Developing New Isotope-coded MS-cleavable Cross-linkers for Elucidating Protein Structures

Clinton Yu[1], Wynne Kandur[2], Athit Kao[1], Scott Rychnovsky[2], Lan Huang[1]*

[1]Department of Physiology and Biophysics, University of California, Irvine, CA 92697

[2]Department of Chemistry, University of California, Irvine, CA 92697

Supplemental Materials.

Characterization of Synthesized Chemicals. ^{1}H NMR and ^{13}C NMR spectra were recorded at ambient temperature at 500 MHz and 125 MHz, respectively, on a Bruker DRX500 NMR instrument. ^{1}H and ^{13}C NMR data are reported as follows: chemical shifts are reported in ppm on a δ scale and referenced to internal tetramethylsilane or residual solvent (TMS: δ 0.00; $CHCl_3$: δ 7.27), multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, qu = quintet, m = multiplet), coupling constants (Hz), and integration. Infrared (IR) spectra were obtained using a FT-IR spectrometer. Accurate mass spectra were acquired on a Waters LCT Premier quadrupole time-of-flight spectrometer and were obtained by peak matching. Gas chromatography/mass spectrometry (GC/MS) was performed with a Thermo-Finnigan Trace Mass Spectrometer Plus quadrupole system with a fused silica capillary column (30 m ´ 0.32 mm ´ 0.25 mm) wall-coated with DB-5 (J & W Scientific) using electron ionization (70 eV) or a Waters GCT Premier orthogonal acceleration time-of-flight spectrometer using chemical ionization. Analytical thin layer chromatography was performed on EMD Chemicals Inc. silica gel 60 F_{254} plates. Liquid

chromatography was performed using forced flow (flash chromatography) of the indicated solvent system on Sorbent Technologies silica gel (SiO_2) 60 (230–400 mesh). Unless otherwise noted, all reactions were carried out under an atmosphere of argon in flame-dried glassware. Solvents were distilled from CaH_2 or filtered through alumina before use.[1]

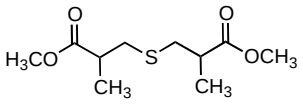
Experimental Procedures for the synthesis of d_0- and d_{10}- DMDSSO

General Procedure 1: Diester Hydrolysis. The starting diester was dissolved in 2:1 or 4:1 THF:H_2O, depending on solubility. The mixture was cooled to 0 °C and lithium hydroxide (98%, 5 equiv.) dissolved in minimal H_2O was added slowly. The reaction was monitored by thin layer chromatography, while stirring vigorously. Upon completion, the mixture was diluted with diethyl ether and water. The aqueous layer was acidified with HCl to pH 1 (monitored by pH paper). The layers were separated and the aqueous layer was extracted with diethyl ether (3x). The organic layers were combined, washed with brine, dried over Na_2SO_4, filtered and evaporated *in vacuo*.

General Procedure 2: Di-NHS Ester Synthesis. The crude diacid (0.085 mmol) was dissolved in dry DMF (0.5 ml) in a flame dried round bottom flask under argon. NHS (0.34 mmol) and pyridine (0.68 mmol) were added before a slow addition of TFAA (0.34 mmol) at 0° C. The mixture was left to warm slowly overnight while stirring under argon. The mixture was diluted with dichloromethane and 1M HCl. After the layers were separated, the organic layer was washed with 1 M HCl (2x) and dilute $NaHCO_3$ solution

(2x). The organic layer was dried with Na_2SO_4, filtered, and evaporated *in vacuo*. This procedure is adapted from a published procedure.[2]

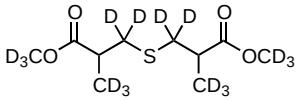
General Procedure 2: Oxidation to Sulfoxide. The di-NHS ester (0.026 mmol) was dissolved in $CDCl_3$ (1 ml) and *m*-CPBA (77% w/w, 0.026 mmol) was added slowly while monitoring by LRMS ESI + and ^1H NMR. When this reaction was run on larger scale than 30 mg starting material, the solution was cooled in an ice bath prior to *m*-CPBA addition. Once the reaction was complete, the solution was diluted with dichloromethane (2 ml) and washed with 10% sodium bicarbonate solution (3 x 2 ml). The organic layer was dried over Na_2SO_4, filtered, and evaporated *in vacuo* to afford the product.



