N-Terminal Protein Modification Using Simple Aminoacyl Transferase Substrates

Anne M. Wagner, Mark W. Fegley, John B. Warner, Christina L. J. Grindley,

Nicholas P. Marotta, & E. James Petersson

Department of Chemistry, University of Pennsylvania
231 South 34th Street, Philadelphia, Pennsylvania 19104-6323 USA.

Additional General Information. *N*-Boc-L-leucine (Boc-Leu-OH), *N*-Boc-L-*p*-azidophenylalanine (Boc-Azf-OH), *N*-Boc-L-naphthylalanine (Boc-Nap-OH), *N*-Boc-, *N*-Me-L-phenylalanine (Boc-Mef-OH), *N*-Acetyl-L-phenylalanine (Acf-OH), and *N*-Fmoc-methoxycoumarinylalanine (Fmoc-Mcm-OH) were purchased from Bachem (Torrence, CA). All solvents were purchased from Fisher Scientific (Pittsburgh, PA). QuikChange® site-directed mutagenesis kits were purchased from Stratagene (La Jolla, CA). DNA oligomers were purchased from Integrated DNA Technologies, Inc (Coralville, IA). Bradford reagent assay kits were purchased from BioRAD (Hercules, CA). Protease inhibitor cocktail was purchased from Sigma-Aldrich (St. Louis, MO). All other reagents were purchased from Fisher Scientific (Pittsburgh, PA). Infrared spectra were recorded on a Nicolet 6700 FT-IR instrument.

Donor (4b-i) Synthesis. As described in the main text, the carboxy-terminus of the amino acid (or analog) was activated as a cyanomethyl ester using chloroacetonitrile to give **1b-i**. Next, the protected amino acid was attached to one of the available hydroxyls (2′ OH or 3′ OH) of DMT-A without preference. The resulting protected adenylate (**3b-i**) was deprotected using a

50/50 solvent mixture of trifluoroacetic acid and THF (1 mL) or neat TFA (1 mL) with TIPSH present as a scavenger, stirred for 24 h, concentrated under reduced pressure, extracted using DCM and water as mentioned in the main text, and HPLC purified. The final product, deprotected adenylate (**4b-i**), was purified *via* C18 column purification using an HPLC.

(*S*)-cyanomethyl 2-((*tert*-butoxycarbonyl)amino)-4-methylpentanoate (Boc-Leu-OCH_2CN, 1b). Chloroacetonitrile (5 mL) and DIPEA (160 mg, 0.21 mL, 1.2 mmol) were added to Boc-Leu-OH (252 mg, 1.09 mmol) and stirred for 12 h. The solvent was removed under reduced pressure and SiO_2 flash chromatography (20% ethyl acetate in hexanes) afforded 284 mg of a pale yellow oil in 96% yield. R_f 0.5 in 20% ethyl acetate in hexanes; ^{1}H NMR (500 MHz, $CDCl_3$): δ 5.00 (d, J = 7.7 Hz, 1H), 4.80 – 4.65 (m, 2H), 4.30 – 4.29 (m, 1H), 1.70 – 1.65 (m, 1H), 1.60 – 1.47 (m, 2H), 1.39 (s, 9H), 0.91 – 0.89 (m, 6H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 172.2, 155.5, 114.21, 80.3, 53.2, 52.0, 41.0, 28.3, 24.8, 22.8, 21.7; HRMS (ESI) m/z calcd for $C_{13}H_{22}N_2O_4Na$ $[M + Na]^+$ 293.147, found 293.151.

(*S*)-cyanomethyl 3-(4-azidophenyl)-2-((*tert*-butoxycarbonyl)amino)propanoate (Boc-Azf-OCH_2CN, 1c). Chloroacetonitrile (5 mL) and DIPEA (99 mg, 0.13 mL, 0.76 mmol) were added to Boc-Azf-OH (212 mg, 0.691 mmol) and stirred for 12 h. The solvent was removed under reduced pressure and SiO_2 flash chromatography (30% ethyl acetate in hexanes) afforded 230 mg of a pale yellow oil in 96% yield. R_f 0.5 in 30% ethyl acetate in hexanes; ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.13 (d, J = 8.3 Hz, 2H), 6.97 (d, J = 8.4 Hz, 2H), 5.00 (d, J = 7.7 Hz, 1H), 4.80 – 4.65 (m, 2H), 4.62 – 4.58 (m, 1H), 3.12 – 3.01 (m, 2H), 1.40 (s, 9H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 170.7, 155.1, 139.4, 132.0, 130.7, 119.5, 114.0, 80.7, 54.4, 49.0, 37.5, 28.3; HRMS (ESI) m/z calcd for $C_{16}H_{19}N_5O_4Na$ $[M + Na]^+$ 368.133, found 368.135.

