Supporting Information for:

Characterization of the deoxyguanosine-lysine cross-link of methylglyoxal

Katya V. Petrova, Amy D. Millsap, Donald F. Stec, and Carmelo J. Rizzo*

Departments of Chemistry and Biochemistry, Center in Molecular Toxicology, and Vanderbilt-Ingram Cancer Center, Vanderbilt University, Nashville, TN 37235

Table of Contents

General experimental ......................................................2
Synthesis of cross-link standards (1) ................................................5
Formation of CE-dGuo and cross-link 1 in deuterated buffer ...............................................12

Figure S1. 1H and 13C NMR spectra (DMSO-d6) of (S)-2. .........................................................14
Figure S2. COSY spectrum (CD3OD) of (S)-2. .................................................................15
Figure S3. 1H and 13C NMR spectra (CD3OD) of (R)-2. ......................................................16
Figure S4. 1H and 13C NMR spectra (CD3OD) of (S)-[13C315N1]-2. .................................................17
Figure S5. 1H and 13C NMR spectra (DMSO-d6) of (S)-4. ..............................................................18
Figure S6. 1H and 13C NMR spectra (DMSO-d6) of ((R)-4) .....................................................19
Figure S7. 1H and 13C NMR spectra (DMSO-d6) of (S)-[13C315N1]-4 ............................................20
Figure S8. 1H NMR, 13C and DEPT-135 spectra (DMSO-d6) of (S)-1. ........................................21
Figure S9. COSY spectrum of (S)-1. ................................................................................22
Figure S10. 1H and 13C NMR spectra of (R)-1 .............................................................................23
Figure S11. 1H and 13C NMR spectra (DMSO-d6) of (S)-[13C315N1]-1. ........................................24
Figure S12. 13C NMR and 13C–15N-decoupled spectra (DMSO-d6) of (S)-[13C315N1]-1. ..............25
Figure S13. 1H NMR spectra of (R)- and (S)-CE-dGuo adducts (D2O) ........................................26
Figure S14. 1H spectra of synthetic standard (R)-CE-dGuo adduct vs. the (R)-CE-dGuo adduct isolated from reaction of dGuo and methylglyoxal in deuterated buffer. .......................27
Figure S15. 1H spectra of synthetic cross-link standard (R)-1 vs. the (R)-1 cross-link isolated from reaction of dGuo, N-acetyllysine and methylglyoxal in deuterated buffer. .................28
Figure S16. 1H NMR spectra of both diastereomers of 1,N2-cyclic dGuo adduct (MG-dGuo) ..........29
Figure S17. LC-ESI-MSn of the CE-dGuo adducts ..................................................................30
Figure S18. Product ion mass spectra of cross-link 1 and (S)-[13C315N1]-1 ..................................31
Figure S19. MS calibration curve of 1 vs. (S)-[13C315N1]-1 ......................................................33
Figure S20. Reconstructed UPLC-ESI-MS ion chromatogram from the reaction dGuo, methylglyoxal, and AcAVAGKAGAR (1:1:1) at pH 7.4 and 37 °C. ........................................34
Figure S21. CID mass spectrum of the Ac-AVAGKAGAR peptide (m/z = 421.8 [M+2H]+) ..........36
Figure S22. CID mass spectrum of the cross-link 11 (m/z = 560 [M+2H]+) ..................................37
Figure S23. CID mass spectrum of the cross-link 12 (m/z = 1101.3 [M+H]+) ...............................38
Figure S24. CID mass spectrum of the arginine adduct 13 (m/z = 457.8 [M+2H]+) .....................39
Figure S25. CID mass spectrum of the arginine adduct 14 (m/z = 448.8 [M+2H]+) ..........................40
Figure S26. CID mass spectrum of the arginine + CE-Lys adduct 15 (m/z = 493.8 → 484.8 → 462.7 [M+2H]+) .....................................................................................................................41
Figure S27. CID Mass spectrum of the unknown Arginine + 84 adduct ([M+2H]+ = 463.74) ......42
**General Experimental.** All commercially obtained chemicals were used as received.

**NMR.** NMR experiments were acquired using a 14.0 T Bruker magnet equipped with a Bruker Avance-III console operating at 600.13 MHz (Bruker, Billerica, MA). All spectra were acquired in 3 mm NMR tubes using a Bruker 5 mm TCI cryogenically cooled NMR probe. Chemical shifts were referenced internally to DMSO-$d_6$ (2.5 ppm), CD$_3$OD (3.1 ppm) or D$_2$O (4.70 ppm), which also served as the $^2$H lock solvents. Typical 1D $^1$H NMR experimental conditions included 32K data points, 13 ppm sweep width, a recycle delay of 1.5 seconds and 32-256 scans depending on sample concentration. Typical 1D $^{13}$C NMR experimental conditions included 32K data points, 240 ppm sweep width, a 5° pulse tip angle, 2-3 sec recycle delay and 512-2048 scans depending on sample concentration. Experimental conditions for 2D $^1$H-$^1$H COSY included 2048 x 512 data matrix, 13 ppm sweep width, recycle delay of 1.5 seconds and 4 scans per increment. The data was processed using squared sinebell window function, symmetrized, and displayed in magnitude mode. Multiplicity-edited HSQC experiments were acquired using a 1024 x 256 data matrix, a $J_{C-H}$ value of 145 Hz, which resulted in a multiplicity selection delay of 34 ms, a recycle delay of 1.5 seconds and 16 scans per increment along with GARP decoupling on $^{13}$C during the acquisition time (150 ms). The data was processed using a $\pi/2$ shifted squared sine window function and displayed with the CH/CH$_3$ signals phased positive and the CH$_2$ signals phased negative. Acquisition parameters for $^{13}$C-$^{15}$N decoupling experiments were similar to those listed above for a standard $^{13}$C NMR experiment with the exception of an additional 200 µs pulse applied to the $^{15}$N channel during acquisition using the WALTZ-16 pulse scheme.

