Supplementary Data

Arginine side-chain modification that occurs during copper-catalysed azide-alkyne click reactions resembles an advanced glycation end product

Anne C. Conibear[1], Karine Farbiarz[1], Rupert L. Mayer[2,3], Maria Matveenko[1], Hanspeter Kählig[4] and Christian F. W. Becker[1*]

[1] University of Vienna, Faculty of Chemistry, Institute of Biological Chemistry, Währinger Straße 38, 1090 Vienna, Austria.

[2] University of Vienna, Faculty of Chemistry, Institute of Analytical Chemistry, Währinger Straße 38, 1090 Vienna, Austria.

[3] University of Vienna, Faculty of Chemistry, Mass Spectrometry Centre, Währinger Straße 38, 1090 Vienna, Austria.

[4] University of Vienna, Faculty of Chemistry, Institute of Organic Chemistry, Währinger Straße 38, 1090 Vienna, Austria.

[*]Corresponding author:

Christian F. W. Becker

Institute of Biological Chemistry, University of Vienna, Währinger Straße 38, 1090 Vienna, Austria

E-mail: christian.becker@univie.ac.at

Phone: + 43-1-4277-70510

Fax: + 43-1-4277-9705

Synthesis of peptides P1-8

Peptides P1-P8 were synthesized by manual solid phase peptide synthesis (SPPS) on Rink Amide MBHA (P1, P2 and P3), Wang-polystyrene (P5, P6 and P7) or *bis*(2-sulfanylethyl)amido (SEA)-linked (P4) resin using fluorenylmethoxycarbonyl (Fmoc)-protected amino acids and standard side-chain protecting groups [Arg(Pbf), Asn(Trt), Asp(*t*Bu), Cys(Trt), Gln(Trt), Glu(*t*Bu), His(Trt), Lys(Boc), Ser(*t*Bu), Thr(*t*Bu), Trp(Boc) and Tyr(*t*Bu)]. Azide functionalities were introduced using Lys(N_3) and alkyne functionalities with propargylglycine. Peptides P1, P2 and P3 had a polyethylene glycol (PEG_3 or PEG_4) linker between the peptide and the azide moiety. In brief, amino acid residues (2.5 eq.) were coupled with 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU, 2.4 eq., 0.5 M in DMF) and diisopropylethylamine (DIPEA, 5 eq.) and then deprotected with piperidine (20% v/v in DMF). After the final coupling and deprotection, the resin was washed with DCM, dried under vacuum and the peptide was cleaved with TFA/triisopropylsilane/dimethylsulfide/H_2O (92.5 : 2.5 : 5 : 2.5 v/v/v/v) for 2-3 h. The cleaved peptide was precipitated with ice-cold diethylether and the pellet dissolved in water/acetonitrile (ACN, 50:50 containing 0.1% TFA) before lyophilisation. Crude peptides were purified by RP-HPLC using a Kromasil C4 column and a 1%/min gradient of ACN 5-45% at 20 mL/min and the fractions were analysed by electrospray ionisation mass spectrometry (ESI-MS).

Analytical methods

Liquid chromatography-mass spectrometry (LC-MS) was used to monitor formation of the DHA-arginine adduct and was performed on a Waters AutoPurification HPLC/MS system equipped with a 3100 Mass Detector, 2545 Binary Gradient Module, 2767 Sample Manager, and 2489 UV/visible detector. Separation was achieved with a Kromasil 300-5-C4 or 300-5-C18

column (500 × 4.6 mm, 5 µm particle size) using a flow rate of 1 mL/min, 5-65% linear gradient of ACN + 0.05% TFA in ddH_2O + 0.05% TFA over 10 min. Mass spectra were acquired by electrospray ionization (ESI-MS) in positive ion mode. Analytical RP-HPLC was performed on a Dionex Ultimate 3000 instrument with Kromasil 300-5-C4 (150 × 4.6 mm, 5 µm particle size) or Kromasil 300-5-C18 (150 × 4.6 mm, 5 µm particle size) column with a linear 5-65% gradient of ACN + 0.08% TFA in ddH_2O + 0.1% TFA over 30 min at 1 mL/min, monitoring absorption at 214 and 280 nm.

Standard CuAAC reaction conditions

Oxygen was removed from solvents by bubbling with argon and all stock solutions were prepared freshly before the reaction. Stock solutions of $CuSO_4$ (200 mM in ddH_2O), TBTA (100 mM in anhydrous DMF) and sodium ascorbate (1500 mM in ddH_2O) were prepared and the peptide(s) were dissolved in a 2 mL Eppendorf tube flushed with argon. In a separate argon-flushed vessel, the $CuSO_4$ was first combined with the ligand TBTA and mixed and then the sodium ascorbate was added to reduce the Cu(II) to Cu(I). This mixture was then added to the peptide solution and the reaction mixture was stirred at room temperature under argon. Approximate reagent concentrations in the final reaction mixture: peptide alkyne 3 mM (1 eq.), peptide azide, N^{α}-Fmoc-L-arginine, N^{α}-Boc-L-arginine, N^{α}-Boc-L-lysine or N^{α}-Boc-L-glutamine 8 mM (2.5 eq.), $CuSO_4$ 19 mM (6 eq.), sodium ascorbate 120 mM (40 eq.), TBTA 21 mM (6.5 eq.) and DMF : H_2O ~ 4 : 1 v/v. For CuAAC ligations between P2 and P8 to determine the effects of ascorbate concentration and aminoguanidine on DHA-arginine adduct formation, final reaction concentrations were peptide alkyne P8 3 mM (1 eq.), peptide azide P2 3 mM (1 eq.), $CuSO_4$ 16 mM (6 eq.), sodium ascorbate 107 mM (40 eq.) or 27 mM (10 eq.), TBTA 17 mM (6.5 eq.), aminoguanidine 0 mM (0 eq.) or 27 mM (10 eq.) and DMF : H_2O ~ 4 : 1 v/v. A1iquots

for LC-MS analysis were diluted with ACN/H2O 50 : 50 + 0.1 % TFA and separated on a C18 column.

MS/MS analysis of modified peptide P1*

Modified peptide P1* was dissolved in water (~0.1 mg/mL) and analysed by LC-MS/MS on an UltiMate 3000 RSLC nano UPLC system coupled via an ESI interface to an LTQ Velos Orbitrap mass spectrometer (both Thermo Scientific, Germany, Bremen). The peptide (5 µL) was trapped on a C4 PepMap300 µ-precolumn (Thermo Scientific), washed with loading eluent (97.95% H_2O, 2% ACN, 0.05% TFA) and then eluted with a 30 min gradient from 2% to 64% ACN by mixing the respective eluents A (97.9% H_2O, 2% ACN and 0.1% FA) and B (79.9% ACN, 20% H_2O and 0.1% FA). As analytical column a 15cm C4 Accucore column (Thermo Scientific) was utilized at a flow rate of 300 nL/min. The m/z range for the MS method was set as 400 to 1400 for MS1 at a resolution of 60000 and the six most abundant signals were selected for fragmentation. Collision-induced dissociation fragmentation was carried out in the linear ion trap with 35% normalized collision energy.