(*S*)-cyanomethyl 2-((*tert*-butoxycarbonyl)amino)-3-(naphthalen-2-yl)propanoate (Boc-Nap-OCH_2CN, 1d). Chloroacetonitrile (2.5 mL) and DIPEA (25 mg, 33 μL, 0.19 mmol) were added to Boc-NapAla-OH (54.8 mg, 0.174 mmol) and stirred for 12 h. The solvent was removed under reduced pressure and SiO_2 flash chromatography (30% ethyl acetate in hexanes) afforded 61 mg of a pale yellow solid in 99% yield. R_f 0.5 in 30% ethyl acetate in hexanes; ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.83-7.81 (m, 3H), 7.62 (s, 1H), 7.51 – 7.46 (m, 2H), 7.29 (dd, J = 8.4, 1.5 Hz, 1H), 5.00 (d, J = 7.4 Hz, 1H), 4.79 – 4.64 (m, 3H), 3.31 (dd, J = 14.1, 5.9 Hz), 3.25 (dd, J = 14.0, 6.5 Hz, 1H), 1.41 (s, 9H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 170.9, 155.2, 133.6, 132.8, 132.8, 128.8, 128.3, 127.8, 127.2, 126.5, 126.2, 114.0, 80.7, 54.5, 49.0, 38.2, 28.4; HRMS (ESI) m/z calcd for $C_{20}H_{22}N_2O_4Na$ $[M + Na]^+$ 377.147, found 377.149.

(*S*)-cyanomethyl 2-((*tert*-butoxycarbonyl)(methyl)amino)-3-phenylpropanoate (Boc-Mef-OCH_2CN, 1e). Chloroacetonitrile (1.1 mL) and DIPEA (258 mg, 0.341 mL, 1.99 mmol) were added to Boc-MePhe-OH (497 mg, 1.78 mmol) in tetrahydrofuran (5 mL) and stirred for 12 h. The solvent was removed under reduced pressure and SiO_2 flash chromatography (20 - 35% ethyl acetate in hexanes) afforded 505 mg of a pale yellow oil in 89% yield. R_f 0.3 in 20% ethyl acetate in hexanes; ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.28 – 7.25 (m, 2H), 7.20 – 7.16 (m, 3H), 4.72 – 4.68 (m, 2H), 4.53 – 4.52 (m, 1H), 3.27 (m, 1H), 3.11 – 3.02 (m, 1H), 2.67 (d, J = 17 Hz, 3H), 1.36 (d, J = 12 Hz, 9H)**;** ^{13}C NMR (125 MHz, $CDCl_3$): δ 169.9, 155.5, 136.8, 128.9, 128.6, 128.4, 126.8, 126.7, 114.2, 80.8, 61.5, 48.7, 35.4, 33.0, 28.1; HRMS (ESI) m/z calcd for $C_{17}H_{22}N_2O_4Na$ $[M + Na]^+$ 341.147, found 341.149.

(*S*)-cyanomethyl 2-acetamido-3-phenylpropanoate (Acf-OCH_2CN, 1f). Chloroacetonitrile (2 mL) and DIPEA (630 mg, 0.841 mL, 4.87 mmol) were added to *N*-Ac-Phe-OH (501 mg, 2.42 mmol) and stirred for 12 h. The solvent was removed under reduced pressure and SiO_2 flash

chromatography (50% ethyl acetate in hexanes) afforded (535 mg) of a pale yellow oil in 90% yield. R_f 0.2 in 50% ethyl acetate in hexanes; ^{1}H NMR (500 MHz, $CDCl_3$): δ 7.26 (t, J = 7 Hz, 2H), 7.22 (t, J = 2.3 Hz, 1H), 7.07 (d, J = 7.1 Hz, 2H), 6.02 (d, J = 7.4 Hz, 1H), 4.83 (dd, J = 14.0, 6.5 Hz, 1H), 4.66 (dd, J = 15.7, 4.7 Hz, 2H), 3.10 – 3.01 (m, 2H), 1.91 (s, 3H); ^{13}C NMR (125 MHz, $CDCl_3$): δ 170.6, 170.1, 135.2, 129.3, 139.0, 127.7, 114.0, 53.6, 53.2, 49.0, 37.7, 23.0; HRMS (ESI) m/z calcd for $C_{13}H_{14}N_2O_3Na$ $[M + Na]^+$ 269.090, found 269.091.

(*S*)-cyanomethyl 2-azido-3-phenylpropanoate (Boc-N$_3$f-OCH$_2$CN, 1g). Chloroacetonitrile (1 mL) and DIPEA (310 mg, 0.411 mL, 2.41 mmol) were added to 2-azido-3-phenylpropanoate and stirred for 12 h. The solvent was removed under reduced pressure and SiO_2 flash chromatography (40 - 50% ethyl acetate in hexanes) afforded (250 mg) of a white solid in 50% yield. R_f 0.5 in 50% ethyl acetate in hexanes; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.37 (t, J = 7.6 Hz, 2H), 7.32-7.29 (m, 1H), 7.25 (t, J = 7.6 Hz, 2H), 4.77 (t, J = 1.0 Hz, 1H), 4.18 (dd, J = 6.2, 1.8 Hz, 1H), 3.21 (dd, J = 14.0, 5.8 Hz, 1H), 3.08 (dd, J = 13.9, 8.4 Hz, 1H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 168.8, 135.1, 129.3, 129.1, 127.8, 113.7, 62.9, 49.3, 37.7; HRMS (ESI) m/z calcd for $C_{11}H_{10}N_4O_2Na$ $[M + Na]^+$ 253.070, found 253.071.