**HPLC** Analysis and purification of the nucleoside adducts were conducted on a Beckman HPLC gradient system with a diode array UV detector monitoring at 260 nm using a Phenomenex Luna 5 µm C18(2) 100A columns (4.6 x 250 mm, 5 µm) at a 1.0 mL min$^{-1}$ for analysis and (250 x 10 mm, 5 µm) at a 3 mL min$^{-1}$ for purification. The mobile phase consisted of either H$_2$O or 100 mM ammonium formate buffer and either CH$_3$CN or MeOH using the following gradients: **HPLC gradient I:** initially 90 % H$_2$O, 10 % CH$_3$CN; a 25 min linear gradient to 20 % H$_2$O, 80 % CH$_3$CN; followed by a 5 min linear gradient to the initial conditions. **HPLC gradient II:** initially 99 % H$_2$O, 1 % CH$_3$CN; a 15 min linear gradient to 90 % H$_2$O, 10 % CH$_3$CN; a 5 min linear gradient to 80 % H$_2$O, 20 % CH$_3$CN; isocratic at 80 % H$_2$O,
20 % CH$_3$CN for 5 min; 10 min linear gradient to 20 % H$_2$O, 80 % CH$_3$CN; isocratic at 20 % H$_2$O, 80 % CH$_3$CN for 5 min; followed by a 5 min linear gradient to the initial conditions. **HPLC gradient III:** initially 99 % ammonium formate buffer, 1 % MeOH; a 15 min linear gradient to 90 % buffer, 10 % MeOH; a 3 min linear gradient to 88 % buffer, 12 % MeOH; isocratic at 88 % buffer, 12 % MeOH for 10 min; a linear gradient to 20 % buffer, 80 % MeOH for 3 min; isocratic at 20 % buffer, 80 % MeOH for 3 min; followed by a 5 min linear gradient to the initial conditions. **HPLC gradient IV (flow rate 2.5 mL min$^{-1}$):** initially 99 % ammonium formate buffer, 1 % CH$_3$CN; a 15 min linear gradient to 90 % buffer, 10 % CH$_3$CN; a 7 min linear gradient to 80 % buffer, 20 % CH$_3$CN; a linear gradient to 20 % buffer, 80 % CH$_3$CN for 3 min; isocratic at 20 % buffer, 80 % CH$_3$CN for 2 min; followed by a 5 min linear gradient to the initial conditions.

**Solid Phase Extraction:** SPE purification was performed on a Phenomenex, Strata-X 33 µm Polymeric Reversed Phase (200 mg/3mL and 100 mg/3mL) using the manufacturer’s suggested protocols.

**Mass spectrometry.** Low resolution mass spectra were obtained on an LTQ with electrospray ionization. High resolution mass spectra for compounds (S)-2, (R)-2, (S)-4, (R)-4, (S)-1, and (R)-1 were recorded on a Waters Synapt hybrid quadrupole oa-TOF high resolution mass spectrometer operated in “w” mode, posESI. A post-acquisition gain correction factor was applied using sodium formate as the lock mass. High-resolution FAB mass spectra for the isotopically labeled compound [$^{13}$C$_3$$^{15}$N$_1$]-(S)-2, [$^{13}$C$_3$$^{15}$N$_1$]-(S)-4, and [$^{13}$C$_3$$^{15}$N$_1$]-(S)-1 were obtained from the University of the Notre Dame Mass Spectrometry Center using nitrobenzyl alcohol (NBA) as a matrix. Samples analysis were carried out using a Waters Acquity UPLC system (Waters, Milford, MA) made up of a binary solvent manager, refrigerated sample manager, and a heated column manager and equipped with a Kinetex™ 2.6 µm C18 100 Å, LC Column 100 × 2.1 mm column (Phenomenex). Tandem mass spectrometric detection was performed using LTQ linear ion trap mass spectrometer (Thermo-Electron, San Jose, CA) equipped with an Ion Max API electrospray ionization (ESI) source and a 50 µm ID stainless steel capillary. The instrument was tuned and calibrated every six to eight weeks over the mass range of m/z 195 to m/z 1822 using a mixture of caffeine, MRFA, and Ultramark-1621 following the manufacturer’s autotune.
procedure. Data acquisition and quantitative spectral analysis were done using Thermo-Finnigan Xalibur version 2.0 SUR1 and Thermo-Finnigan LCQuan version 2.5.5, respectively.

**Nucleoside Analysis and Quantitation.** The mass spectrometer was operated in negative ionization mode. Quantitation was based on selected reaction monitoring (SRM) detection at normalized collision energy of 20 % (cross-link (S)-1: m/z 508 →392, (S)-[13C3 15N1]-1: m/z 512 →396). The following optimized parameters were used for the detection of analyte, calibrant, and internal standard: N2 sheath gas, 65 psi; N2 auxiliary gas, 6 psi; source spray voltage, 4.5 kV; capillary temperature, 350 °C; capillary offset, 35 eV; source current, 6.5 µA; tube lense offset, 80 V; AGC max inject time, 5 ms. Individual MS instrument parameters were optimized by infusing the crosslink 1 with the syringe pump into the MS source through a mixing tee at a flow rate of 0.025 mL/min. The LC solvent (1:1 A/B) flowed at 0.15 mL/min. Other MS parameters were optimized to achieve maximum signal intensity. Liquid chromatography was carried out at 45° C at a flow of to 0.2 µL min⁻¹ using gradient mixture of 1:99 CH₃CN in 10 mM ammonium acetate buffer (solvent A) and 90:10 CH₃CN in 10 mM ammonium acetate buffer (solvent B). The gradient program began at 1% B, then a 2 min linear gradient to 10% B, a 3 min linear gradient to 50% B, a 2 min linear gradient to 100% B, then isocratic at 100% B for 0.5 min, then a 0.5 min linear gradient back to 1% B, followed by a 3 min isocratic period to allow the column to re-equilibrate to the starting conditions (11 min, **UPLC gradient I**). Samples (10 µL) were injected through a 10 µL injection loop into a six-port switching valve injector that diverted the column eluent to waste for the first 2.0 min of **UPLC gradient I**.

**Peptide cross-link analysis.** The mass spectrometer was operated in positive ionization mode. LTQ MS parameters were optimized for maximum response during infusion of a standard solution of the synthetic AcAVAGKAGAR peptide, so that some parameters may have varied slightly from experiment to experiment. Typically, the tune parameters were as follows: N2 sheath gas, 60 psi; N2 auxiliary gas, 10 psi; source spray voltage, 4.5 kV; capillary temperature, 350° C; capillary offset, 35 eV; source current, 6.5 µA; tube lense offset, 80 V; AGC max inject time, 5 ms. A full scan was obtained in the range of 125–2000 amu and MS/MS spectra were generated by collision induced dissociation of the peptide ions at normalized collision energy of 35 % to generate a series of b- and y-ions as major
fragments. UPLC-ESI+-MS/MS analyses were performed in the SRM mode using the transition corresponding to a major fragment ion observed upon CID fragmentation of peptide-methylglyoxal-dGuo cross-links or CID fragmentation of peptide-methylglyoxal adducts in a linear ion trap mass spectrometer. Liquid chromatography was carried out at 45° C at a flow of to 0.15 mL min⁻¹ using gradient mixture of 0.1 % (v/v) trifluoroacetic acid in water (solvent A) and 0.08 % (v/v) trifluoroacetic acid in CH₃CN (solvent B). The gradient program began at 1% B, then a 2 min linear gradient to 10 % B, a 3 min linear gradient to 50 % B, a 3 min linear gradient to 100% B, then isocratic at 100 % B for 0.5 min, then a 0.5 min linear gradient back to 1 % B, followed by a 3 min isocratic period to allow the column to re-equilibrate to the starting conditions (12 min, UPLC gradient II). Samples (10 µL) were injected through a 10 µL injection loop into a six-port switching valve injector that diverted the column eluent to waste for the first 2.0 min of UPLC gradient II. Predicted CID fragment ions were calculated using the MS/MS Fragment Ion Calculator from the Institute for Systems Biology (http://db.systemsbiology.net:8080/proteomicsToolkit/)