MS raw data were subjected to an automated spectra search by Proteome Discoverer 1.4 (Thermo Scientific, Germany, Bremen) against the peptide sequence of P1* with 10 ppm mass tolerance at the MS1 level and 0.8 Da fragment mass tolerance at the MS2 level. Sequest was chosen as the search engine with a DHA modification allowed on arginine as variable modification as well as modification with PEG_3-K(N_3) at the C-terminus. Target false discovery rates were set to 0.05 as relaxed and 0.01 as strict parameters, respectively. Positive hits were filtered for high confidence peptides only and the MS2 match tolerance refined for 0.4 Da. The observed b- and y- ions, including neutral losses, are shown in Table S1. Peaks for b_{12}^{2+} and b_{12}-

H_2O^{2+} were assigned manually due to limitations of the software for detecting the PEG_3-K(N_3) fragment.

Table S1. b- and y-ion series for LC-MS/MS sequencing of modified peptide P1*.

MS2 fragment b-ion series	m/z	delta mass (Da)	Peptide sequence	MS2 fragment y-ion series	m/z	delta mass (Da)
b_1^+	fragment not found		**S**			
b_2^+	274.12	-0.09	**W**	y_{12}-NH_3^{3+}	599.29	0.22
				y_{12}-NH_3^{2+}	898.43	-0.09
b_2-H_2O^+	256.11	-0.12		y_{12}-H_2O^{3+}	598.97	-0.11
				y_{12}-H_2O^{2+}	897.94	0.24
				y_{12}^{3+}	604.97	-0.04
b_3^+	387.20	-0.09	**L**	y_{11}-NH_3^{2+}	805.40	-0.23
b_3-H_2O^+	369.19	-0.32		y_{11}^{2+}	813.91	-0.24
b_4^+	458.24	-0.04	**A**	y_{10}-NH_3^{2+}	748.85	-0.21
b_4-H_2O^+	440.23	-0.14		y_{10}^{2+}	757.37	-0.23
b_5^+	621.30	-0.09	**Y**	y_9-NH_3^{2+}	713.34	0.08
b_5-H_2O^+	603.29	0.11		y_9^{2+}	721.85	-0.25
b_6^+	718.36	-0.17	**P**	y_8^+	fragment not found	
b_7^+	775.38	-0.09	**G**	y_7-NH_3^{2+}	583.28	-0.22
b_7^{2+}	388.19	-0.20		y_7-H_2O^{2+}	582.78	0.04
b_7-H_2O^+	757.37	-0.23				
b_8^+	846.41	-0.07	**A**	y_6-NH_3^{2+}	554.76	-0.27
b_8^{2+}	423.71	-0.27		y_6-H_2O^{3+}	369.85	0.34
b_8-H_2O^+	828.40	0.25				
b_9^+	945.48	-0.03	**V**	y_5-NH_3^{2+}	519.25	0.21
				y_5-NH_3^+	1037.49	-0.05
b_9^{2+}	473.24	-0.18		y_5-H_2O^{2+}	518.76	-0.28
				y_5^{2+}	527.76	-0.07
b_{10}^+	1032.51	0.02	**S**	y_4-NH_3^{2+}	469.71	-0.01
b_{10}^{2+}	516.76	-0.19		y_4-NH_3^+	938.42	-0.08
b_{10}-H_2O^{2+}	507.76	-0.19		y_4^{2+}	478.23	-0.23
				y_4^+	955.45	-0.09

b_{11}^{2+}	598.29	0.02	**Y**	$y_3\text{-}NH_3^{2+}$	426.20	-0.35
				$y_3\text{-}NH_3^{+}$	851.39	-0.00
$b_{11}\text{-}H_2O^{2+}$	589.29	-0.35		y_3^{2+}	434.71	-0.01
				y_3+	868.42	-0.02
b_{12}^{2+}	763.48	-0.37	**R(DHA)**	$y_2\text{-}NH_3^{+}$	688.32	0.03
$b_{12}\text{-}H_2O^{2+}$	754.78	-0.06		y_2^{+}	705.35	-0.14
			PEG_3-K(N_3)	y_1^{+}	fragment not found	

Modification of peptide P1 with ascorbate and DHA

Peptide P1 (0.5 mg, 300 nmol) samples were treated for 60 min with four different reagent combinations to determine whether copper is necessary for modification of the peptide by ascorbate or DHA: 1) Standard CuAAC reagents as described above; 2) DHA (120 mM, 40 eq.); 3) Standard CuAAC reagents in which the sodium ascorbate was replaced with DHA (120 mM, 40 eq.); and 4) Sodium ascorbate (120 mM, 40 eq.). All reaction mixtures were dissolved in DMF : H_2O ~ 4 : 1 v/v and were flushed with argon and stirred for 60 min. A1iquots for LC-MS analysis were diluted with ACN/H2O 50 : 50 + 0.1 % TFA and separated on a C4 column. As shown in Figure S1 below, peptide P1 was modified to P1* in the presence of standard CuAAC reagents. No modification was observed on P1 if the ascorbate was replaced with DHA, either in the presence or absence of copper and TBTA. Ascorbate alone, however, resulted in modification of P1 to P1*, although to a lesser extent than in the presence of copper and TBTA.

