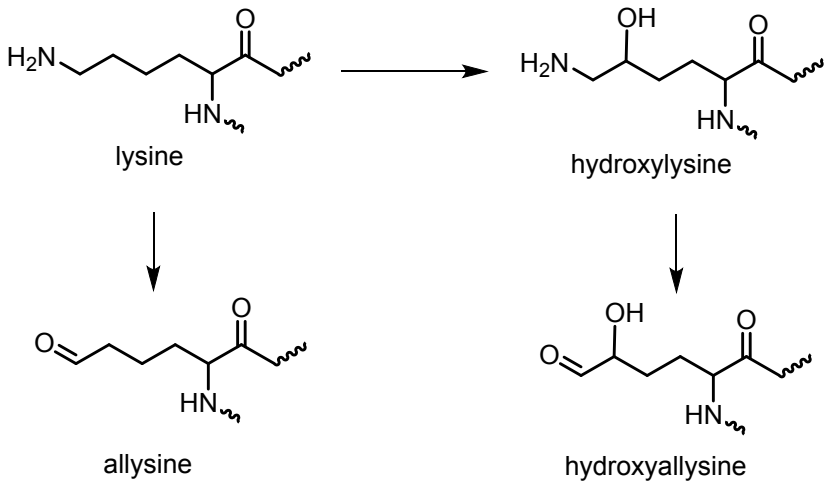
Essential but sparse collagen hydroxylysyl post-translational modifications detected by DNP NMR

Wing Ying Chow,[a] Rui Li,[b] Ieva Goldberga,[b] David G. Reid,[b] Rakesh Rajan,[b] Jonathan Clark,[c] Hartmut Oschkinat,*[a] Melinda J. Duer,*[b] Robert Hayward,[d] Catherine M. Shanahan[d]

a. WYC, HO - Leibniz-Forschungsinstitut für Molekulare Pharmakologie (FMP), Campus Buch, Robert-Roessle Str. 10, Berlin 13125, Germany
b. RL, IG, DGR, RR, MJD* - Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK
c. JC, Babraham Institute, Babraham Research Campus, Cambridge CB22 3AT, UK
d. RH, CMS - BHF Centre of Research Excellence, Cardiovascular Division, King's College London, London SE5 9NU, UK
Authors for correspondence: HO E-mail: oschkinat@fmp-berlin.de; Tel: +49 30 94793 160
MJD mjd13@cam.ac.uk; Tel: +44 1223 736394

Scheme S1 Post-translational initial modifications of Hyl and Lys leading to formation of crosslinks

The first category of vital Hyl modifications is oxidative deamination of Hyl, and Lys, to the aldehydes allysine and hydroxyallysine respectively, following a sequence- and tissue- specific pattern. These oxidations are followed by apparently spontaneous crosslinking involving either Lys or Hyl, and either of the aldehydes,[1,2] believed to result initially in immature crosslinks between two residues, summarized in ESI Scheme S2. Further lysyl residues can then participate in forming mature crosslinks between three residues, containing pyridinium and hydroxypyridinium heterocyclics.[3,4] Hydroxyallysine overabundance makes fibrotic skin resistant to degradation.[5]

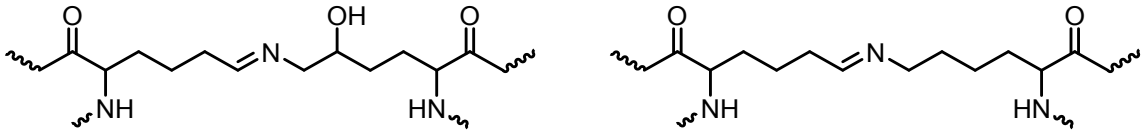


Scheme S2 Further post-translational modifications of Lys and Hyl

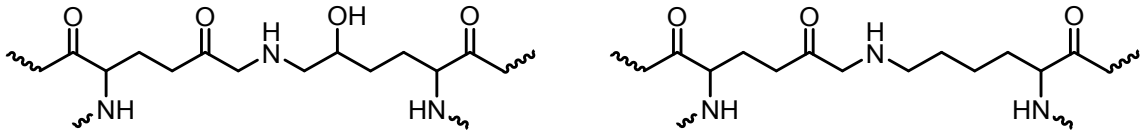
Proposed structures of enzymatically initiated collagen crosslinks.

1) Immature crosslinks:

Aldimines: dehydro-hydroxylysino-norleucine (deH-HLNL) and dehydro-lysino-norleucine (deH-LNL)

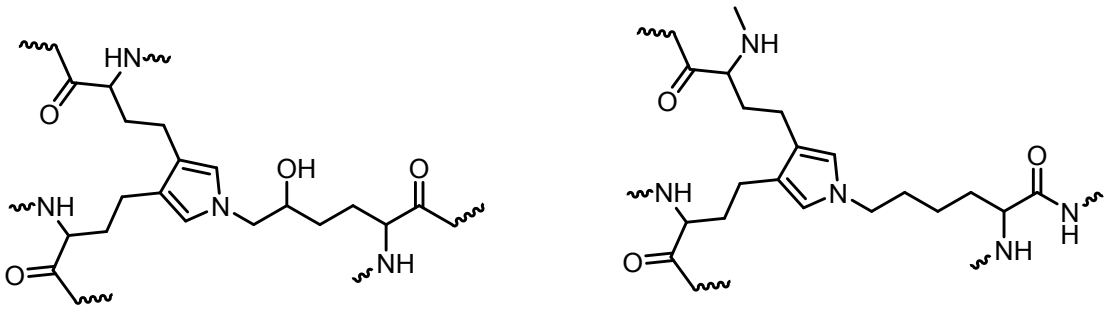


Ketoamines: hydroxylysino-keto-norleucine (HLKNL) and lysino-keto-norleucine (LKNL)