**Synthesis of cross-link standards (1) via Scheme 1.**

\[
\text{N}^2\text{acetyl-N}^6-[\text{N}-(\text{9H-fluoren-9-ylmethylene}xy)\text{carbonyl}]\text{-L-alanyl-L-lysine (}(\text{S})-2)\.
\]

A solution of diisopropylethylamine (50.4 µL, 0.29 mmol) in dry DMSO (4.0 mL) was added to dry flask containing FMOCL-Ala (45.15 mg, 0.145 mmol) and HBTU (55.0 mg, 0.145 mmol) and the resulting solution was stirred at room temperature for 4h. Nα-acetyl-L-lysine methyl ester (38.09 mg, 0.160 mmol) was added and the reaction was monitored by HPLC using gradient I. After 1 h, the reaction mixture was poured onto crushed ice (50 mL) and extracted with methylene chloride (100 mL), washed with brine, concentrated and kept overnight under high vacuum. Purification by flash chromatography on silica, eluting with 3-4 % methanol in methylene chloride, afforded (S)-2 (63.5 mg, 88 %). FAB HRMS⁺ m/z
calcd for C$_{27}$H$_{34}$N$_3$O$_6$ [M+H]$^+$, 496.2448; found, 496.2434; m/z calcd for C$_{27}$H$_{33}$N$_3$NaO$_6$ [M+H+Na]$^+$, 518.2267; found, 518.2251. $^1$H NMR (400 MHz, CD$_3$OD) δ 1.32 (d, 3H, J = 7.2 Hz, CH$_3$), 1.30-1.38 (m, 2H, HNCH$_2$CH$_2$), 1.43-1.55 (m, 2H, HNCH$_2$CH$_2$), 1.61-1.71 (m, 1H, CH$_2$CHNHAc), 1.74-1.81 (m, 1H, CH$_2$CHNHAc), 1.96 (s, 3H, NHCOC$_3$H), 3.14-3.20 (m, 2H, NCH$_2$), 3.67 (s, 3H, COOCH$_3$), 4.07 (q, 1H, J = 7.1 Hz, CHCH$_3$), 4.23 (t, 1H, J = 6.7 Hz, CH$_2$CH$_2$), 4.35 (dd, 1H, J = 5.0 Hz, J = 8.8 Hz, CHNHAc), 4.39 (d, 1H, J = 6.8 Hz, CH$_2$OCO), 7.31 (dt, 2H, J = 1.0, 7.5 Hz, Ar), 7.39 (t, 2H, J = 7.4 Hz, Ar), 7.66 (t, 2H, J = 6.7 Hz, Ar), 7.77 (t, 2H, J = 7.5 Hz, Ar). $^{13}$C NMR (150.9 MHz, DMSO-d$_6$) δ 172.8, 172.3, 169.5, 155.7, 143.9, 143.8, 140.7, 127.6, 127.1, 125.3, 120.1, 65.6, 51.9, 51.7, 50.1, 46.7, 38.1, 30.6, 28.6, 22.6, 22.2, 18.4.

N$_2$-acetyl-N$_6$-[N-((9H-fluoren-9-ylmethylloxycarbonyl)-D-alanyl]-L-lysine ((R)-2). The diastereomeric (R)-2 was synthesized from FMOC-D-Ala in 78 % yield using the same procedure as for (S)-2. FAB HRMS$^+$ m/z calcd for C$_{27}$H$_{34}$N$_3$O$_6$ [M+H]$^{+1}$, 496.2448; found, 496.2430; m/z calcd for C$_{27}$H$_{33}$N$_3$NaO$_6$ [M+H+Na]$^{+1}$, 518.2267; found, 518.2250. $^1$H NMR (400 MHz, CD$_3$OD) δ 1.33 (d, 3H, J = 7.5 Hz, CH$_3$), 1.29-1.36 (m, 2H, HNCH$_2$CH$_2$CH$_2$), 1.45-1.50 (m, 2H, HNCH$_2$CH$_2$), 1.60-1.70 (m, 1H, CH$_2$CHNHAc), 1.73-1.82 (m, 1H, CH$_2$CHNHAc), 1.95 (s, 3H, NHCOC$_3$H), 3.14-3.19 (m, 2H, NCH$_2$), 3.67 (s, 3H, COOCH$_3$), 4.05 (q, 1H, J = 7.1 Hz, CHCH$_3$), 4.22 (t, 1H, J = 6.8 Hz, CH$_2$CH$_2$), 4.32 (dd, 1H, J = 5.1, 8.9 Hz, CHNHAc), 4.39 (d, 1H, J = 6.3 Hz, CH$_2$OCO), 7.29 (dt, 2H, J = 1.0, 7.4 Hz, Ar), 7.36 (t, 2H, J = 7.4 Hz, Ar), 7.64 (t, 2H, J = 6.6 Hz, Ar), 7.77 (t, 2H, J = 7.5 Hz, Ar). $^{13}$C NMR (100.6 MHz, CD$_3$OD) δ 175.6, 174.3, 173.4, 158.2, 145.4, 145.2, 142.6, 128.8, 128.2, 126.2, 121.0, 67.9, 55.9, 53.8, 52.7, 52.2, 43.7, 39.9, 32.1, 29.9, 24.1, 22.3, 18.5.

N$_2$-acetyl-N$_6$-[N-((9H-fluoren-9-ylmethylloxycarbonyl)-L-[13C$_3$15N$_1$]-alanyl]-L-lysine ([13C$_3$15N$_1$]-S)-2). Isotopically labeled [13C$_3$15N$_1$]-S)-2 was synthesized from [13C$_3$15N$_1$]-FMOC-L-Ala in 77 % yield using the same procedure as for (S)-2. FAB HRMS$^+$ m/z calcd for C$_{24}$H$_{34}$N$_2$O$_6$C$_3$$_{15}$N$_1$
[M+H]$^+$, 500.2513; found, 500.2507; m/z calcd for C$_{24}$H$_{33}$N$_2$NaO$_6^{13}$C$_3^{15}$N$_1$ [M+H+Na]$^+$, 522.2333; found, 522.2336. $^1$H NMR (600 MHz, CD$_3$OD) δ 1.18-1.41 (m, 5H, HNCH$_2$CH$_2$CH$_2$, CH$_3$), 1.46-1.51 (m, 2H, HNCH$_2$CH$_2$), 1.62-1.68 (m, 1H, CH$_2$CHNHAc), 1.74-1.79 (m, 1H, CH$_2$CHNHAc), 1.95 (s, 3H, NHCOCH$_3$), 3.15-3.19 (m, 2H, NCH$_2$), 3.66 (s, 3H, COOCH$_3$), 3.92-3.94 and 4.16-4.18 (m, 1H, J(CH) = 6 Hz, J(CH) = 141.9 Hz, CHCH$_3$), 4.20 (t, 1H, J = 6.8 Hz, CHCH$_2$), 4.34 (dd, 1H, J = 5.0, 8.7 Hz, CHNHAc), 4.36 (d, 1H, J = 6.3 Hz, CH$_2$OCO), 7.29 (t, 2H, J = 7.5 Hz, Ar), 7.37 (t, 2H, J = 7.5 Hz, Ar), 7.66 (t, 2H, J = 8.2 Hz, Ar), 7.77 (t, 2H, J = 7.6 Hz, Ar). $^{13}$C NMR (150.9 MHz, CD$_3$OD) δ 175.7 (d, J(CC) = 26 Hz, $^{13}$CO), 174.2, 173.4, 158.3, 145.4, 145.2, 142.6, 128.8, 128.2, 126.2, 121.0, 67.9, 53.8, 52.6, 52.2 (ddd, J(CN) = 12.1 Hz, J(CC) = 17.2 Hz, J(CC) = 35.5 Hz, 15NH$^{13}$CH($^{13}$CH$_3$)$^{13}$CO), 48.4, 39.8, 32.0, 29.8, 23.9, 22.3, 18.5 (d, J(CC) = 35.3 Hz, 13CH$_3$).