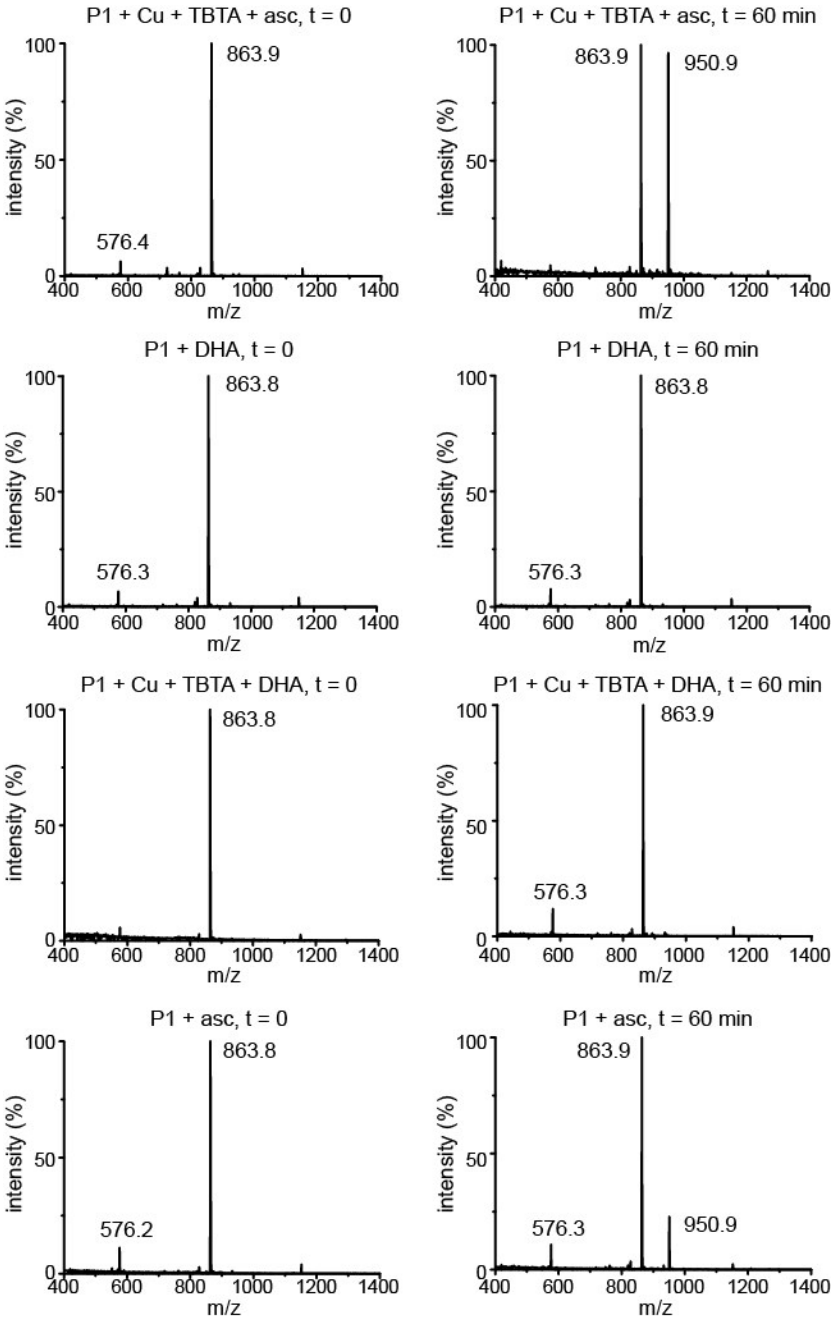

Figure S1: Mass spectra of peptide P1 treated with CuAAC reagents, DHA, CuAAC reagents and DHA, or ascorbate for 60 min. **P1**: MW 1725.9, $[M+2H]^{2+}_{calc}$ 863.9, $[M+2H]^{2+}_{obs}$ 863.9, $[M+3H]^{3+}_{calc}$ 576.3, $[M+3H]^{3+}_{obs}$ 576.3; **P1***: MW 1899.9, $[M+2H]^{2+}_{calc}$ 950.9, $[M+2H]^{2+}_{obs}$ 950.9.

Formation of DHA-arginine adduct on peptides P1-P7

Peptides P1-P7 (7 mM) were incubated with CuAAC reagents as described above for 60 min. Aliquots were analysed by LC-MS and the mass spectra of the peptide peaks at t = 60 min are shown in figure S2 below.

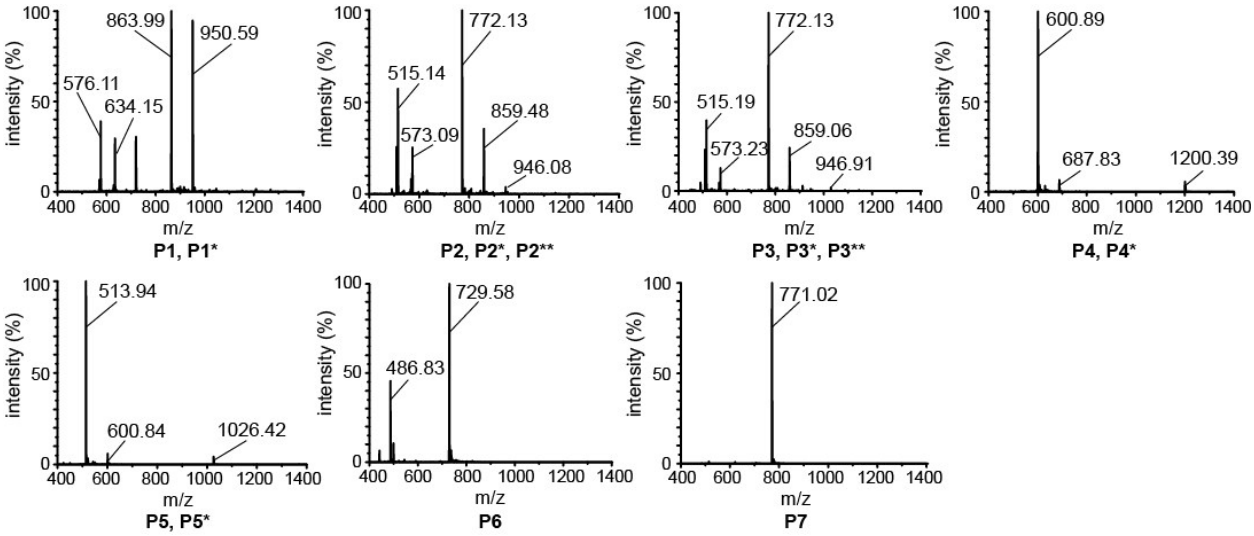

Figure S2: Mass spectra of short peptides treated with CuAAC reagents for 60 min. **P1**: MW 1725.94, $[M+2H]^{2+}_{calc}$ 863.97, $[M+2H]^{2+}_{obs}$ 863.99, $[M+3H]^{3+}_{calc}$ 576.31, $[M+3H]^{3+}_{obs}$ 576.11; **P1***: MW 1899.94, $[M+2H]^{2+}_{calc}$ 950.97, $[M+2H]^{2+}_{obs}$ 950.59, $[M+3H]^{3+}_{calc}$ 634.31, $[M+3H]^{3+}_{obs}$ 634.15; **P2**: MW 1542.78, $[M+2H]^{2+}_{calc}$ 772.41, $[M+2H]^{2+}_{obs}$ 772.13, $[M+3H]^{3+}_{calc}$ 515.27, $[M+3H]^{3+}_{obs}$ 515.14; **P2***: MW 1716.89, $[M+2H]^{2+}_{calc}$ 859.39, $[M+2H]^{2+}_{obs}$ 859.48, $[M+3H]^{3+}_{calc}$ 573.30, $[M+3H]^{3+}_{obs}$ 573.09; **P2****: MW 1890.00, $[M+2H]^{2+}_{calc}$ 946.39, $[M+2H]^{2+}_{obs}$ 946.08; **P3**: MW 1542.78, $[M+2H]^{2+}_{calc}$ 772.41, $[M+2H]^{2+}_{obs}$ 772.13, $[M+3H]^{3+}_{calc}$ 515.27, $[M+3H]^{3+}_{obs}$ 515.19; **P3***: MW 1716.89, $[M+2H]^{2+}_{calc}$ 859.39, $[M+2H]^{2+}_{obs}$ 859.06; $[M+3H]^{3+}_{calc}$ 573.30, $[M+3H]^{3+}_{obs}$ 573.23; **P3****: MW 1890.00, $[M+2H]^{2+}_{calc}$ 946.39, $[M+2H]^{2+}_{obs}$ 946.91; **P4**: MW 1200.40, $[M+H]^{+}_{calc}$ 1201.48, $[M+H]^{+}_{obs}$ 1200.39, $[M+2H]^{2+}_{calc}$ 601.20, $[M+2H]^{2+}_{obs}$ 600.89; **P4***: MW 1374.51, $[M+2H]^{2+}_{calc}$ 688.26, $[M+2H]^{2+}_{obs}$ 687.83; **P5**: MW 1026.17, $[M+H]^{+}_{calc}$

1027.56, $[M+H]^{+}_{obs}$ 1026.42, $[M+2H]^{2+}_{calc}$ 514.19, $[M+2H]^{2+}_{obs}$ 513.94; **P5***: MW 1200.28, $[M+2H]^{2+}_{calc}$ 601.14, $[M+2H]^{2+}_{obs}$ 600.84; **P6**: MW 1457.72, $[M+2H]^{2+}_{calc}$ 729.86, $[M+2H]^{2+}_{obs}$ 729.58, $[M+3H]^{3+}_{calc}$ 486.91, $[M+3H]^{3+}_{obs}$ 486.83; **P7**: MW 1540.68, $[M+2H]^{2+}_{calc}$ 771.34, $[M+2H]^{2+}_{obs}$ 771.02.