(*S*)-cyanomethyl 2-((*tert*-butoxycarbonyl)amino)-3-(7-methoxy-2-oxo-2*H*-chromen-4-yl) propanoate (Boc-Mcm-OCH$_2$CN, 1h). A solution of 20% piperidine in tetrahydrofuran (4 mL) was added to Fmoc-Mcm-OH (200 mg, 0.439 mmol) and was stirred for 15 minutes. The solvent was removed under reduced pressure and ethyl acetate (10 mL) and 2M sodium hydroxide was added until pH > 8 followed by addition of Boc anhydride (110 mg, 0.526 mmol) to the white residue and stirred overnight. The organic layer was extracted and washed with saturated sodium bicarbonate (2 x 10 mL) and all aqueous fractions were combined and acidified with 1 M sodium bisulfate until pH < 2. The aqueous layer was then extracted with ethyl acetate

(3 x 10 mL), dried with magnesium sulfate, and concentrated under reduced pressure. Chloroacetonitrile (1 mL) and DIPEA (62 mg, 82 μL, 0.48 mmol) were added to the white residue and stirred for 12 h. The solvent was removed under reduced pressure and SiO_2 flash chromatography (50% - 65% ethyl acetate in hexanes) afforded 56 mg of a pale yellow oil in 35% yield. R_f 0.4 in 50% ethyl acetate in hexanes; ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.37 (d, J = 8.7 Hz, 1H), 6.90 (dd, J = 8.9, 2.4 Hz, 1H), 6.85 (d, J = 2.1 Hz, 1H), 6.12 (s, 1H), 5.16 (d, J = 7.7 Hz, 1H), 4.79 (d, J = 4.2 Hz, 2H), 4.69-4.68 (m, 1H), 3.88 (s, 3H), 3.32 (dd, J = 14.1, 5.2 Hz, 1H), 3.12 (dd, J = 13.9, 8.4 Hz, 1H), 1.41 (s, 9H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 170.0, 163.2, 160.8, 155.9, 155.1, 150.5, 125.3, 113.7, 113.3, 113.0, 112.3, 101.6, 81.3, 56.0, 52.8, 49.5, 34.6, 28.4; HRMS (ESI) m/z calcd for $C_{20}H_{22}N_2O_7Na$ $[M + Na]^+$ 425.133, found 425.133.

(*S*)-cyanomethyl 3-(4-benzoylphenyl)-2-((*tert*-butoxycarbonyl)amino)propanoate (Boc-Bzf-OCH$_2$CN, 1h). Chloroacetonitrile (1 mL) and DIPEA (20 mg, 34 µL, 0.21 mmol) were added to Boc-*p*-benzoylphenylalanine (61 mg, 0.19 mmol). The reaction was stirred for 18 h, then concentrated under reduced pressure. SiO_2 flash chromatography (40% ethyl acetate in hexanes) afforded 70 mg of a pale yellow oil in 93% yield. R_f 0.5 in 50% ethyl acetate in hexanes. ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.79 (dd, J = 7.1, 4.9 Hz, 4H), 7.59 (t, J = 7.6 Hz, 1H), 7.48 (t, J = 7.6 Hz, 2H), 7.29 (d, J = 8.0 Hz, 2H), 5.00 (d, J = 7.4 Hz, 1H), 4.83 (d, J = 15.7 Hz, 1H), 4.72 – 4.69 (m, 2H), 3.25 (dd, J = 13.7, 5.8 Hz, 1H), 3.16 (dd, J = 13.1, 6.5 Hz, 1H), 1.43 (s, 9H); ^{13}C NMR (125 MHz, $CDCl_3$) δ 196.4, 170.6, 155.1, 140.3, 137.6, 136.9, 132.7, 130.8, 130.2, 129.4, 128.5, 113.9; HRMS (ESI) m/z calcd for $C_{23}H_{24}N_2NaO_5$ $[M + Na]^+$ 431.158, found 431.158.

(*S*)-(2*R*,3*R*,4*R*,5*R*)-2-(6-amino-9*H*-purin-9-yl)-5-((bis(4-methoxyphenyl)(phenyl) methoxy)methyl)-4-hydroxytetrahydrofuran-3-yl 2-((*tert*-butoxycarbonyl)amino)-4-methylpentanoate (Boc-Leu-(DMT)-A, 3b). Tetrahydrofuran (5 mL) was added to Boc-Leu-OCH_2CN **1b** (226 mg, 0.837 mmol), (DMT)-A (120 mg, 0.211 mmol), and TBAAc (15 mg, catalyst) and stirred for 24 h. The solvent was removed under reduced pressure and chromatography (80 - 100% ethyl acetate in petroleum ether, 5% methanol in ethyl acetate), afforded 122 mg of a white solid in 74% yield. R_f 0.5 in 5% methanol in ethyl acetate; ^1H and ^{13}C NMR for **3b** shown below; LRMS (ESI) m/z calcd for $C_{42}H_{50}N_6O_9Na$ $(M + Na)^+$ 805.4, found 805.3.

(*S*)-(2*R*,3*R*,4*R*,5*R*)-2-(6-amino-9*H*-purin-9-yl)-5-((bis(4-methoxyphenyl)(phenyl)methoxy) methyl)-4-hydroxytetrahydrofuran-3-yl 3-(4-azidophenyl)-2-((*tert*-butoxycarbonyl)amino) propanoate (Boc-Azf-(DMT)-A, 3c). Tetrahydrofuran (5 mL) was added to Boc-Azf-OCH_2CN **1c** (101 mg, 0.290 mmol), (DMT)-A (43 mg, 75 μmol), and TBAAc (2 mg, catalyst) and stirred for 24 h. Solvent was removed under vacuum and preparative TLC (100% ethyl acetate), afforded 44 mg of white solid in 69% yield. R_f 0.3 - 0.5 in ethyl acetate; ^1H and ^{13}C NMR for **3c** shown below; LRMS (ESI) m/z calcd for $C_{42}H_{47}N_9O_9Na$ $(M + Na)^+$ 880.3, found 880.4.