Methyl N2-acetyl-N6-[N-(6,9-dihydro-6-oxo-9[2-deoxy-β-D-2-erythro-pentofuranosyl]-1H-purin-2-yl)-L-alanyl]-L-lysine ((S)-4). A solution of (S)-2 (75.0 mg, 0.15 mmol) in anhydrous DMSO (400 µL) was added to a solution of 2-fluoro-O$^6$-(2-trimethylsilylethyl)-2'-deoxyinosine$^{36}$ (3,14.5 mg, 0.039 mmol) and diisopropylethylamine (200 µL) in DMSO (100 µL). The reaction was stirred at 60-65°C and monitored by analytical HPLC using gradient I or II. After 3 days, the reaction mixture was cooled and concentrated in vacuo using a centrifugal evaporator. The residue was dissolved in 5% aqueous acetic acid (1.5 mL) and stirred at room temperature. After 15 min, the reaction mixture separated into an aqueous and oil layers; the mixture was stirred for 1 h, then the water layer was separated and neutralized to pH 7.0 with 1% NaOH. The product mixture was filtered through a 0.45 µm filter cartridge, rinsed with water (3 × 1 mL), and lyophilized. The resulting solid residue was dissolved in water (3 mL) and purified by semi-preparative HPLC using gradient II (flow rate 3.0 mL/min). The fractions were immediately frozen after collection. Nucleoside (S)-4 (7.5 mg, 37%) was obtained as a white solid. The purity of the product was judged to be > 99% by HPLC (gradient I or II).
flow rate 1.0 mL min⁻¹). FAB HRMS⁻, m/z calcd for C₂₂H₃₂N₇O₈ [M-H]⁻, 522.2312; found, 522.2324; FAB HRMS⁺, m/z calcd for C₂₂H₃₃N₇NaO₈ [M+H+Na]⁺, 546.2288; found, 546.2289. ¹H NMR (600 MHz, DMSO-d₆) δ 1.28 (d, 3H, J = 6.8 Hz, CH₃), 1.22-1.32 (m, 2H, HNCH₂CH₂CH₂), 1.35-1.41 (m, 2H, HNCH₂), 1.52-1.58 (m, 1H, CH₂CHNHAc), 1.61-1.67 (m, 1H, CH₂CHNHAc), 1.82 (s, 3H, NHCOCH₃), 2.15-2.19 (m, 1H, H-2′′), 2.51-2.54 (m, 1H, H-2′), 3.02-3.11 (m, 2H, NCH₂), 3.46-3.48 (m, 1H, H-5′′), 3.52-3.54 (m, 1H, H-5′), 3.59 (s, 3H, COOCH₃), 3.81 (q, 1H, J = 2.6 Hz, H-4′), 4.14-4.17 (m, 1H, CHNHAc), 4.34 (br s, 1H, H-3′), 4.38 (quintet, 1H, J = 6.9 Hz, CHCH₃), 4.91 (s., 1H, 5′-OH), 5.30 (s, 1H, 3′-OH), 6.14 (dd, 1H, J = 6.6, 7.5 Hz, H-1′), 6.79 (d, 1H, J = 7.0 Hz, NHCHCH₃), 7.93 (s, 1H, H-8), 8.16 (t, 1H, J = 5.6 Hz, NHCH₂), 8.20 (d, 1H, J = 7.4 Hz, NHAc), 10.59 (br s, 1H, NH1). ¹³C NMR (150.9 MHz, DMSO-d₆) δ 172.8, 171.8, 169.5, 156.6, 151.6, 150.3, 135.5, 116.8, 87.6, 82.5, 70.9, 61.8, 51.9, 49.6, 39.6, 38.2, 30.5, 28.5, 22.7, 22.2, 19.4.

Methyl N₂-acetyl-N₆-[N-(6,9-dihydro-6-oxo-9-[2-deoxy-β-D-2-erythro-pentofuranosyl]-1H-purin-2-yl)-D-alanyl]-L-lysine ((R)-4). Following the procedure described above for (S)-4, the reaction of (R)-2 with 3 afforded (R)-4 in 31 % yield. The purity of the product was judged to be > 99 % by HPLC (gradient I or II, flow rate 1.0 mL min⁻¹). FAB HRMS⁻ m/z calcd for C₂₂H₃₂N₇O₈ [M-H]⁻, 522.2312; found, 522.2332; FAB HRMS⁺ m/z calcd for C₂₂H₃₂N₇Na₂O₈ [M+H⁺], 568.2108; found, 568.2100. ¹H NMR (600 MHz, DMSO-d₆) δ 1.30 (d, 3H, J = 6.9 Hz, CH₃), 1.23-1.28 (m, 2H, HNCH₂CH₂CH₂), 1.35-1.41 (m, 2H, HNCH₂), 1.52-1.58 (m, 1H, CH₂CHNHAc), 1.61-1.66 (m, 1H, CH₂CHNHAc), 1.82 (s, 3H, NHCOCH₃), 2.16-2.19 (m, 1H, H-2′′), 2.58-2.63 (m, 1H, H-2′), 3.02-3.11 (m, 2H, NCH₂), 3.46-3.48 (m, 1H, H-5′′), 3.52-3.54 (m, 1H, H-5′), 3.59 (s, 3H, COOCH₃), 3.80 (sextet, 1H, J = 2.4 Hz, H-4′), 4.14-4.18 (m, 1H, CHNHAc), 4.33 (br t, 1H, J = 2.6 Hz, H-3′), 4.38 (quintet, 1H, J = 6.9 Hz, CHCH₃), 4.89 (s., 1H, 5′-OH), 5.27 (s, 1H, 3′-OH), 6.13 (dd, 1H, J = 6.2, 7.8 Hz, H-1′), 6.88 (br s, 1H, J = 7.0 Hz, NHCHCH₃), 7.91 (s, 1H, H-8), 8.13 (t, 1H, J = 5.4 Hz, H-8), 8.16 (t, 1H, J = 5.6 Hz, NHCH₂), 8.20 (d, 1H, J = 7.4 Hz, NHAc), 10.59 (br s, 1H, NH1).
NHCH$_2$), 8.20 (d, 1H, $J$ = 7.4 Hz, NHAc), 10.67 (br s, 1H, NH$_1$). $^{13}$C NMR (150.9 MHz, DMSO-$d_6$) $\delta$

172.8, 171.9, 169.5, 156.8, 151.7, 150.3, 135.8, 117.3, 87.7, 82.7, 71.0, 61.8, 51.9, 51.7, 49.7, 39.6, 38.3, 30.5, 28.5, 22.7, 22.2, 19.3.