Modification of Boc- and Fmoc-protected amino acids under CuAAC conditions

Treatment of N$^{\alpha}$-Fmoc-L-arginine, N$^{\alpha}$-Boc-L-arginine, N$^{\alpha}$-Boc-L-lysine and N$^{\alpha}$-Boc-L-glutamine (9 mM) with standard CuAAC reagents as described above was carried out for 3 days. Formation of the DHA-adduct was monitored by LC-MS and the mass spectra of the amino acid peaks before the addition of CuAAC reagents (t = 0) and after 3 days are shown in figure S3 below. Whereas a DHA adduct formed with N$^{\alpha}$-Boc-L-arginine (MW 448.43), no DHA adducts could be detected for N$^{\alpha}$-Boc-L-lysine (MW 420.42) or N$^{\alpha}$-Boc-L-glutamine (MW 420.37).

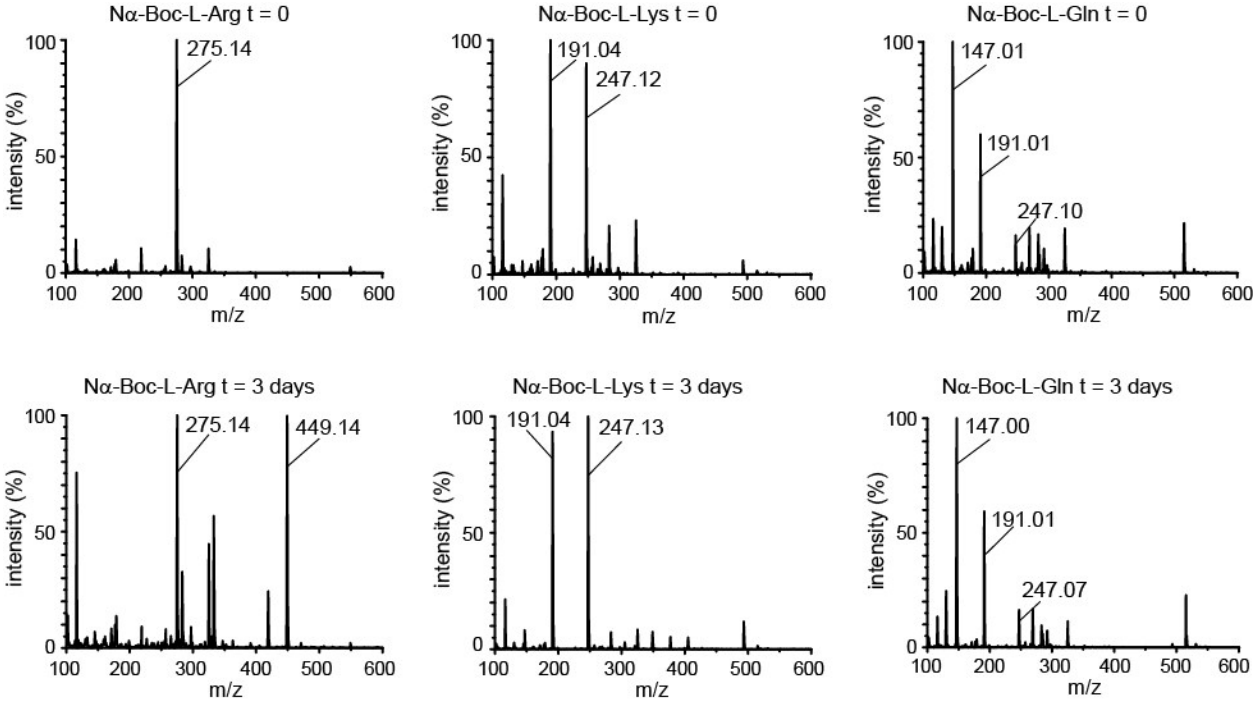

Figure S3: Mass spectra of N$^{\alpha}$-Boc-L-arginine, N$^{\alpha}$-Boc-L-lysine and N$^{\alpha}$-Boc-L-glutamine treated with CuAAC reagents for 3 days. N$^{\alpha}$-Boc-L-arginine: MW 274.32, $[M+H]^+_{calc}$ 275.40, $[M+H]^+_{obs}$ 275.14; N$^{\alpha}$-Boc-L-arginine*: MW 448.43, $[M+H]^+_{calc}$ 449.51, $[M+H]^+_{obs}$ 449.14; N$^{\alpha}$-Boc-L-lysine: MW 246.31, $[M+H]^+_{calc}$ 247.39, $[M+H]^+_{obs}$ 247.12 and 247.13, $[M-t\text{Bu}+H]^+_{calc}$ 191.28, $[M-t\text{Bu}+H]^+_{obs}$ 191.04; N$^{\alpha}$-Boc-L-glutamine: MW 246.26, $[M+H]^+_{calc}$ 247.34, $[M+H]^+_{obs}$ 247.10

and 247.07, $[M\text{-}t\text{Bu}+H]^+_{calc}$ 191.24, $[M\text{-}t\text{Bu}+H]^+_{obs}$ 191.01, $[M\text{-Boc}+H]^+_{calc}$ 147.23, $[M\text{-Boc}+H]^+_{obs}$ 147.01 and 147.00.

NMR characterisation of N^{α}-Fmoc-L-arginine*

All 1- and 2-dimensional NMR spectra were acquired for solutions in dimethylsulfoxide-d_6 at 298 K on a Bruker AV III HDX 700 MHz NMR spectrometer (Bruker BioSpin, Rheinstetten, Germany) using a 5 mm helium cooled quadruple resonance inverse cryoprobe (QCI-F: ^{1}H, ^{13}C, ^{15}N, ^{19}F) with z axis gradients and automatic tuning and matching accessory. The resonance frequency for ^{1}H NMR was 700.40 MHz, for ^{13}C NMR 176.12 MHz, and for ^{15}N NMR 70.97 MHz. Spectra were processed using Topspin software 3.2 (Bruker BioSpin) and referenced to the residual, non-deuterated solvent signal for ^{1}H (δ=2.50), the carbon signal of the solvent for ^{13}C (δ=39.52), and from the calibrated proton shifts via the ratio of the gyromagnetic ratios for ^{15}N, respectively. The chemical shifts of the assigned resonances of the major product (Fmoc-arginine-DHA structure 2) are shown in Table S2 and spectra are shown in figures S4-S10.