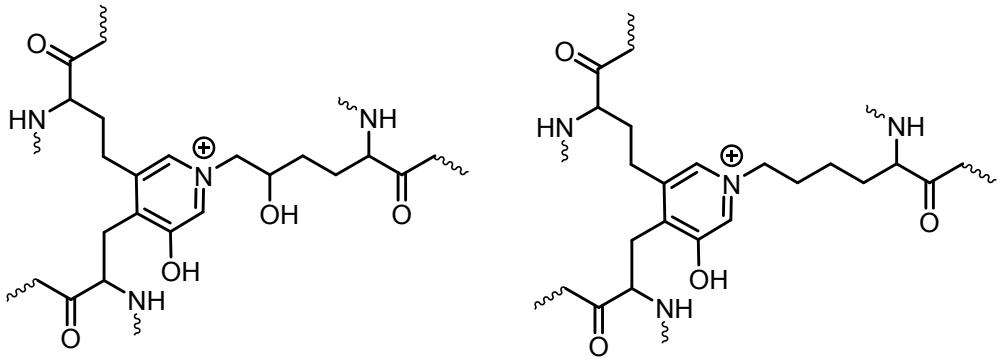


2) Mature crosslinks:

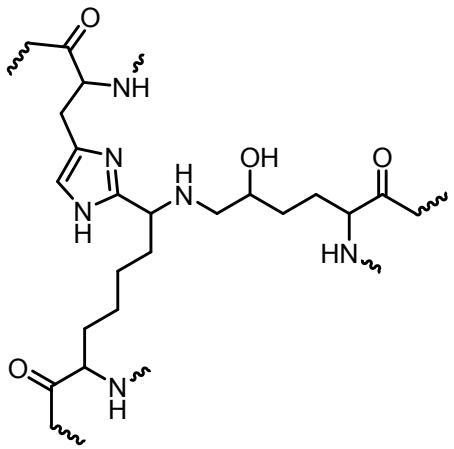
Pyrroles – hydroxylysyl pyrrole (PYL) and lysyl pyrrole (DPL)



Pyridinolines – hydroxylysyl pyridinoline (PYD) and lysyl pyridinoline (DPD)



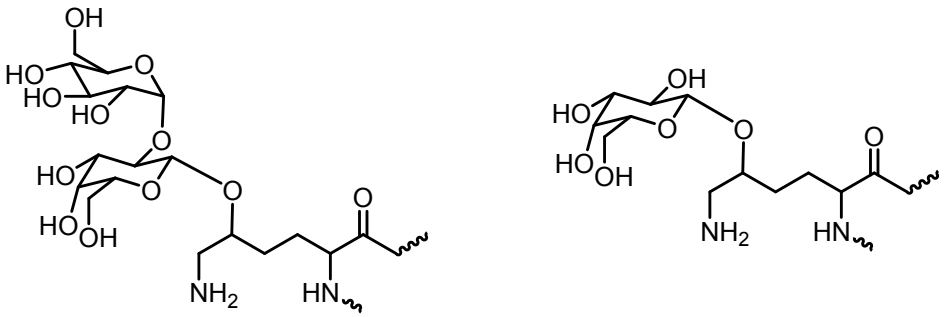
Histidinyl hydroxylysino-norleucine (HHL)



Scheme S3 Enzymatic glycosylation products.

The second category of specific Hyl modification is tissue- and sequence-specific enzymatic O-glycosylation at the C_δ-OH group to form (5R)-5-O-(β-D-galactopyranosyl)-O-Hyl (GalHyl), and α-D-glucopyranosyl-(1→2)-β-D-galactopyranosyl-O-Hyl (GlcGalHyl), comprising at most ca. 0.5% of ECM collagen. The structural consequences of the attachment of bulky hydrophilic sugar groups to collagen have not yet been clarified, but the importance of collagen glycosylation is underlined by a number of connective tissue disorders arising from mutational loss-of-function of the relevant glycosylating enzymes.[6,7]

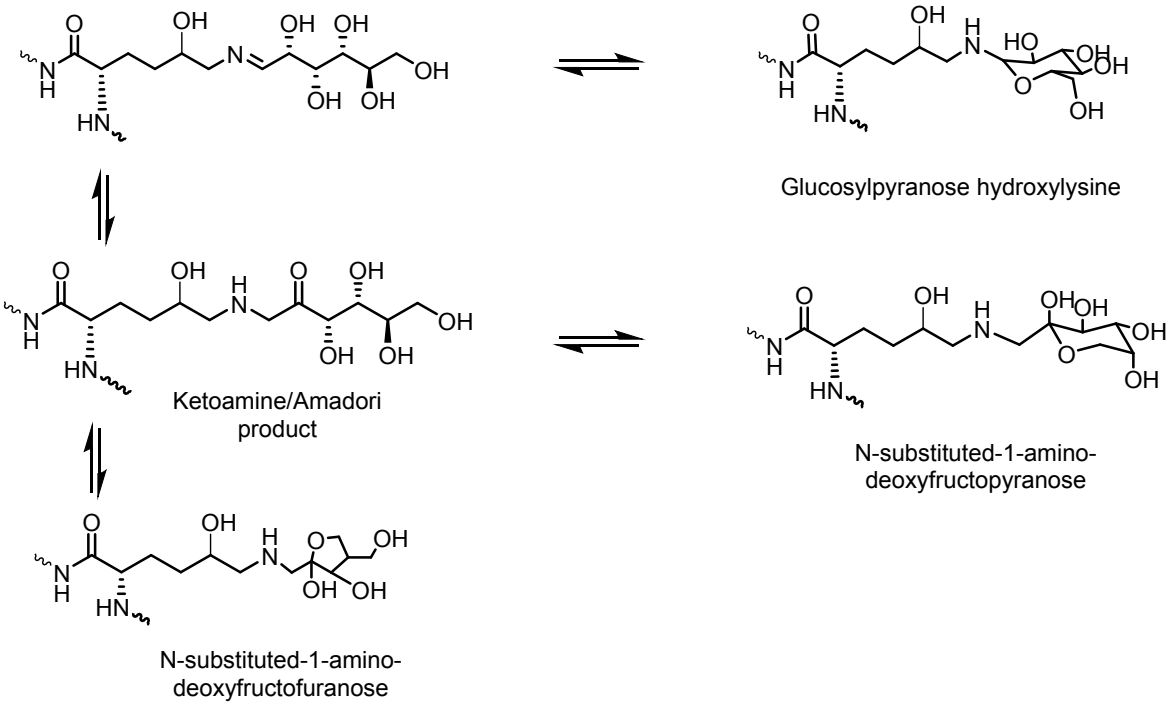
α-D-glucosyl-β-D-galactosyl-5R-O-hydroxylysine and β-D-galactosyl-5R-O-hydroxylysine