Methyl N$_2$-acetyl-N$_6$-[N-(6,9-dihydro-6-oxo-9[2-deoxy-$\beta$-D-2-erythro-pentofuranosyl]-1H-purin-2-yl)-L-alanyl]-L-lysine ([$^{13}$C$_3$$^{15}$N$_1$]-(S)-4). Following the procedure described above for (R)-4, reaction of [$^{13}$C$_3$$^{15}$N$_1$]-(S)-2 (51.01 mg, 0.10 mmol) with 3 afforded [$^{13}$C$_3$$^{15}$N$_1$]-(S)-4 in 47% yield. The purity of the product was judged to be $> 99\%$ by HPLC (gradient system II, flow rate 1.0 mL min$^{-1}$). FAB HRMS$^+$ m/z calcd for C$_{19}$H$_{34}$N$_6$O$_8$$^{13}$C$_3$$^{15}$N$_1$ [M+H]$^+$, 528.2534; found, 528.2550; m/z calcd for C$_{19}$H$_{33}$N$_6$NaO$_8$$^{13}$C$_3$$^{15}$N$_1$ [M+H+Na]$^+$, 550.2354; found, 550.2392. $^1$H NMR (600 MHz, DMSO-$d_6$) $\delta$

1.18-1.19 and 1.28-1.29 (m, 3H, $J_{(CH)}$ = 129.1 Hz, CH$_3$), 1.22-1.29 (m, 2H, HNCH$_2$CH$_2$CH$_2$), 1.35-1.40 (m, 2H, HNCH$_2$), 1.53-1.58 (m, 1H, CH$_2$CHNHAc), 1.61-1.66 (m, 1H, CH$_2$CHNHAc), 1.82 (s, 3H, NHCOCH$_3$), 2.15-2.19 (m, 1H, H-2’′), 2.51-2.54 (m, 1H, H-2’), 3.01-3.10 (m, 2H, NCH$_2$), 3.46-3.55 (m, 2H, H-5′, H-5’′), 3.59 (s, 3H, COOCH$_3$), 3.80-3.82 (m, 1H, H-4′), 4.14-4.18 (m, 1H, CHNHAc), 4.3-4.34 (m, 2H, H-3′), 4.26 and 4.50 (m, 1H, $J_{(CH)}$ = 142.1 Hz, CHCH$_3$), 4.90 (s., 1H, 5′-OH), 5.28 (s, 1H, 3′-OH), 6.13 (t, 1H, $J$ = 7.0 Hz, H-1′), 6.70 and 6.85 (dd, 1H, $J_{(NH)}$ = 7 Hz, $J_{(NH)}$ = 92.1 Hz, NHCH$_2$CH$_3$), 7.93 (s, 1H, H-8), 8.15 (d, 1H, $J$ = 4.4 Hz, NHCH$_2$), 8.20 (d, 1H, $J$ = 7.4 Hz, NHAc), 10.57 (br s, 1H, NH$_1$). $^{13}$C NMR (150.9 MHz, DMSO-$d_6$) $\delta$

175.3, 172.8 (d, $J_{(CC)}$ = 51.9 Hz, $^{13}$CO), 171.8, 169.4, 156.5, 151.4, 150.3, 135.4, 116.8, 87.6, 82.4, 70.9, 61.8, 51.8, 49.5 (ddd, $J_{(CN)}$ = 12.1 Hz, $J_{(CC)}$ = 16.6 Hz, $J_{(CC)}$ = 35.2 Hz, $^{15}$NH$_{13}$CH($^{13}$CH$_3$)CO), 39.6, 38.2, 30.5, 28.5, 22.7, 22.2, 19.4 (d, $J_{(CC)}$ = 35.0 Hz, $^{13}$CO).

$\text{N}2$-acetyl-N$_6$-[N-(6,9-dihydro-6-oxo-9$\beta$-D-2-erythro-pentofuranosyl)-1H-purin-2-yl]-L-alanyl]-L-lysine ([S]-1). A solution of ([S]-4 (7.25 mg, 0.0138 mmol) was stirred in an aqueous solution of LiOH (150 mM) and THF (750 $\mu$L, 1:2.5 v/v) for 2 h at room temperature. The reaction was monitored by analytical HPLC using gradient II and 100 mM ammonium formate and CH$_3$CN as mobile
phase. When \((S)-4\) was no longer detectable, the mixture was concentrated under high vacuum using a centrifugal evaporator, then neutralized and very carefully acidified to pH 5.5 with 5 % acetic acid. Purification was achieved by solid phase extraction (Phenomenex Strata-X 33µ 200 mg/3 mL), starting with 0.1 % formic acid, then 5 % MeOH, and 100 % MeOH. The desired compound eluted in the 100 % MeOH fraction, which was concentrated under high vacuum using a centrifugal evaporator. The residue was dissolved in water (3 mL) and lyophilized to afforded \((S)-1^{37}\) (6.0 mg, 85 %) as a white solid. The purity of the product was judged to be > 99 % by HPLC (gradient II, 100 mM HCOONH₄/ CH₃CN, flow rate 1.0 mL min⁻¹). UV \(\lambda_{max}\) (1:1 MeOH/H₂O) 254 nm (ε 12,482); FAB HRMS\(^{-}\) \(m/z\) calcd for C₂₁H₃₀N₇O₈ [M-H]⁻, 508.2156; found, 508.2147; \(^1\)H NMR (600 MHz, DMSO-\(d_6\)) δ 1.30 (d, 3H, \(J = 6.8\) Hz, CH₃), 1.25-1.32 (m, 2H, HNCH₂CH₂CH₂), 1.35-1.43 (m, 2H, HNCH₂CH₂), 1.50-1.56 (m, 1H, CH₂CHNHAc), 1.62-1.68 (m, 1H, CH₂CHNHAc), 1.82 (s, 3H, NHCOCH₃), 2.16-2.20 (m, 1H, H-2′′), 2.50-2.55 (m, 1H, H-2′), 3.0-3.11 (m, 2H, NCH₂), 3.48 (dd, 1H, \(J = 11.6\) Hz, H-5′′), 3.52 (dd, 1H, \(J = 4.9\), 11.6 Hz, H-5′), 3.81 (sextet, 1H, \(J = 2.4\) Hz, H-4′), 4.10 (sextet, 1H, \(J = 4.3\) Hz, CHNHAc), 4.34 (quintet, 1H, \(J = 6.9\) Hz, CHCH₃), 6.14 (dd, 1H, \(J = 6.4, 7.6\) Hz, H-1′), 6.79 (d, 1H, \(J = 7.2\) Hz, NHCHCH₃), 7.93 (s, 1H, H-8), 8.04 (d, 1H, \(J = 7.7\) Hz, NHAc), 8.16 (t, 1H, \(J = 5.6\) Hz, NHCH₂), 10.61 (br s, 1H, NH1). \(^{13}\)C NMR (150.9 MHz, DMSO-\(d_6\)) δ 173.9, 171.8, 169.3, 156.6, 151.6, 150.3, 135.5, 116.8, 87.7, 82.5, 70.9, 61.8, 51.8, 49.6, 40.0, 38.3, 30.7, 28.6, 22.8, 22.4, 19.4.