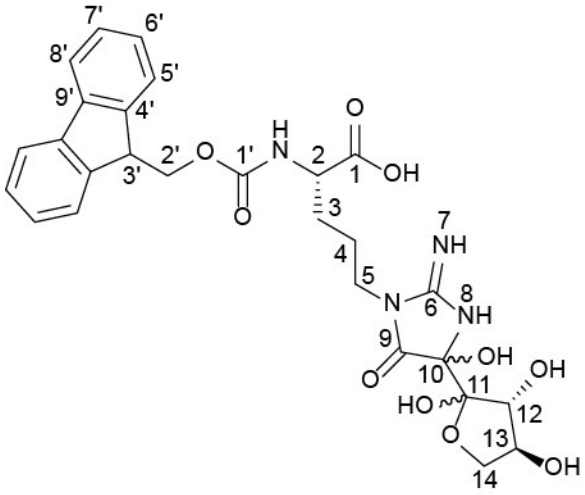

Table S2. NMR chemical shifts of Fmoc-arginine-DHA structure.

Resonance	^{1}H (ppm)	$J_{H,H}$ (Hz)	^{13}C (ppm)	^{15}N (ppm)
1-C			173.58	
1-OH	12.61	broad		
2-CH	3.94	m	53.63	
2-NH	7.67	d 8.0		88.3
3-CH_2	1.75	m	27.50	
	1.58			
4-CH_2	1.57	m	24.47	
5-CH_2	3.60	m	38.63	
6-C			157.20	
8-NH	10.44	s		107.0
9-C			170.91	
10-C			87.48	
10-OH	7.49	s		
11-C			101.35	
12-CH	4.34	d 6.8	75.49	
13-CH	4.06	d 7.1/d 7.5/d 6.8	74.06	
14-CH_2	3.82	d 7.1/d 8.6	69.35	
	3.17	d 7.5/d 8.6		
1′-C			156.17	
2′-CH_2	4.27	m	65.67	
3′-CH	4.23	m	46.63	
4′-C			143.79	
5′-CH	7.72	m	125.28	
6′-CH	7.33	m	127.12	
7′-CH	7.42	m	127.68	
8′-CH	7.90	m	120.15	
9′-C			140.73	

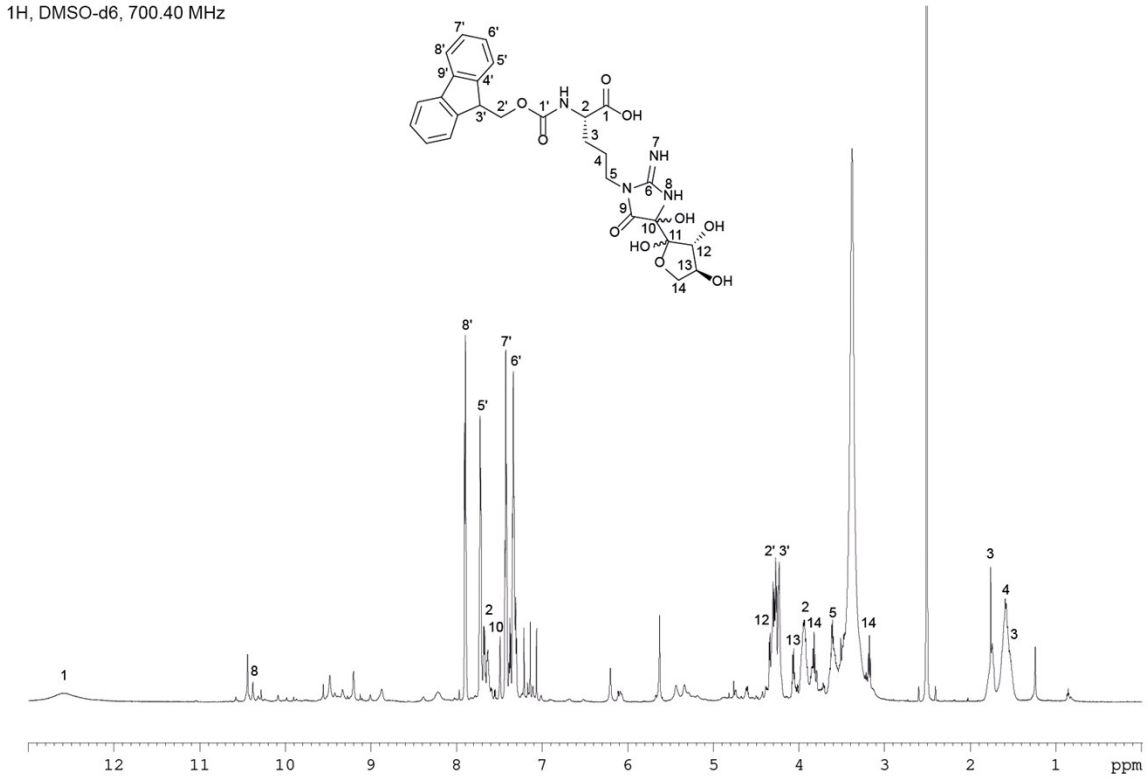

Figure S4: ^{1}H NMR spectrum of N$^{\alpha}$-Fmoc-L-arginine* with peak assignments marked.