Scheme S4 Non-enzymatic early glycation products.

A third category of collagen modification is non-enzymatic glycation, the spontaneous reaction of metabolic sugars with collagen, commonly at the 6-amino group (N_ζ) of Lys and Hyl. This widespread and highly prevalent process is usually deleterious[8,9] and has been linked for instance to bone fragility in ageing.[10] A few chemically well-defined adducts, such as Lys or Hyl coupled via N_ζ to the 1-carbon of deoxyfructose, are generated in early-stage glycation by glucose; these subsequently react into a large heterogeneous group of advanced glycation end products (AGEs). AGE generation underlies many of the tissue pathologies of normal ageing, and hyperglycemic conditions such as long-term diabetes.[3]

The initial glycation reactions between hydroxylysine and glucose:



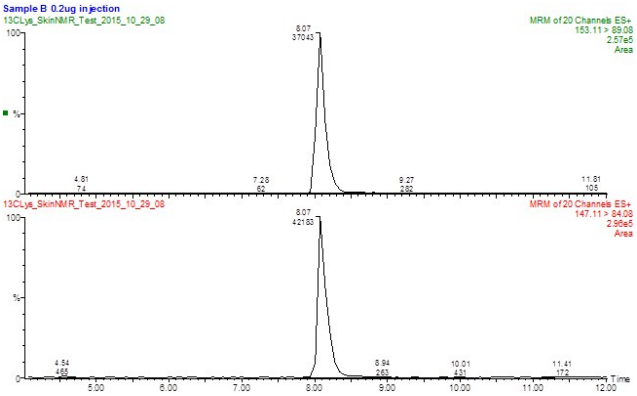
Materials and Methods

Skin labelling with (U-$^{13}C_6$)-Lysine

Two 8-week-old male C57/Bl6 mice were fed MouseExpress L-Lysine ($^{13}C_6$, 99%) irradiated mouse feed (CK Isotopes Ltd) for 56 days. Each mouse consumed approximately 4.5 g feed per day on average over this time period. After day 56 the animals we culled using a schedule 1 method. The fur was plucked from an area of skin of about 1 cm^2 from behind the shoulders which was then cut out for processing. The sample was washed in 3% triton for 5 days, water for 2 days, 0.2% SDS for 5 days and then water for 2 days. Samples were then stored frozen in PBS.

The ^{13}C-lysine content was measured by HPLC mass spectrometry after hydrolysis with 6M hydrochloric acid under standard conditions for amino acid analysis. The HPLC methodology was as described by Bawazeer[11] and Pesek[12].

In brief, the samples were made up in (acetonitrile, 30%water, 0.1% formic acid) and 5ul injected (0.2ug tissue equivalent) onto a Cogent Diamond Hydride column (4um, 100A, 150 x 2.1mm). A gradient of 100% (Acetonitrile, 5% water, 0.1% formic acid, 0.005% trifluoroacetic acid) to 100% (Water, 0.1% formic acid) over 20 minutes was run at a flow rate of 0.1ml/min and passed into an esi probe of a Quattro Ultima mass spectrometer and the fragmentations 147.11>84.08 for $^{12}C_6$-Lys and 153.11>89.08 for $^{13}C_6$-Lys monitored. (Source temperature 120^{o}C, desolvation temperature 350^{o}C, cone voltage 3kV, Capillary voltage 35V, collision gas was argon, collision voltage 20V). The areas of the MS peaks for $^{12}C_6$-Lys (from incorporation prior to beginning the isotope feeding, top trace at M_r = 147.11 Area = 42183, below) and $^{13}C_6$-Lys (bottom trace at M_r = 153.11 Area = 37043, below) were used to calculate the percentage $^{13}C_6$-Lys of the total lysine content ($^{12}C_6$-Lys plus $^{13}C_6$-Lys areas) in the tissue samples.



The $^{13}C_6$-lysine incorporation into the mouse skin samples was found to be ca. 45% of the total lysine content in the skin. The other 55% of lysine content is unlabelled as it was already incorporated into the tissue prior to the start of the feeding of isotopically-enriched diet at 8 weeks.

ssNMR rotor packing

No grinding or cryomilling procedures were used. For the DNP NMR experiments, the solid samples were freeze-dried, then cut into small pieces if necessary prior to rotor filling. For all NMR experiments, 3.2 mm outer diameter zirconia rotors with Vespel caps were used. For DNP NMR experiments, AMUPol[13] was prepared as 16-20 mM solutions in D_2O/H_2O (75% D_2O, no glycerol was used), then 12.5-15 µL of the radical solution was added to the sample-filled

rotor. The rotors were closed with a Vespel cap, then spun in both vertical orientations (with the cap facing upwards and downwards) with a desktop mini-centrifuge to distribute the radical solution within the rotor. The packed rotors were allowed to equilibrate with the radical solution overnight in the fridge before the start of DNP NMR experiments.