**N2-acetyl-N6-[N-(6,9-dihydro-6-oxo-9[2-deoxy-β-D-2-erythro-pentofuranosyl]-1H-purin-2-yl)-D-alanyl]-L-lysine (\((R)-1\)).** Following the procedure described above for \((S)-1\), alkaline hydrolysis of \((R)-4\) afforded \((R)-1^{37}\) in 84 % yield after purification by solid phase extraction. The purity of the product was judged to be > 99 % by HPLC (gradient II, HCOONH₄/ CH₃CN, flow rate 1.0 mL min⁻¹). FAB HRMS\(^{-}\) \(m/z\) calcd for C₂₁H₃₀N₇O₈ [M-H]⁻, 508.2156; found, 508.2150; \(^1\)H NMR (600 MHz,
DMSO-$d_6$ δ 1.30 (d, 3H, $J = 6.8$ Hz, CH$_3$), 1.25-1.30 (m, 2H, HNCH$_2$CH$_2$CH$_2$), 1.35-1.42 (m, 2H, HNCH$_2$), 1.50-1.56 (m, 1H, CH$_2$CHNHAc), 1.62-1.68 (m, 1H, CH$_2$CHNHAc), 1.82 (s, 3H, NHCOCH$_3$), 2.16-2.20 (m, 1H, H-2$''$), 2.58-2.63 (m, 1H, H-2$'$), 3.0-3.07 (m, 2H, NCH$_2$), 3.44 (dd, 1H, $J = 4.8$, 11.5 Hz, H-5$''$), 3.53 (dd, 1H, $J = 5.1$, 11.6 Hz, H-5$'$), 3.80 (sextet, 1H, $J = 2.5$ Hz, H-4$'$), 4.10 (sextet, 1H, $J = 4.3$ Hz, CHNHAc), 4.33 (quintet, 1H, $J = 2.7$ Hz, H-3$'$), 4.37 (quintet, 1H, $J = 6.9$ Hz, CHCH$_3$), 6.13 (dd, 1H, $J = 6.2$, 7.9 Hz, H-1$'$), 6.81 (d, 1H, $J = 7.2$ Hz, NHCH$_3$), 7.91 (s, 1H, H-8), 8.03 (d, 1H, $J = 7.7$ Hz, NHAc), 8.16 (t, 1H, $J = 5.6$ Hz, NHCH$_2$), 10.61 (br s, 1H, NH1). 13C NMR (150.9 MHz, DMSO-$d_6$) δ 173.9, 171.8, 169.3, 156.6, 151.5, 150.3, 135.8, 117.0, 87.7, 82.7, 71.0, 61.9, 51.8, 49.6, 39.1, 38.3, 30.8, 28.6, 22.8, 22.4, 19.4.

$N_2$-acetyl-$N_6$-[N-(6,9-dihydro-6-oxo-9[2-deoxy-$\beta$-D-2-erythro-pentofuranosyl]-1H-purin-2-yl)-L-$[^{13}C_3^{15}N_1]$-alanyl]-L-lysine ($[^{13}C_3^{15}N_1]$-(S)-1). Following the procedure described above for (S)-1, alkaline hydrolysis of [$^{13}C_3^{15}N_1$]-(S)-4 afforded [$^{13}C_3^{15}N_1$]-(S)-1 in 49 % yield after purification by solid phase extraction. The purity of the product was judged to be > 99 % by HPLC (gradient II, flow rate 1.0 mL min$^{-1}$). FAB HRMS$^+$ $m/z$ calcd for C$_{18}$H$_{32}$N$_6$O$_8$$^{13}$C$_3$$^{15}$N$_1$ [M+H]$^+$, 514.2378; found, 514.2372; $m/z$ calcd for C$_{18}$H$_{31}$N$_6$NaO$_8$$^{13}$C$_3$$^{15}$N$_1$ [M+H+Na]$^+$, 536.2197; found, 536.2206. 1H NMR (600 MHz, DMSO-$d_6$) δ 1.16-1.17 and 1.37-1.39 (m, 3H, $J_{(CH)} = 128.1$ Hz, CH$_3$), 1.19-1.27 (m, 2H, HNCH$_2$CH$_2$CH$_2$), 1.33-1.41 (m, 2H, HNCH$_2$), 1.57-1.59 (m, 2H, CH$_2$CHNHAc), 1.81 (s, 3H, NHCOCH$_3$), 2.20-2.24 (m, 1H, H-2$''$), 2.53-2.57 (m, 1H, H-2$'$), 3.02-3.08 (m, 2H, NCH$_2$), 3.46-3.56 (m, 2H, H-5$'$,H-5$''$), 3.76-3.77 (m, 1H, H-4$'$), 3.92 (q, 1H, $J = 6.1$ Hz, CHNHAc), 4.41 (m, 1H, H-3$'$), 4.28 and 4.52 (m, 1H, $J_{(CH)} = 141.3$ Hz, CHCH$_3$), 5.05 (s., 1H, 5$'$-OH), 5.81 (s, 1H, 3$'$-OH), 6.11 (t, 1H, $J = 6.9$ Hz, H-1$'$), 7.39 (d, 1H, $J = 7.4$ Hz, NHAc), 7.84 (s, 1H, H-8), 8.06 (br d, 1H, $J = 3.9$ Hz, NHCH$_2$), 11.65 (br s, 1H, NH1). 13C NMR (150.9 MHz, DMSO-$d_6$) δ 175.7, 172.3 (d, $J_{(CC)} = 52$ Hz, $^{13}$CO), 168.5, 157.5, 152.7, 150.7, 136.0,
117.5, 87.9, 83.0, 71.0, 62.1, 54.5, 50.4 (ddd, \(J_{(CN)}\) = 12.1 Hz, \(J_{(CC)}\) = 16.6 Hz, \(J_{(CC)}\) = 35.2 Hz, \(^{15}\text{NH}_{13}\text{CH}^{(13}\text{CH}_{3})^{13}\text{CO}), 40.4, 38.3, 31.7, 28.3, 23.3, 22.2, 19.2 (d, \(J_{(CC)}\) = 35.3 Hz, \(^{13}\text{CH}_{3}).