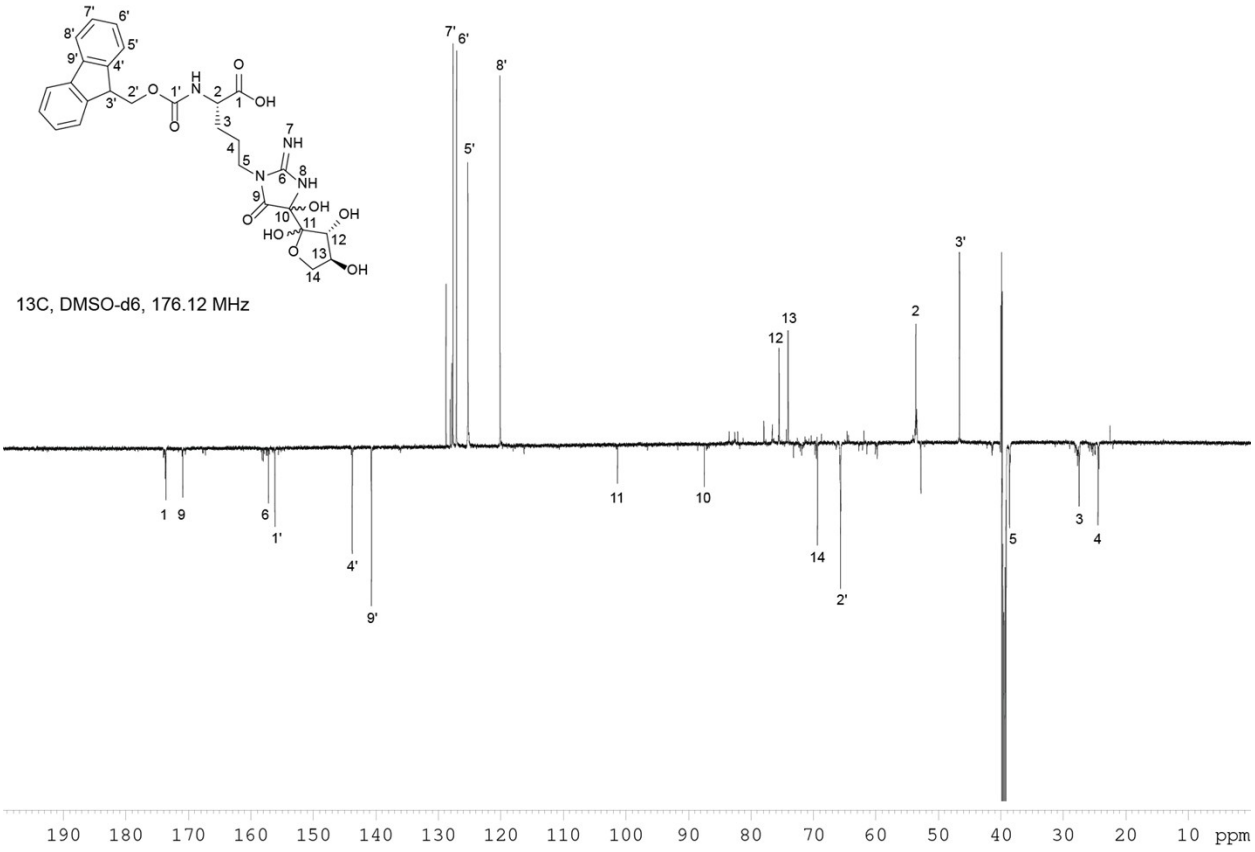

Figure S5: ^{13}C NMR spectrum of N^{α}-Fmoc-L-arginine* with peak assignments marked.

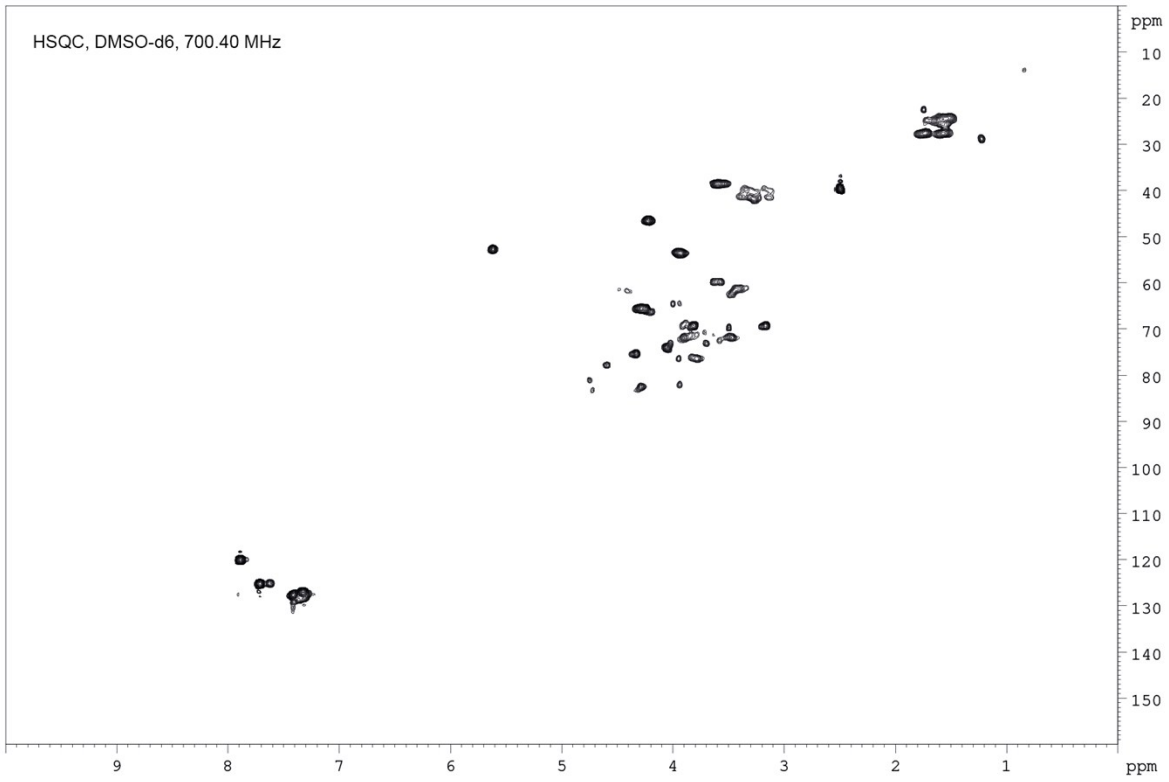

Figure S6: HSQC NMR spectrum of N^{α}-Fmoc-L-arginine*.

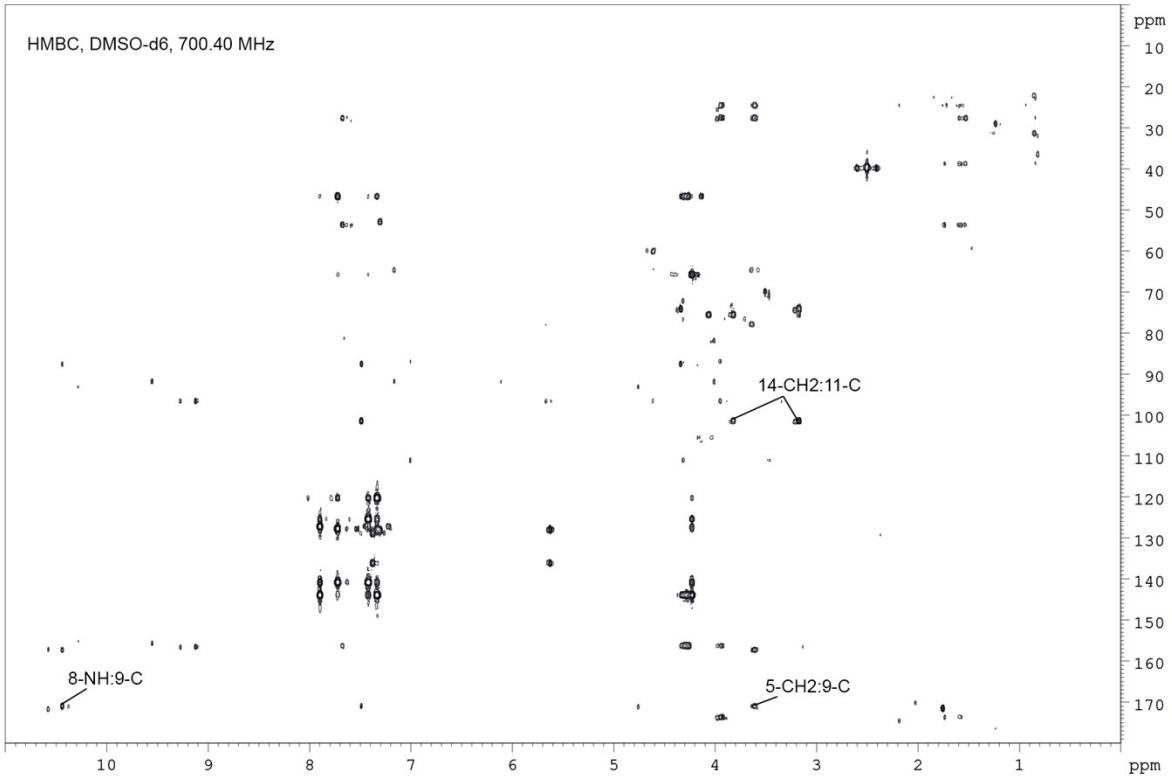

Figure S7: HMBC NMR spectrum of N^{α}-Fmoc-L-arginine*.