Specific conditions for each sample:

1. Mouse skin labelled in lysine (no glycerol)

Sample mass: 16.5 mg

AMUPol solution preparation: 0.166 mg of solid AMUPol was added to 14 µL of water (75% D_2O, 25% H_2O) to yield a 16.3 mM AMUPol solution. All of this solution was added to the solid sample in the rotor and incubated in the fridge (4°C) overnight prior to DNP NMR experiments.

2. Mouse skin labelled in lysine (with glycerol)

Sample mass: 16.3 mg

AMUPol solution preparation: 0.170 mg of solid AMUPol was added to 14 µL of the standard DNP NMR solvent mix (volume ratio of 6:3:1 glycerol:D_2O:H_2O; specifically 8.4 µL deuterated glycerol, 4.2 µL D_2O, 1.4 µL H_2O) to yield a 16.7 mM AMUPol solution. All of this solution was added to the solid sample in the rotor and incubated in the fridge (4°C) overnight prior to DNP NMR experiments.

3. VSMC

Sample mass: 10.0 mg

AMUPol solution preparation: 0.445 mg of solid AMUPol was added to 30 µL of water (75% D_2O, 25% H_2O) to yield a 20.4 mM AMUPol solution. 12.5 µL of this solution was added to the solid sample in the rotor and incubated in the fridge (4°C) overnight prior to DNP NMR experiments.

DNP NMR spectroscopy

NMR experiments were performed on Bruker Avance III wide-bore spectrometers operating at a static magnetic field of 9.4 T, with ^{1}H Larmor frequencies of 400 MHz. The 400 MHz spectrometer is equipped for DNP NMR, with a gyrotron operating at 9.7 T and a 3.2 mm (^{1}H-^{13}C-^{15}N) triple resonance low-temperature magic angle spinning (LTMAS) probe, with the MAS rate set to 8889 Hz unless otherwise specified. The sample temperature was set to 100 K using the LTMAS cooling cabinet with a variable temperature gas flow of 2000 litres per hour, which has been previously shown to yield an actual experimental temperature of 110 K. Auxiliary temperature readouts suggest a deviation of within 5 K under MAS and microwave irradiation conditions. Experimental conditions were based on those previously optimized in solid state DNP NMR experiments on biomolecular systems.[14]

NMR parameters on the 400 MHz spectrometer were: ^{1}H 90° pulse length 2.2-2.5 µs, ^{13}C 90° pulse length 4.1-4.75 µs, 80+ kHz ^{1}H decoupling. DNP-enhanced experiments were carried out under 30 mA microwave irradiation.

^{1}H-^{13}C cross polarisation (CP) experiments are widely used in ssNMR. This technique is based on dipolar coupling and therefore specifically detects immobile and solid species. 1500 µs ^{1}H-^{13}C ramped CP[15] contact time was used on the 400 MHz spectrometer.

^{13}C-^{13}C 2D homonuclear correlation spectra were carried out with a dipolar-assisted rotary resonance (DARR) mixing sequence [16]. Mixing times of 5 and 20 ms were recorded for both samples. We note under the low-temperature conditions of DNP NMR, mixing times cannot be directly compared to the same mixing time at room temperature. Each DARR experiment was signal averaged for 7-15 hours.

^{13}C-^{13}C 2D single quantum-double quantum (SQ-DQ) correlation spectra were carried out on the mouse skin by CP followed by a SPC5[17] mixing sequence. The MAS rate was set to 8300 Hz for this experiment to avoid overlap of spinning sidebands with bona fide signals. DQ coherence was excited by applying the SPC5(2) sequence for 1.2 ms (10 times). The SQ-DQ experiment was signal averaged for 6 hours.

^{15}N-(^{13}C)-^{13}C pseudo-2D heteronuclear correlation (effectively ^{15}N-^{13}C correlation) experiments were carried out with 90° pulse lengths of: ^{1}H 2.4 µs, ^{13}C 4.33 µs, ^{15}N 4.27 µs. Magnetization was first excited by ^{1}H-^{15}N CP with a 450 µs contact pulse length, with a slight ramp shape on the ^{15}N channel, followed by indirect ^{15}N evolution, during which ^{13}C magnetisation was refocussed by a 90°-180°-90° sequence and ^{1}H was decoupled at 80 kHz. Magnetization was then transferred by ^{15}N-^{13}C SPECIFIC-CP[18 19] with a 4.5 ms contact pulse length, with a square shape on ^{15}N and a tangential shape on ^{13}C. At this point there was no indirect evolution on ^{13}C, instead, a DARR sequence[16] (mixing time of 10 ms) was immediately applied on the ^{13}C channel, followed by ^{13}C acquisition with 80 kHz ^{1}H decoupling. A 219 ppm ^{15}N sweep width was used here to ensure the full ^{15}N range was recorded without aliasing. 60 indirect evolution increments were collected, with 1024 transients per increment. Thus, the signal averaging time was ~51 hours.