**Formation of CE-dGuo in deuterated buffer.** A D\(_2\)O solution of NaH\(_2\)PO\(_4\) monohydrate (1.9 mL, 100 mM) was added to a D\(_2\)O solution of anhydrous Na\(_2\)HPO\(_4\) (8.1 mL, 100 mM) and the pH was adjusted to 7.4 with a NaH\(_2\)PO\(_4\) solution (100 mM). The phosphate buffer (100 mM, pH 7.4) was lyophilized, then D\(_2\)O (10 mL) was added to the solid residue and the solution lyophilized again. This procedure was repeated three times. dGuo\(\cdot\)H\(_2\)O (14.25 mg, 0.053 mmol) was dissolved in D\(_2\)O phosphate buffer (2.5 mL) and sonicated for 15 min and the solution was lyophilized. The solid was dissolved in 2.5 mL D\(_2\)O and the solution was lyophilized; this procedure was repeated three times, after which the solid residue was dried under vacuum for 2 days. D\(_2\)O (2.5 mL) was added and the pH was adjusted to 7.4 with 10% NaOD in D\(_2\)O. The reaction mixture was incubated at 37°C and methylglyoxal (3.78 mg, 0.0525 mmol) was added in portions (0.0015 mmol every 2 hr.) over 5 days (7 additions per day); the reaction was stirred at 37°C for 2 additional days after the addition was complete and monitored by HPLC (gradient II, 100 mM HCOONH\(_4\)/MeOH, flow rate 1.0 mL min\(^{-1}\)). The reaction mixture was then lyophilized and the solid residue was dissolved in H\(_2\)O (1.5 mL) and the pH was adjusted to 7.0. The solution was filtered through a 0.45 \(\mu\)m filter cartridge and purified by HPLC (gradient III: 100 mM HCOONH\(_4\)/MeOH, flow rate 2.5 mL min\(^{-1}\), see supporting information section). The fractions were immediately frozen after collection to afford (\(R\))-CE-dGuo (3.4 mg, 20 % yield) as a white solid. Under these chromatographic conditions, (\(S\))-CE-dGuo co-eluted with one of the 2:1 methylglyoxal-dGuo adducts and was not isolated. \(^1\)H NMR (600 MHz, D\(_2\)O) \(\delta\) 1.33 (s, 3H, CH\(_3\)), 2.29-2.33 (m, 1H, H-2'), 2.93-2.98 (m, 1H, H-2'), 3.68 (dd, 1H, \(J = 6.3\) Hz, \(J = 12.0\) Hz, H-5'), 3.72 (dd, 1H, \(J = 4.2\) Hz, \(J = 12.2\) Hz, H-5'), 3.94-3.97 (m, 1H, H-4'), 4.50-4.52 (m, 1H, H-3'), 6.19 (t, 1H, \(J = 6.9\) Hz, H-1'), 7.81 (s, 1H, H-8).

An authentic standard of (\(R\))-CE-dGuo was prepared as previously described for comparison.\(^{38}\) A sample of (\(R\))-CE-dGuo was dissolved in deuterated phosphate buffer (pH 7.4, 100 mM) and incubated
at 37°C over 7 days in an NMR tube. No exchange of the proton between the $N^2$ and carboxylate groups was observed.