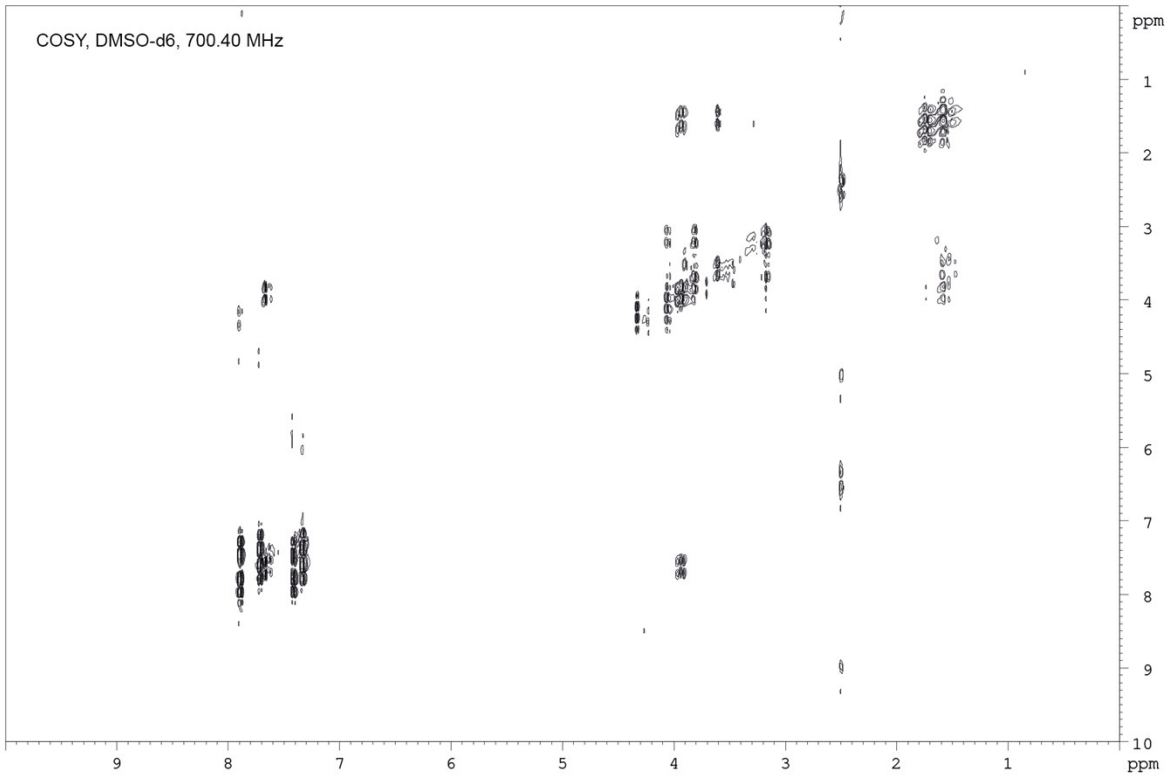

Figure S8: COSY NMR spectrum of N^{α}-Fmoc-L-arginine*.

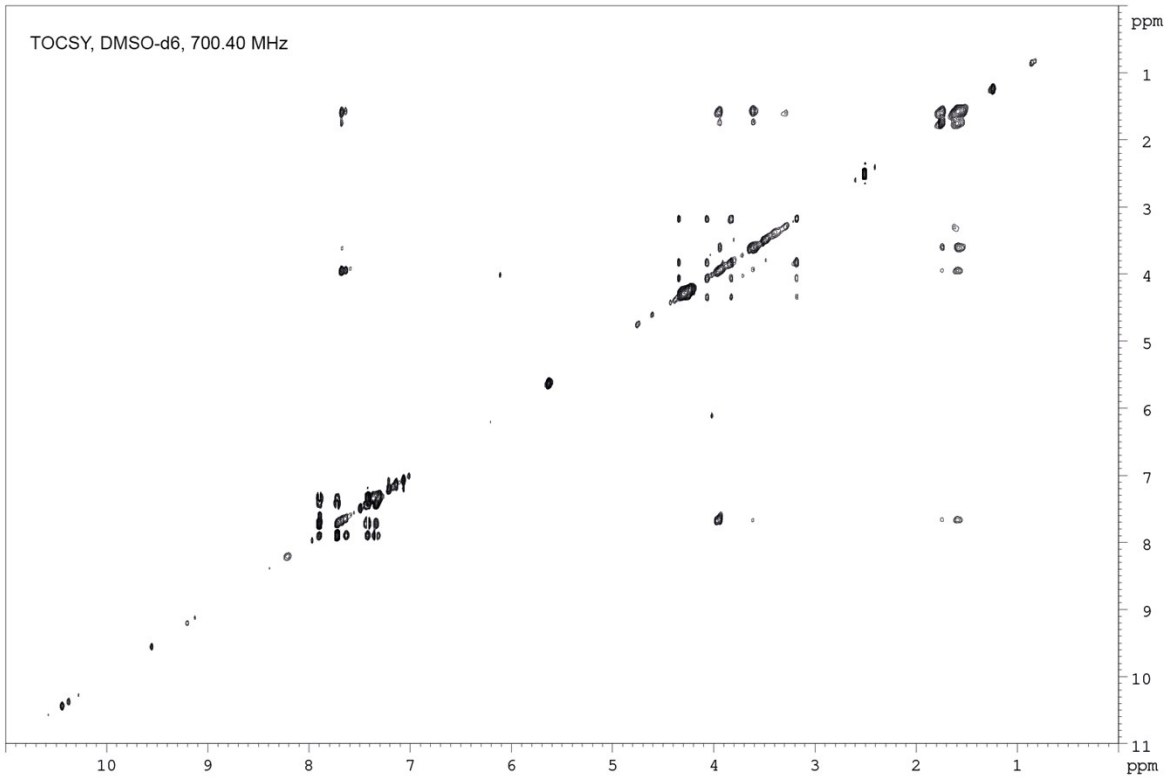

Figure S9: TOCSY NMR spectrum of N^{α}-Fmoc-L-arginine*.