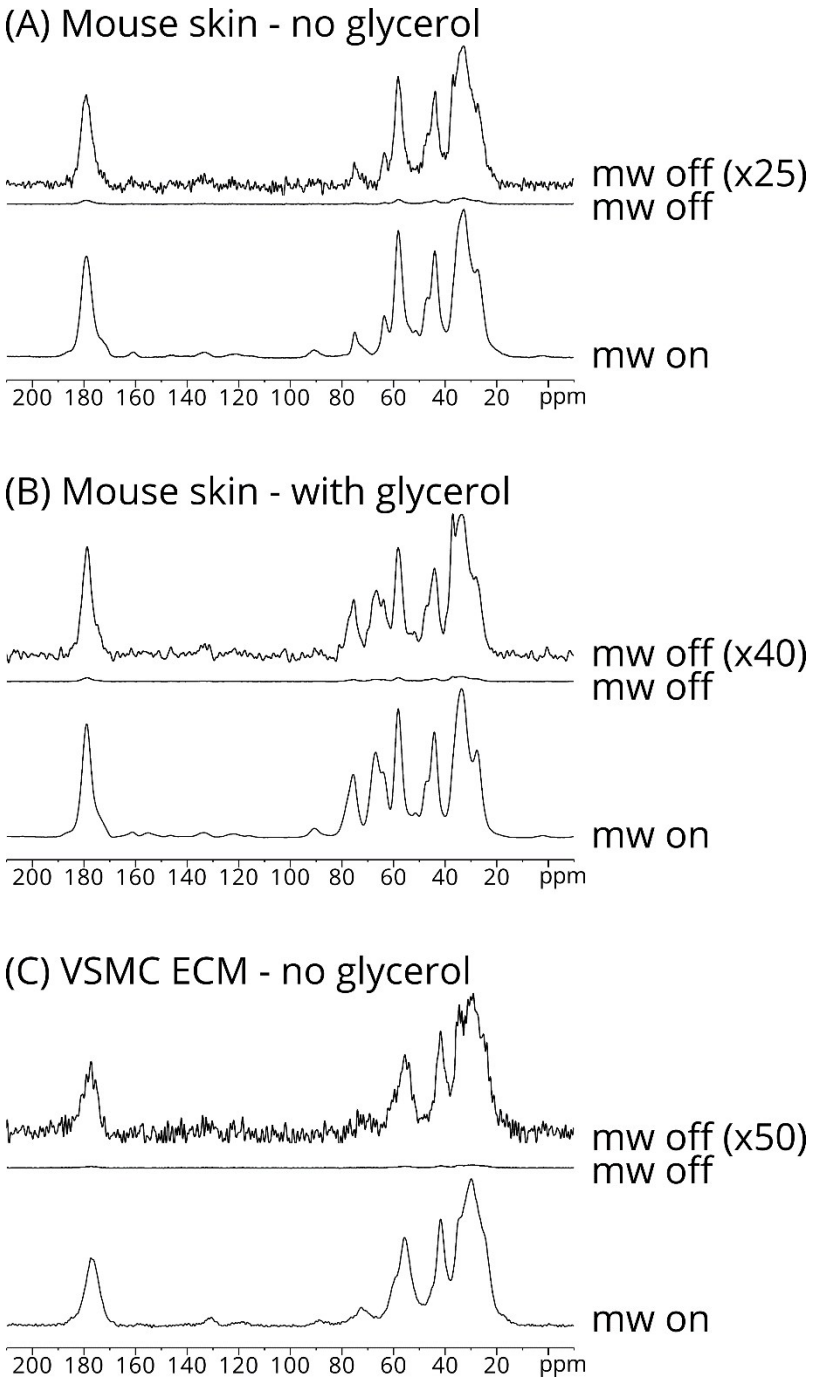


Fig. S1 1D ^{13}C MAS DNP-enhanced CP spectra on AMUPol-doped samples at 110K on a 400 MHz spectrometer, with and without microwave irradiation. (A) Mouse skin with ^{13}C-Lys enrichment, no glycerol. ε = 26.5 ± 1.5. (B) Mouse skin with ^{13}C-Lys enrichment, with glycerol. ε = 39 ± 5. (C) VSMC ECM with ^{13}C,^{15}N-Lys enrichment. ε = 57 ± 4. From (A) and (B) we note that glycerol does not make a large difference in enhancement of this type of samples. From all sets of microwave on versus microwave off spectra, we note that the DNP enhancement is generally fairly uniform on this type of samples.