(*S*)-(2*R*,3*R*,4*R*,5*R*)-2-(6-amino-9*H*-purin-9-yl)-5-((bis(4-methoxyphenyl)(phenyl) methoxy) methyl)-4-hydroxytetrahydrofuran-3-yl 2-((*tert*-butoxycarbonyl)amino)-3-(naphthalen-2-yl)propanoate (Boc-Nap-(DMT)-A, 3d). Tetrahydrofuran (5 mL) was added to Boc-Nap-OCH_2CN **1d** (58 mg, 0.16 mmol), (DMT)-A (26 mg, 46 μmol), and TBAAc (6.4 mg, catalyst) and stirred for 24 h. The solvent was removed under reduced pressure and preparative TLC (5% methanol in ethyl acetate), afforded 36 mg of a white solid in 91% yield. R_f 0.3 - 0.5 in 5% methanol in ethyl acetate; ^1H and ^{13}C NMR for **3d** shown below; LRMS (ESI) m/z calcd for $C_{49}H_{50}N_6O_9Na$ $(M + Na)^+$ 889.9, found 889.4.

(*S*)-(2*R*,3*R*,4*R*,5*R*)-2-(6-amino-9*H*-purin-9-yl)-5-((bis(4-methoxyphenyl)(phenyl) methoxy) methyl)-4-hydroxytetrahydrofuran-3-yl 2-((*tert*-butoxycarbonyl)(methyl)amino)-3-phenylpropanoate (Boc-Mef-(DMT)-A, 3e). Tetrahydrofuran (3.5 mL) and DIPEA (44 mg, 60 μL, 0.35 mmol) were added to Boc-Mef-OCH_2CN **1e** (110 mg, 0.347 mmol), (DMT)-A (50 mg, 89 μmol), and TBAAc (3 mg, catalyst) and stirred for 12 h. The solvent was removed under reduced pressure and SiO_2 flash chromatography (50 - 100% ethyl acetate in hexanes), afforded 10 mg of a white solid in 14% yield. R_f 0.4 - 0.5 in ethyl acetate; ^1H and ^{13}C NMR for **3e** shown below; LRMS (ESI) m/z calcd for $C_{46}H_{50}N_6NaO_9$ $(M + Na)^+$ 853.4, found 853.5.

(*S*)-(2*R*,3*R*,4*R*,5*R*)-2-(6-amino-9*H*-purin-9-yl)-5-((bis(4-methoxyphenyl)(phenyl) methoxy)methyl)-4-hydroxytetrahydrofuran-3-yl 2-acetamido-3-phenylpropanoate (Acf-(DMT)-A, 3f). Tetrahydrofuran (5 mL) and DIPEA (250 mg, 0.30 mL, 1.9 mmol) were added to Boc-Acf-OCH_2CN **1f** (400 mg, 1.93 mmol), (DMT)-A (73 mg, 0.13 mmol) and TBAAc (5 mg, catalyst) and stirred for 12 h. The solvent was removed under reduced pressure and SiO_2 flash chromatography (5% methanol in ethyl acetate), afforded 75 mg of a clear solid that was mix of DMTA and product. Subsequent HPLC analysis suggests approximately 38% of this mixture was product. R_f 0.1 - 0.2 5% methanol in ethyl acetate; ^1H and ^{13}C NMR for **3f** shown below; LRMS (ESI) m/z calcd for $C_{42}H_{43}N_6O_8$ $(M + H)^+$ 759.3 $C_{42}H_{42}N_6NaO_8$ $(M + Na)^+$ 781.3, found 759.3, 781.4.

(*S*)-(2*R*,3*R*,4*R*,5*R*)-2-(6-amino-9*H*-purin-9-yl)-5-((bis(4-methoxyphenyl)(phenyl)methoxy) methyl)-4-hydroxytetrahydrofuran-3-yl 2-azido-3-phenylpropanoate (Boc-N_3f-(DMT)-A, 3g). Tetrahydrofuran (3.5 mL) and DIPEA (44 mg, 60 μL, 0.35 mmol) were added to Boc-N_3f-OCH_2CN **1g** (80 mg, 0.35 mmol), (DMT)-A (50 mg, 89 μmol), and TBAC (3 mg, catalyst) and stirred for 12 h. The solvent was removed under reduced pressure and purification by silica flash

chromatography (75 - 100% ethyl acetate in hexanes), afforded 53 mg of a white solid in 80% yield. R_f 0.3 - 0.5 in ethyl acetate; ^{1}H and ^{13}C NMR for **3g** shown below; LRMS (ESI) m/z calcd for $C_{40}H_{39}N_8O_7$ $(M + H)^+$ 743.3, $C_{40}H_{39}N_8O_7Na$ $(M + Na)^+$ 765.3, found 743.4, 765.4; FTIR (Film) ν_{max} 3334, 3179, 3063, 3032, 2953, 2934, 2837, 2112 ($-N_3$), 1757, 1644, 1606, 1508, 1252, 1177 cm^{-1}.