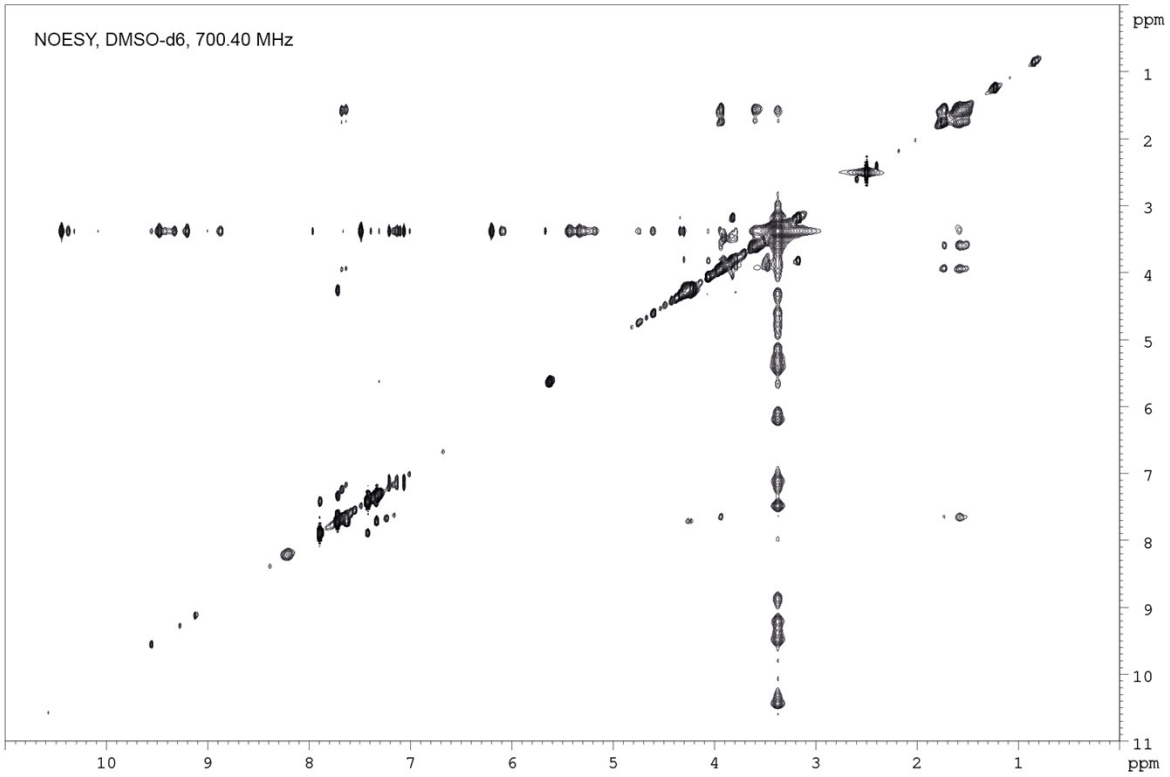

Figure S10: NOESY NMR spectrum of N$^{\alpha}$-Fmoc-L-arginine*.

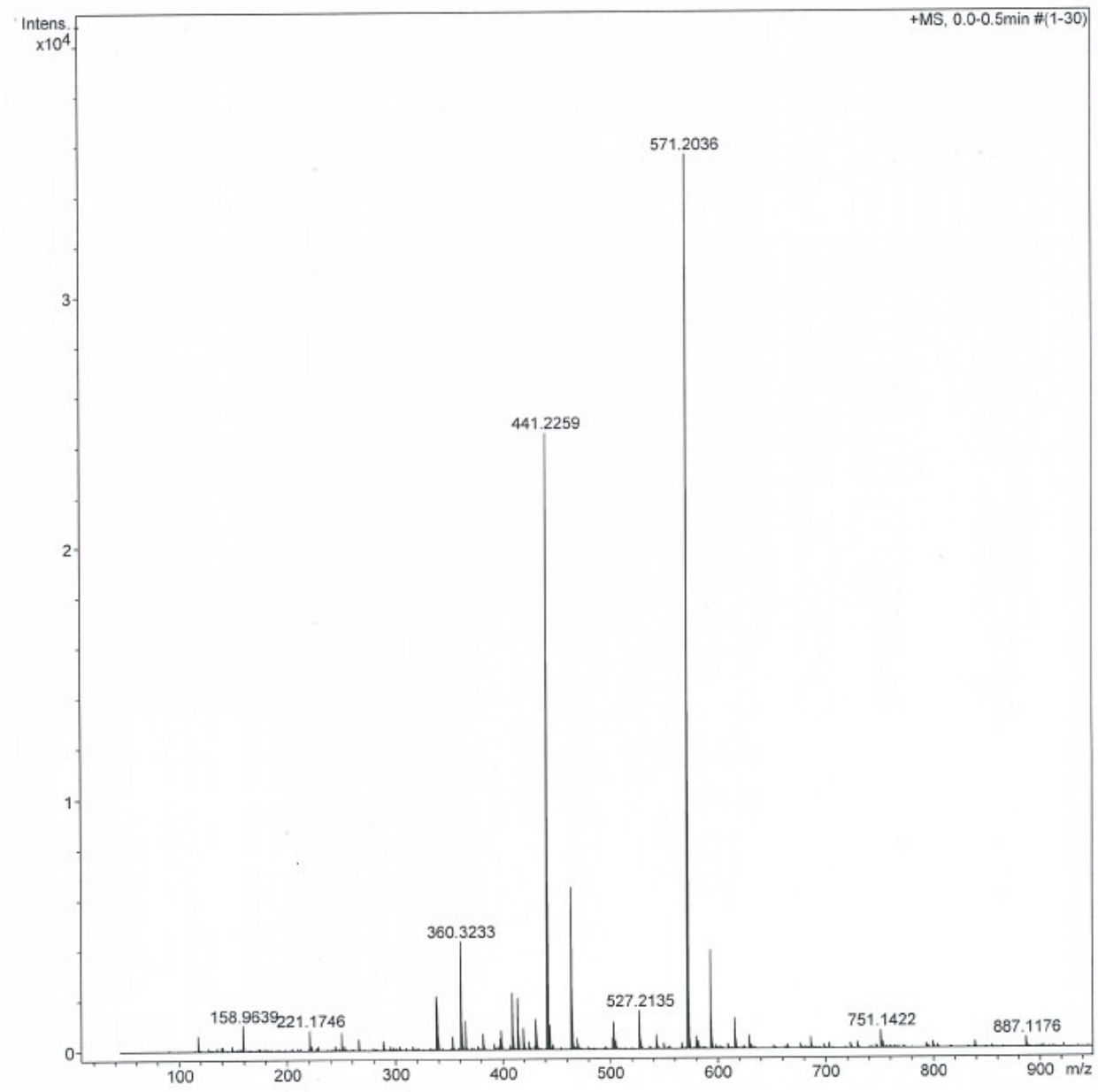

Figure S11: High resolution MS spectrum of N^{α}-Fmoc-L-arginine*. $[M+H]^{+}_{calc}$ 571.2035, $[M+H]^{+}_{obs}$ 571.2036), elemental composition predicted ($C_{27}H_{31}N_4O_{10}$), found ($C_{27}H_{31}N_4O_{10}$).