3-22

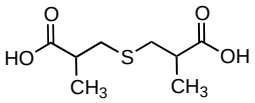
Diester 3-22. Methylmethacrylate (0.57 mL, 5.3 mmol) and thioacetic acid (0.47 mL, 6.8 mmol) were combined in a flame-dried vial under argon, which was sealed and heated to 80 °C for 6 h. The vial was allowed to cool to rt and the mixture was evaporated to remove any excess thioacetic acid. To the crude reaction mixture was added methanol (135 mL) and triethylamine (4.2 mL, 30 mmol), after which methyl methacrylate (1.9 mL, 18 mmol) was added dropwise. The reaction mixture was let stir for 24 h, after which time the solution was evaporated to remove methanol and triethylamine. The crude mixture was taken on to the next step assuming a quantitative yield. ^1H NMR (500 MHz, $CDCl_3$) δ 3.71 (s, 6H), 2.87–2.82 (m, 2H), 2.68–2.63 (m, 2H), 2.60–2.55 (m, 2H), 1.28–1.24 (m, 6H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 175.6, 51.9, 40.2, 36.0, 16.8; IR (thin film)

2976, 2953, 2936, 1730 cm^{-1}; Accurate Mass ES+ *m / z* calcd for $C_{10}H_{18}O_4SNa$ $(M+Na)^+$ 257.0823, found 257.0820.



3-25

Deuterated diester 3-25.* A similar procedure to preparation of **3-22** was performed starting with d_8-methylmethacrylate (1.00 g, 9.25 mmol) and thioacetic acid (0.720 mL, 10.2 mmol). The second step in the one-pot procedure was conducted with d_8-methylmethacrylate (1.00 g, 9.25 mmol), methanol (3.25 mL), and triethylamine (3.85 mL, 27.8 mmol). After evaporation, the crude mixture was taken on to the next step assuming a quantitative yield. ^{1}H NMR (500 MHz, $CDCl_3$) δ 2.64 (s, 2H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 175.6, 40.0; Accurate Mass ES+ *m / z* calcd for $C_{10}H_2D_{16}O_4SNa$ $(M+Na)^+$ 273.1827, found 273.1819.

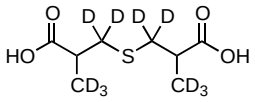


S3-1

d_0-Diacid S3-1. Diester **3-22** was subjected to general procedure 1 in 1:2 H_2O/THF. The mixture was stirred overnight at rt before workup to yield a white solid (0.83 g, 75% over two steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 12.27 (br s, 2H), 2.84 (dd, *J* = 9.5, 12.2, 2H), 2.79–2.77 (m, 2H), 2.60 (dd, *J* = 4.4, 12.2, 2H), 1.28 (app d, *J* = 6.7, 6H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 182.0, 40.1, 34.6, 17.5; IR (thin film) 2920, 2354, 1724, 1697, 1449 cm^{-1}; Accurate Mass ES- *m / z* calcd for $C_8H_{14}O_4SNa$ $(M+Na)^+$ 229.0511, found 229.0507.

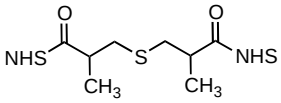
---

* The synthesis of **3-22** is an improved procedure that uses more methanol in the second step than the older procedure, which is described here for the synthesis of **3-25**. We recommend the stoichiometry of the **3-22** synthesis for both substrates.



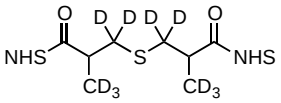
S3-2

d_{10}-Diacid S3-2. Crude diester **3-25** was subjected to general procedure 1 in 1:4 H_2O/THF. The mixture was stirred vigorously overnight at rt before workup to yield a white solid (1.72 g, 86% over two steps). ^{1}H NMR (500 MHz, $CDCl_3$) δ 11.32 (br s, 2H), 2.72 (s, 1H), 2.71 (s, 1H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 181.9, 181.7, 40.0, 39.9; IR (thin film) 2919, 2544, 2229, 1696, 1421 cm^{-1}; Accurate Mass ES- *m / z* calcd for $C_8H_2D_{10}O_4SNa$ $(M-H)^-$ 215.1162, found 215.1171.



S3-3

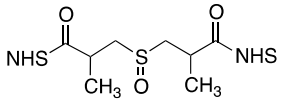
d_0-Di-NHS ester S3-3. d_0-Diacid (0.020 g, 0.085 mmol) was subjected to general procedure 2. The mixture was stirred overnight at rt before workup to yield a colorless oil (0.033 g, 97%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 3.06–2.96 (m, 4H), 2.83 (br s, 8H), 2.78–2.72 (m, 2H), 1.43 (s, 3H), 1.42 (s, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 170.5, 170.5, 169.2, 169.1, 38.4, 38.3, 35.7, 35.7, 25.8, 16.8, 16.6; IR (thin film) 2918, 1811, 1781, 1741, 1070 cm^{-1}; Accurate Mass ES+ *m / z* calcd for $C_{16}H_{20}N_2O_8SNa$ $(M+Na)^+$ 423.0838, found 423.0834.



S3-4

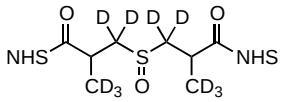
d_{10}-Di-NHS ester S3-4. d_{10}-Diacid (0.020 g, 0.092 mmol) was subjected to general procedure 2. The mixture was stirred overnight at rt before workup to yield a colorless oil (0.032 g, 85%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 2.99 (s, 1H), 2.98 (s, 1H), 2.83 (br s, 8H);

^{13}C NMR (125 MHz, $CDCl_3$) δ 170.5, 169.2, 38.01, 37.97, 25.8; IR (thin film) 2924, 1812, 1782, 1737, 1065 cm^{-1}; Accurate Mass ES+ *m / z* calcd for $C_{16}H_{10}D_{10}N_2O_8SNa$ $(M+Na)^+$ 433.1465, found 433.1455.