(*S*)-(2*R*,3*R*,4*R*,5*R*)-2-(6-amino-9*H*-purin-9-yl)-5-((bis(4-methoxyphenyl)(phenyl) methoxy) methyl)-4-hydroxytetrahydrofuran-3-yl 2-((*tert*-butoxycarbonyl)amino)-3-(7-methoxy-2-oxo-2*H*-chromen-4-yl)propanoate (Boc-Mcm-(DMT)-A, 3h). Tetrahydrofuran (3.5 ml) and DIPEA (44 mg, 60 μl, 0.35 mmol) were added to Boc-Mcm-OCH_2CN **1h** (56 mg, 0.15 mmol), (DMT)-A (50 mg, 89 μmol), and TBAAc (3 mg, catalyst) and stirred for 12 h. The solvent was removed under reduced pressure and SiO_2 flash chromatography (50 - 100% ethyl acetate in hexanes), afforded 35 mg of a white solid in 43% yield. R_f 0.3 - 0.4 in ethyl acetate; ^{1}H and ^{13}C NMR for **3h** shown below; LRMS (ESI) m/z calcd for $C_{49}H_{50}N_6NaO_{12}$ $(M + Na)^+$ 937.3, found 937.4.

(*S*)-(2*R*,3*R*,4*R*,5*R*)-2-(6-amino-9*H*-purin-9-yl)-5-((bis(4-methoxyphenyl)(phenyl)methoxy) methyl)-4-hydroxytetrahydrofuran-3-yl 3-(4-benzoylphenyl)-2-((*tert*-butoxycarbonyl) amino)propanoate (Boc-Bzf-(DMT)-A, 3i). THF (3.5 mL) and DIPEA (44 mg, 60 µL, 0.35 mmol) were added to Boc-Bzf-OCH_2CN (70 mg, 0.17 mmol), (DMT)-A (50 mg, 89 µmol) and TBAAc (5 mg, catalyst) and stirred for 16 h. The solvent was removed under reduced pressure and SiO_2 flash chromatography (75% - 100% ethyl acetate in hexanes) afforded 66 mg of a white solid in 81% yield. R_f 0.3 in 100% ethyl acetate. LRMS (ESI) m/z calcd for $C_{52}H_{52}N_6NaO_{10}$ $(M + Na)^+$ 943.4, found 943.5.

(*S*)-(2*R*,3*R*,4*R*,5*R*)-2-(6-amino-9*H*-purin-9-yl)-4-hydroxy-5-(hydroxymethyl) tetrahydrofuran-3-yl 2-amino-4-methylpentanoate (Leu-A, 4b). Trifluoroacetic acid (1 mL), tetrahydrofuran (1 mL) and TIPSH (99 mg, 0.13 mL, 0.63 mmol) were added to **3b** (123 mg, 0.157 mmol) following the general deprotection procedure. HPLC/MALDI analysis m/z calcd $C_{16}H_{25}N_6O_5$ $(M + H)^+$ 381.2; Gradient 2 (Main Text); retention time 14.9 min, found 381.1; retention time 16.6 min, found 381.1.

(*S*)-(2*R*,3*R*,4*R*,5*R*)-2-(6-amino-9*H*-purin-9-yl)-4-hydroxy-5-(hydroxymethyl) tetrahydrofuran-3-yl 2-amino-3-(4-azidophenyl)propanoate (Azf-A, 4c). Trifluoroacetic acid (1 mL), tetrahydrofuran (1 mL) and TIPSH (33 mg, 42 μL, 0.21 mmol) were added to **3c** (44 mg, 52 μmol) following the general deprotection procedure. HPLC/MALDI analysis m/z calcd $C_{19}H_{22}N_9O_5$ $(M + H)^+$ 456.2; Gradient 2 (Main Text); retention time 19.6 min, found 456.1; retention time 21.1 min, found 456.1.

(*S*)-(2*R*,3*R*,4*R*,5*R*)-2-(6-amino-9*H*-purin-9-yl)-4-hydroxy-5-(hydroxymethyl) tetrahydrofuran-3-yl 2-amino-3-(naphthalen-2-yl)propanoate (Nap-A, 4d). Trifluoroacetic acid (1 mL), tetrahydrofuran (1 mL) and TIPSH (26 mg, 0.34 mL, 0.17 mmol) were added to **3d** (36 mg, 41 μmol) following the general deprotection procedure. HPLC/MALDI analysis m/z calcd $C_{23}H_{25}N_6O_5$ $(M + H)^+$ 465.2, $C_{23}H_{24}N_6O_5Na$ $(M + Na)^+$ 487.2; Gradient 2 (Main Text); retention time 21.9 min, found 464.6, 486.6; retention time 23.4 min, found 464.6, 486.6.

(*S*)-(2*R*,3*R*,4*R*,5*R*)-2-(6-amino-9*H*-purin-9-yl)-4-hydroxy-5-(hydroxymethyl) tetrahydrofuran-3-yl 2-(methylamino)-3-phenylpropanoate (Mef-A, 4e). Trifluoroacetic acid (1 mL), tetrahydrofuran (1 mL) and oxalic acid (2.9 mg) were added to **3e** (31 mg, 37 μmol) following the general deprotection procedure. No TIPSH was used in this reaction. HPLC/MALDI

analysis m/z calcd $C_{20}H_{25}N_6O_5$ $(M + H)^+$ 429.2, $C_{20}H_{24}N_6O_5Na$ $(M + Na)^+$ 451.2; Gradient 2 (Main Text); retention time 16.3 min, found 429.1; retention time 18.6 min, found 429.1, 451.1.