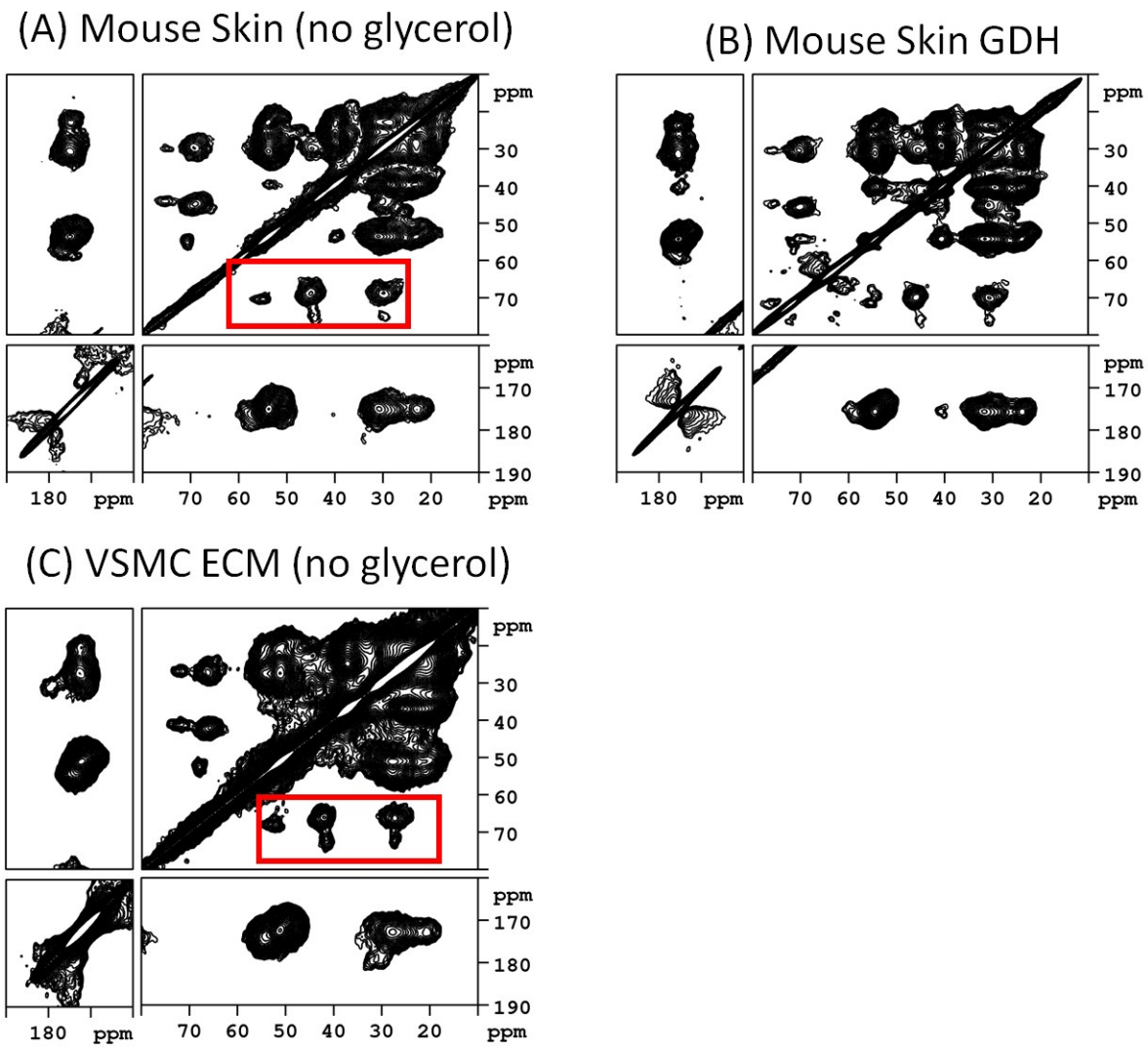


Fig. S2 2D DNP-enhanced DARR ^{13}C - ^{13}C correlations on AMUPol-doped samples at 110K on a 400 MHz spectrometer with mixing time of 5ms. (A) Mouse skin sample with ^{13}C-Lys enrichment (no glycerol), acquisition time 16 h. (B) Mouse skin sample with ^{13}C-Lys enrichment (with glycerol), acquisition time 13 h. (C) VSMC ECM with ^{13},^{15}N-Lys enrichment, no glycerol, acquisition time 14 h. Cross peaks correlating with the on-diagonal Hyl C_δ cross peaks discussed in Fig. 1 are boxed by red rectangles.

With regards to the collagen content of the two samples: the mouse skin sample is generated in vivo and is almost entirely type I collagen, while the VSMC ECM contains considerable amounts of other non-collageneous ECM proteins. To the best of current knowledge, only collagen proteins in the ECM are known to undergo Lys to Hyl conversion and subsequent modifications. Therefore, the two spectra are highly similar despite the differences in protein composition.