3-17

d_0-Sulfoxide 3-17. d_0-Di-NHS ester (0.18 g, 0.45 mmol) was subjected to general procedure 3, which yielded desired product (0.100 g, 54%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 3.56–3.41 (m, 2H), 3.35–3.27 (m, 1.4H), 3.22 (dd, *J* = 9.2, 13.3, 0.6H), 2.90 (dd, *J* = 9.0, 13.2, 1H), 2.85 (s, 8H), 1.61–1.56 (m, 4H), 1.55 (s, 3H); IR (thin film) 2849, 1812, 1776, 1710, 1201 cm^{-1}; Accurate Mass ES+ *m / z* calcd for $C_{16}H_{20}N_2O_9SNa$ $(M+Na)^+$ 439.0787, found 439.0777.



3-18

d_{10}-Sulfoxide 3-18. d_{10}-Di-NHS ester (0.22 g, 0.54 mmol) was subjected to general procedure 3, which yielded desired product (0.20 g, 89%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 3.50–3.39 (m, 2H), 2.84 (s, 8H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 168.9, 33.0, 32.9, 31.9, 31.2, 29.9, 25.8; Accurate Mass ES+ *m / z* calcd for $C_{16}H_{10}D_{10}N_2O_9SNa$ $(M+Na)^+$ 449.1415, found 449.1412.

References

(1) Pangborn, A. B.; Giardello, M. A.; Grubbs, R. H.; Rosen, R. K.; Timmers, F. J. *Organometallics* **1996**, *15*, 1518–1520.

(2) Leonard, N. M.; Brunckova, J. *J. Org. Chem.* **2011**, *76*, 9169–9174.

Supplemental Figure Legends

Supplementary Figure 1. (A) Chemical structure of d_0-DSSO (d_0-Disuccinimidyl Sulfoxide) and d_4-DSSO (d_4-Disuccinimidyl Sulfoxide). (B) Proposed chemical synthesis scheme for d_4-DSSO.

Supplemental Figure 2. Identification of d_0- and d_{10}-DMDSSO inter-linked Ac-Myelin peptides by MS3 sequencing. (A-B) MS3 analysis of α_A and α_T ions observed in MS2 spectrum of the sextuply charged d_0-inter-linked Ac-myelin (i.e. $[\alpha\text{-}\alpha]^{6+}$) as shown in Figure 3B: (A) MS3 spectrum of α_A (m/z 454.57^{3+}). (B) MS spectrum of α_T (m/z 465.23^{3+}). (C-D) MS3 analysis of α_{A*} and α_{T*} ions observed in MS2 spectrum of a sextuply charged d_{10}-inter-linked Ac-myelin (i.e. $[\alpha\text{-}\alpha]^{6+}$) as shown in Figure 3F: (C) MS3 spectrum of α_{A*} (m/z 456.25^{3+}). (D) MS3 spectrum of α_{T*} (m/z 466.91^{3+}).

Supplementary Figure 3. SDS-PAGE electrophoresis separation of cytochrome C products cross-linked by DSSO, d_0-DMDSSO, and d_{10}-DMDSSO respectively.

Supplemental Figure 4. MS2 analysis of d_0/d_{10}-DMDSSO intra-linked and dead-end modified cytochrome C peptides. MS2 spectra of (E) d_0-α_{intra} (m/z 621.3203^{3+}), (F) d_{10}-α_{intra}(m/z 624.6746^{3+}), (G) d_0-α_{DN} (m/z 546.6116^{3+}) and (H) d_{10}-α_{DN} (m/z 549.9661^{3+}). MS3 spectra of the observed MS2 fragments are displayed in Supplemental Figure 7.

Supplementary Figure 5. MS spectra of representative inter-linked, intra-linked, and dead-end modified peptides from cytochrome C, (A-C) when respective cross-linked digests were mixed in 1:1. (D-F) when cytochrome C was cross-linked with 1:1 mixture of d_0/d_{10}-DMDSSO.

Supplementary Figure 6. Cytochrome C lysine residues clustered into 8 'groups' (gray) based on proximity to one another. Gray lines (7) represent regional inter-links identified in both DSSO cross-linking and DMDSSO cross-linking experiments, while red lines (5) represent newly identified regional inter-links presented in DMDSSO studies.

Supplemental Figure 7. Detailed MS3 sequencing of d_0/d_{10} DMDSSO Inter-linked Ac-Myeline and cross-linked Cytochrome C peptides including a selected intra-linked, a dead-end and all inter-linked peptides reported in supplemental Table 1.