(*S*)-(2*R*,3*R*,4*R*,5*R*)-2-(6-amino-9*H*-purin-9-yl)-4-hydroxy-5-(hydroxymethyl) tetrahydrofuran-3-yl 2-acetamido-3-phenylpropanoate (Acf-A, 4f). During HPLC analysis, **3f** was directly deprotected to **4f** by exposure to the 0.1% TFA in the HPLC solvent and confirmed via MALDI. HPLC/MALDI analysis m/z calcd $C_{21}H_{25}N_6O_6$ $(M + H)^+$ 457.2, $(M + Na)^+$ $C_{21}H_{24}N_6NaO_6$ 479.2; Gradient 2 (Main Text); retention time 18.2 min, found 457.3, 479.3; retention time 19.5 min, found 457.3, 479.3.

(*S*)-(2*R*,3*R*,4*R*,5*R*)-2-(6-amino-9*H*-purin-9-yl)-4-hydroxy-5-(hydroxymethyl) tetrahydrofuran-3-yl 2-azido-3-phenylpropanoate (N_3f-A, 4g). Trifluoroacetic acid (1 mL) and TIPSH (21 mg, 30 µl, 0.13 mmol) were added to **3g** (25 mg, 33 µmol) following the general deprotection procedure. HPLC/MALDI analysis m/z calcd $C_{19}H_{21}N_8O_5$ $(M + H)^+$ 441.2, $C_{19}H_{20}N_8NaO_8$ $(M + Na)^+$ 463.1; Gradient 2 (Main Text); retention time 24.5 min, found 440.9, 462.9; retention time 24.7 min, found 440.9, 462.9

(*S*)-(2*R*,3*R*,4*R*,5*R*)-2-(6-amino-9*H*-purin-9-yl)-4-hydroxy-5-(hydroxymethyl) tetrahydrofuran-3-yl 2-amino-3-(7-methoxy-2-oxo-2*H*-chromen-4-yl)propanoate (Mcm-A, 4h). Trifluoroacetic acid (1 mL), tetrahydrofuran (1 mL) and oxalic acid (3.9 mg) were added to **3e** (31 mg, 37 µmol) following the general deprotection procedure, except was worked up in water only and water-soluble portion was HPLC purified. No TIPSH was used in this reaction. HPLC/MALDI analysis m/z calcd $C_{23}H_{25}N_6O_8$ $(M + H)^+$ 513.2, $C_{23}H_{24}N_6O_8Na$ $(M + Na)^+$ 535.2; Gradient 2 (Main Text); retention time 19.9 min, found 513.1, 535.1; retention time 20.9 min, found 513.1, 535.1.

(*S*)-(2*R*,3*R*,4*R*,5*R*)-2-(6-amino-9*H*-purin-9-yl)-4-hydroxy-5-(hydroxymethyl) tetrahydrofuran-3-yl 2-amino-3-(4-benzoylphenyl)propanoate (Bzf-A, 4i). TFA (1 mL) and TIPSH (40 mg, 50 μL, 0.26 mmol) were added to DMT-A benzoylphenylalanine (59 mg, 64 μmol) according the general deprotection procedure. HPLC/MALDI analysis m/z calcd $C_{26}H_{27}N_6O_6$ $(M + H)^+$ 419.2, $C_{26}H_{26}N_6NaO_6$ $(M + Na)^+$ 441.2; Gradient 2 (Main Text), retention time 19.7 min found 519.1, 541.1.