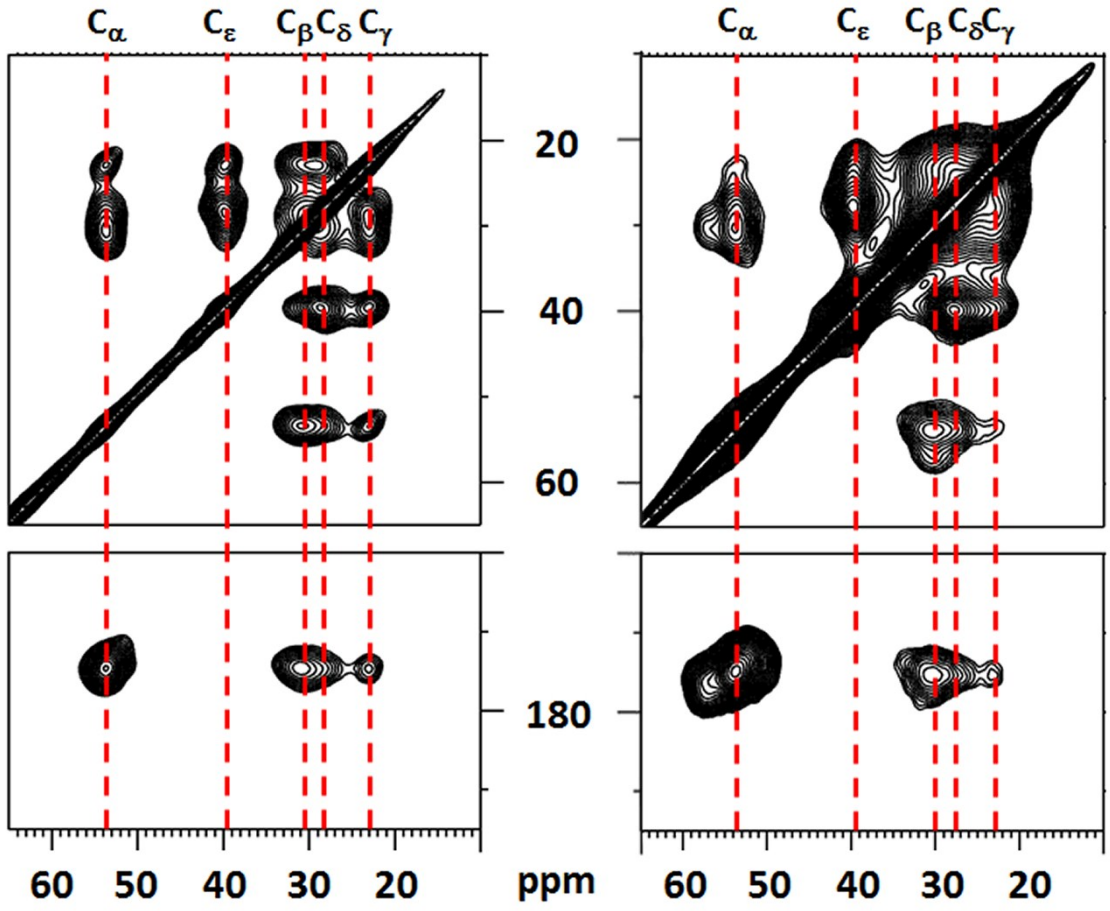


Fig. S3 2D DNP-enhanced DARR ^{13}C - ^{13}C correlations (mixing time 5 ms) of Lys-labelled skin (left) and VSMC ECM (right) showing the expected strong cross peaks arising from the labelled Lys spin systems in collagen (and other proteins). Both spectra were obtained at 110K on a 400 MHz NMR spectrometer.

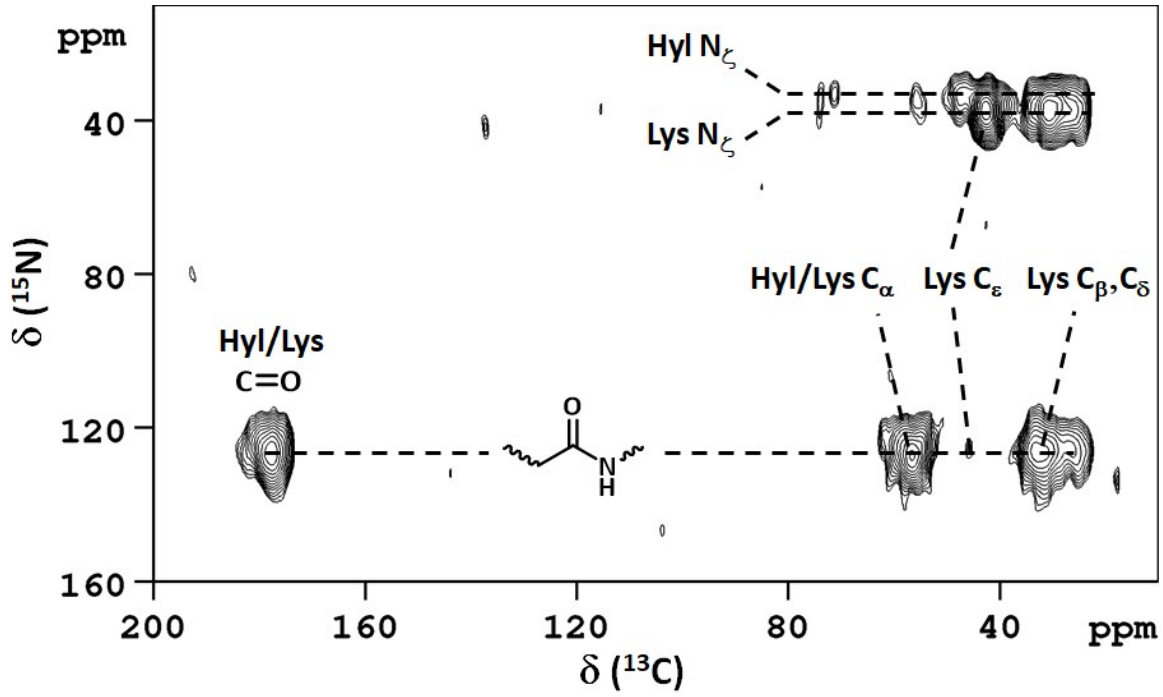
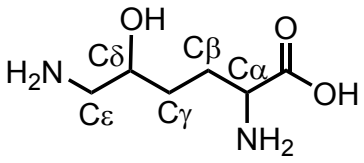


Fig. S4 2D DNP-enhanced ^{15}N - ^{13}C DARR-assisted correlation spectrum of (U-$^{13}C_6$, $^{15}N_2$)-Lys labelled VSMC ECM showing all observed correlations. This spectrum was obtained at 110K on a 400 MHz spectrometer.

Database information supporting assignments of Hyl and adducts

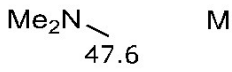
Aqueous solution state ^{13}C NMR shifts of **Hyl** from Human Metabolomics Database (HMDB http://www.hmdb.ca/spectra/nmr_two_d/1319#_blank)

C_α 57.3 ppm, C_β 29.4 ppm, C_γ 32.2 ppm, **C_δ 70.0 ppm**, C_ε 57.3 ppm.

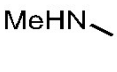


Hydroxylysine

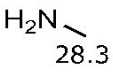
Solution state ^{13}C NMR shifts of various amines shown below (https://www.chem.wisc.edu/areas/reich/nmr/c13-data/cdata.htms), exemplifying the shift of ca. 10 ppm to high frequency of the carbon adjacent to a primary amino group (e.g. C1 of *n*-hexylamine 42.7 ppm) on conversion to a **secondary amino group** (e.g. N-methyl *n*-hexylamine 52.4 ppm).