**Formation of cross-link 1 in deuterated buffer.** A dGuo (14.70 mg, 0.0515 mmol) solution in deuterated phosphate buffer (2.5 mL, 100 mM, pH 7.4) was prepared as described above. N-AcLys (38.68 mg, 205 mmol) then added and the reaction incubated 37°C. Methylglyoxal was added to the reaction in portions (0.0015 mmol every 2 h) over 4 days and the reaction was incubated at 37°C for an additional day after the methylglyoxal addition was complete. The reaction mixture was lyophilized and the solid residue dissolved in H$_2$O (750 µL). The pH was carefully adjusted to 6.0 with 1% NaOH and the mixture purified by Phenomenex, Strata-X 33u Polymeric Reversed Phase (200 mg/3mL and 100 mg/3mL) solid-phase extraction (SPE) cartridge. The product was sequentially eluted with water (5 mL), 5% MeOH (5 mL), then 100% MeOH (7 mL) and the fractions were monitored by HPLC (gradient II, 100 mM HCOONH$_4$/MeOH, flow rate 1.0 mL min$^{-1}$). Methanol was removed in vacuo using a centrifugal evaporator. The residue was dissolved in water and purified by semi-preparative HPLC using gradient IV (100 mM HCOONH$_4$/CH$_3$CN, flow rate 2.5 mL min$^{-1}$). The fractions were immediately frozen after collection to afford (R)-1 (1.3 mg, 4.9 %) and (S)-1 (1.4 mg, 5.3 %) as a white solid. The (R)-1 nucleoside was analyzed by NMR and the spectrum was compared with authentic standard. $^1$H NMR (600 MHz, DMSO-d$_6$) δ 1.28 (s, 3H, CH$_3$), 1.24-1.30 (m, 2H, HNCH$_2$CH$_2$CH$_2$), 1.35-1.44 (m, 2H, HNCH$_2$CH$_2$), 1.50-1.56 (m, 1H, CH$_2$CHNHAc), 1.62-1.68 (m, 1H, CH$_2$CHNHAc), 1.82 (s, 3H, NHCOCH$_3$), 2.16-2.20 (m, 1H, H-2′′), 2.59-2.63 (m, 1H, H-2′), 3.05 (dd, 2H, J = 6.6, 12.8 Hz, NCH$_2$), 3.44 (dd, 1H, J = 4.8, 11.5 Hz, H-5′′), 3.53 (dd, 1H, J = 5.2, 11.6 Hz, H-5′), 3.79 (sextet, 1H, J = 2.5 Hz, H-4′), 4.07 (m, 1H, CHNHaC), 4.33 (quintet, 1H, J = 2.7 Hz, H-3′), 6.13 (dd, 1H, J = 6.3, 7.8 Hz, H-1′), 6.86 (s, 1H, NHCHCH$_3$), 7.90 (s, 1H, H-8), 7.92 (d, 1H, J = 7.0 Hz, NHAc), 8.16 (t, 1H, J = 5.4 Hz, NHCH$_2$).
Figure S1. $^1$H (top) and $^{13}$C (bottom) NMR spectra (DMSO-$d_6$) of $N_2$-acetyl-$N_6$-[N-(9H-fluoren-9-ylmethyloxy)carbonyl]-L-alanyl]-L-lysine ((S)-2).
Figure S2. COSY spectrum (CD$_3$OD) of N2-acetyl-N6-[N-(9H-fluoren-9-ylmethyloxy)carbonyl]-L-alanyl]-L-lysine ((S)-2).
Figure S3. $^1$H (top) and $^{13}$C (bottom) NMR spectra (CD$_3$OD) of N2-acetyl-N6-[(9H-fluoren-9-ylmethyloxy)carbonyl]-D-alanyl]-L-lysine ((R)-2).
Figure S4. $^1$H (top) and $^{13}$C (bottom) NMR spectra (CD$_3$OD) of N2-acetyl-N6-[N-((9H-fluoren-9-ylmethyloxy)carbonyl)-L-$^{13}$C$_3$$^{15}$N$_1$]-alanyl]-L-lysine ((S)-[$^{13}$C$_3$$^{15}$N$_1$]-2).
Figure S5. $^1$H (top) and $^{13}$C (bottom) NMR spectra (DMSO-$d_6$) of methyl N2-acetyl-N6-[N-(6,9-dihydro-6-oxo-9-[2-deoxy-$\beta$-D-2-erythro-pentofuranosyl]-1H-purin-2-yl)-L-alanyl]-L-lysine ((S)-4).
Figure S6. $^1$H (top) and $^{13}$C (bottom) NMR spectra (DMSO-$d_6$) of methyl $N_2$-acetyl-$N_6$-[$(6,9$-dihydro-6-oxo-9-[2-deoxy-$\beta$-D-2-erythro-pentofuranosyl]-1H-purin-2-yl)-D-alanyl]-L-lysine ((R)-4).
Figure S7. $^1$H (top) and $^{13}$C (bottom) NMR spectra (DMSO-$d_6$) of methyl N2-acetyl-N6-[N-(6,9-dihydro-6-oxo-9-[2-deoxy-$\beta$-D-2-erythro-pentofuranosyl]-1H-purin-2-yl)-L-$[^{13}$C$_3$$^{15}$N$_1$]-alanyl]-L-lysine ($($S$)$-$[^{13}$C$_3$$^{15}$N$_1$]$)-4).
Figure S8. $^1$H NMR (top), $^{13}$C (bottom), and DEPT-135 (middle) spectra (DMSO-$d_6$) of N2-acetyl-N6-[N-(6,9-dihydro-6-oxo-9-[2-deoxy-β-D-2-erythro-pentofuranosyl]-1H-purin-2-yl)-L-alanyl]-L-lysine ((S)-1).
Figure S9. COSY spectrum of $N_2$-acetyl-$N_6$-[N-(6,9-dihydro-6-oxo-9-[2-deoxy-$\beta$-D-2-erythro-pentofuranosyl]-1H-purin-2-yl)-L-alanyl]-L-lysine ((S)-1).
Figure S10. $^1$H (top) and $^{13}$C (bottom) NMR spectra of N2-acetyl-N6-[N-(6,9-dihydro-6-oxo-9-[2-deoxy-β-D-2-erythro-pentofuranosyl]-1H-purin-2-yl)-D-alanyl]-L-lysine ((R)-1).
Figure S11. $^1$H (top) and $^{13}$C (bottom) NMR spectra (DMSO-$d_6$) of $N^2$-acetyl-$N^6$-[N-(6,9-dihydro-6-oxo-9-[2-deoxy-$\beta$-D-2-erythro-pentofuranosyl]-1H-purin-2-yl)-L-$[^{13}$C$_3$ $^{15}$N$_1$]-alanyl]-L-lysine ((S)-$[^{13}$C$_3$ $^{15}$N$_1$]-1).
Figure S12. A. $^{13}$C NMR spectrum (DMSO-$d_6$) of $N_2$-acetyl-$N_6$-[N-(6,9-dihydro-6-oxo-9-[2-deoxy-$\beta$-D-2-erythro-pentofuranosyl]-1H-purin-2-yl)-$L$-$^{13}$C$_3^{15}$N$_1$]-alanyl-$L$-lysine ((S)-$^{13}$C$_3^{15}$N$_1$)-1. B. Expanded region between 47-55 ppm of the $^{13}$C–$^{15}$N-decoupled (top) vs. $^{13}$C NMR (bottom) spectra.
Figure S13. $^1$H NMR spectra of (R)- and (S)-CE-dGuo adducts (D$_2$O) synthesized from the reaction dGuo and glyceraldehyde as previously described. (Liebigs Ann. Chem., 1994, 851-853).
Figure S14. A. $^1$H spectra of synthetic standard ($R$)-CE-dGuo adduct (bottom) vs. ($R$)-CE-dGuo adduct (top) isolated from reaction of dGuo and methylglyoxal in deuterated buffer. B. Expanded region between 3.63-4.60 ppm of the synthetic ($R$)-CE-dGuo standard (bottom) vs. ($R$)-CE-dGuo adduct (top) isolated from deuterated buffer.
Figure S15. A. $^1$H spectra of synthetic cross-link standard (R)-1 (bottom) and the (R)-1 cross-link isolated from reaction of dGuo, N$\alpha$-acetyllysine and methylglyoxal in deuterated buffer (top). B. Expanded region between 3.40-4.45 ppm of the synthetic standard (R)-1 (bottom) vs. cross-link (R)-1 (top) isolated in deuterated buffer.
Figure S16. $^1$H NMR spectra of both diastereomers of 1,N$^2$-cyclic dGuo adduct (MG-dGuo). The stereochemistry of the vicinal diol is unknown.
Figure S17. LC-ESI-MSn analysis of CE-dGuo.
Figure S18. Product ion spectra of cross-link (R)-1 (bottom) and (S)-1 (middle) and $^{13}$C$_3$^{15}$N$-(S)-1 (top).
Figure S18 (con’t). UPLC-ESI-MS-SRM (-116 Da) chromatogram and ms³ product ion spectrum of cross-links (R)-1 and (S)-1.
**Figure S19.** Extinction coefficient ($\varepsilon$, 1:1 MeOH:H2O) for crosslink (S)-1 (top right) and calibration curve for cross-link 1 vs. $[^{13}\text{C}_3^{15}\text{N}_1]$-(S)-1 (bottom right).
Figure S20. Reconstructed UPLC-ESI/MS ion chromatograph from the reaction of methylglyoxal, dGuo, and AcAVAGKAGAR (1:1:1) at pH 7.4 and 37 °C.
Figure S20 (con’t). Reconstructed UPLC-ESI/MS ion chromatograph from the reaction of methylglyoxal, dGuo, and AcAVAGKAGAR (1:1:1) at pH 7.4 and 37 °C.
Ac-AVAGKAGAR

\[ [M+H]^+ = 842.47 \]
\[ [M+2H]^{2+} = 421.74 \]

Figure S21. CID mass spectrum of the Ac-AVAGKAGAR peptide (m/z = 421.8 [M+2H]^{2+}).
Figure S22. CID mass spectrum of the cross-link 11 (*m/z* = 560 [M+2H]$^{+2}$).
Figure S23. CID mass spectrum of the cross-link 12 (m/z = 1101.3 [M+H]+1).
Figure S24. CID mass spectrum of the arginine adduct 13 \((m/z = 457.8 \text{ [M+2H]}^2)\).
Figure S25. CID mass spectrum of the arginine adduct 14 ($m/z = 448.8$ [M+2H]$^+$).
Figure S26. CID mass spectrum of the arginine + CE-Lys adduct 15 ($m/z = 493.8 \rightarrow 484.8 \rightarrow 462.7 \ [M+2H]^{2+}$).
Ac-AVAGKAGAR (R+84)

\[ [\text{M+H}]^{+1} = 926.47 \]
\[ [\text{M+2H}]^{+2} = 463.74 \]

Figure S27. CID Mass spectrum of the unknown Arginine + 84 adduct ([M+2H]^{+2} = 463.74)