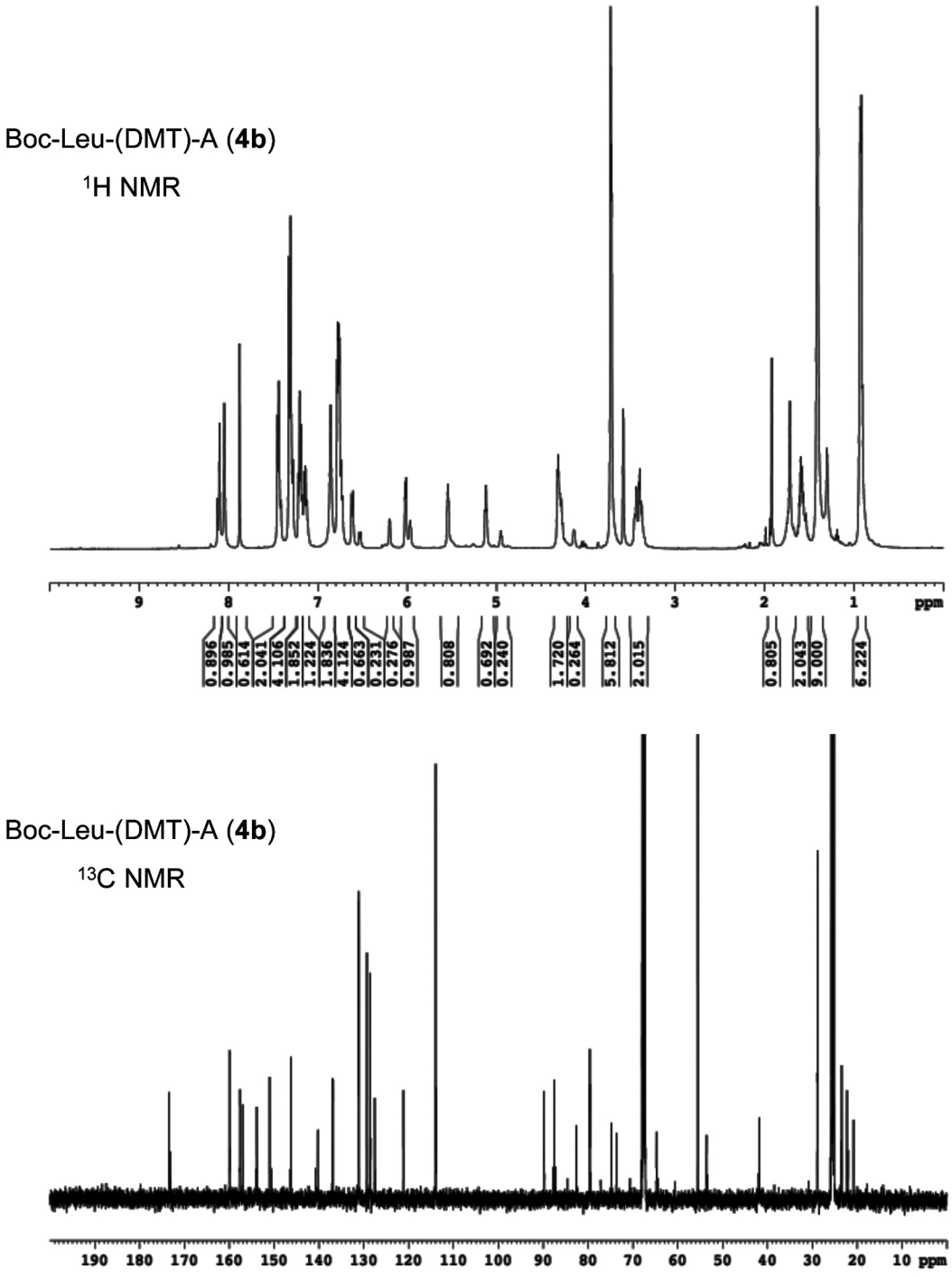


Figure S1. ^{1}H and ^{13}C NMR Characterization of Boc-Leu-(DMT)-A **4b**.