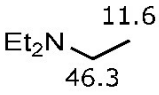
trimethylamine



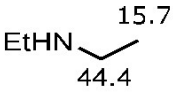
dimethylamine



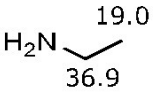
methanamine



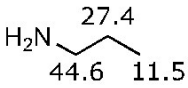
triethylamine



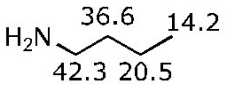
diethylamine



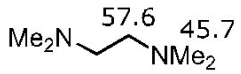
ethanamine



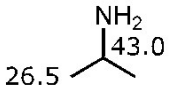
propan-1-amine



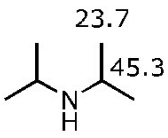
butan-1-amine



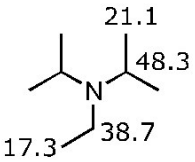
N^1,*N*^{1},*N*^{2},*N*^{2}-tetramethylethane-1,2-diamine



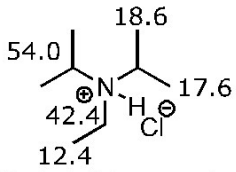
propan-2-amine



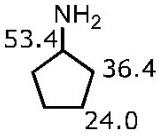
diisopropylamine



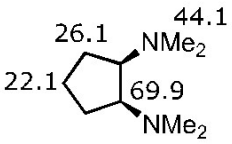
N-ethyl-*N*-isopropyl-propan-2-amine



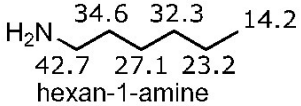
N-ethyl-*N*-isopropyl-propan-2-aminium



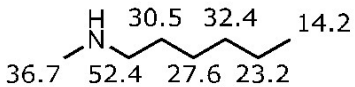
cyclopentanamine



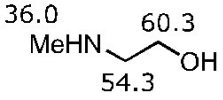
(1*R*,2*S*)-*N*^{1},*N*^{1},*N*^{2},*N*^{2}-tetramethylcyclopentane-1,2-diamine



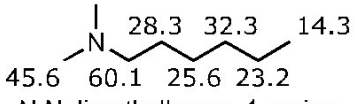
hexan-1-amine



N-methylhexan-1-amine



2-(methylamino)ethan-1-ol



N,*N*-dimethylhexan-1-amine

References

1 D. R. Eyre, M. A. Paz and P. M. Gallop, *Annu. Rev. Biochem.*, 1984, **53**, 717.

2 D. R. Eyre, M. A. Weis and J. J. Wu, *Methods*, 2008, **45**, 65.

3 A. J. Bailey, R. G. Paul and L. Knott, *Mech. Ageing Dev.*, 1998, **106**, 1.

4 D. R. Eyre, I. R. Dickson and K. Van Ness, *Biochem. J.*, 1988, **252**, 495.

5 A. J. van der Slot-Verhoeven, E. A. van Dura, J. Attema, B. Blauw, J. Degroot, T. W. Huizinga, A. M. Zuurmond and R. A. Bank, *Biochim. Biophys. Acta*, 2005, **1740**, 60.

6 M. Sricholpech, I. Perdivara, M. Yokoyama, H. Nagaoka, M. Terajima, K. B. Tomer and M. Yamauchi, *J. Biol. Chem.*, 2012, **287**, 22998.

7 M. Terajima, I. Perdivara, M. Sricholpech, Y. Deguchi, N. Pleshko, K. B. Tomer and M. Yamauchi, *J. Biol. Chem.*, 2014, **289**, 22636.

8 A. G. Huebschmann, J. G. Regensteiner, H. Vlassara and J. E. B. Reusch, *Diabetes Care*, 2006, **29**, 1420.

9 V. M. Monnier, G. T. Mustata, K. L. Biemel, O. Reihl, M. O. Lederer, Z. Y. Dai and D. R. Sell, *Ann. New York Acad. Sci.* , 2005, **1043 Maillard Reaction: Chemistry at the Interface of Nutrition, Aging, and Disease**, 533.

10 M. Saito and K. Marumo, *Osteoporos. In.t*, 2010, **21**, 195.

11 S. Bawazeer, O. B. Sutcliffe, M. R. Euerby, S. Bawazeer and D. G. Watson, *J Chromatogr A*, 2012, **1263**, 61.

12 J. J. Pesek, R. I. Boysen, M. T. W. Hearn and M. T. Matyska, *Anal. Meth.*, 2014, **6**, 4496.

13 C. Sauvee, M. Rosay, G. Casano, F. Aussenac, R. T. Weber, O. Ouari and P. Tordo, *Angewandte Chemie (International ed*, 2013, **52**, 10858.

14 M. Geiger, A. Jagtap, M. Kaushik, H. Sun, D. Stoppler, S. Sigurdsson, B. Corzilius and H. Oschkinat, *Chem. Eur. J.*, 2018, DOI: 10.1002/chem.201801251.

15 G. Metz, X. Wu and S. O. Smith, *J. Magn. Reson. Ser. A*, 1994, **110**, 219.

16 K. Takegoshi, S. Nakamura and T. Terao, *Chem. Phys. Lett.*, 2001, **344**, 631.

17 M. Hohwy, C. M. Rienstra, C. P. Jaroniec and R. G. Griffin, *J. Chem. Phys.*, 1999, **110**, 7983.

18 J. Schaefer, R. A. McKay and E. O. Stejskal, *J. Magn. Reson.*, 1979, **34**, 443.

19 M. Baldus, A. T. Petkova, J. Herzfeld and R. G. Griffin, *Molec. Phys.*, 1998, **95**, 1197.