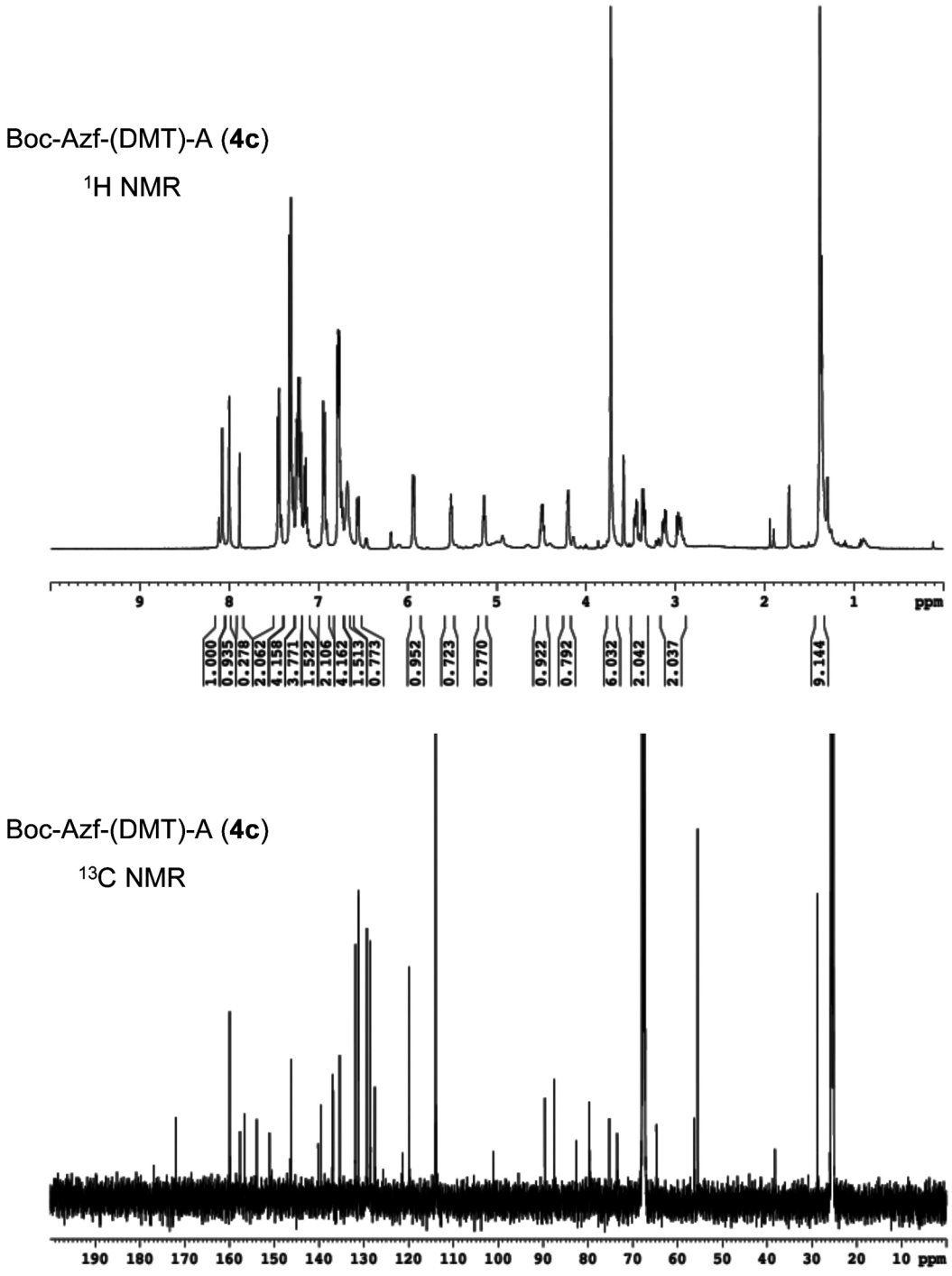


Figure S2. ^{1}H and ^{13}C NMR Characterization of Boc-Azf-(DMT)-A **4c**.

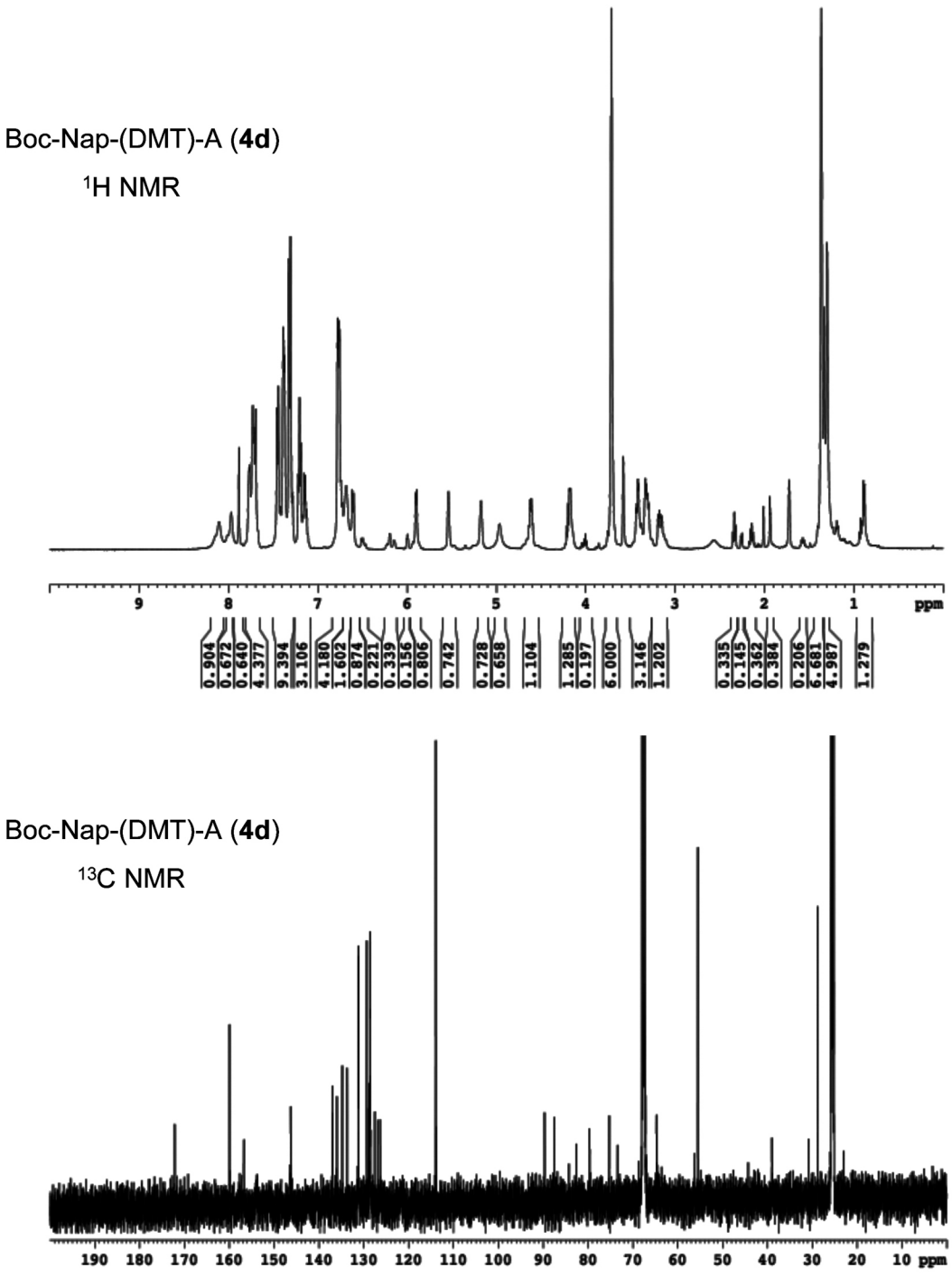


Figure S3. ^{1}H and ^{13}C NMR Characterization of Boc-Nap-(DMT)-A **4d**.

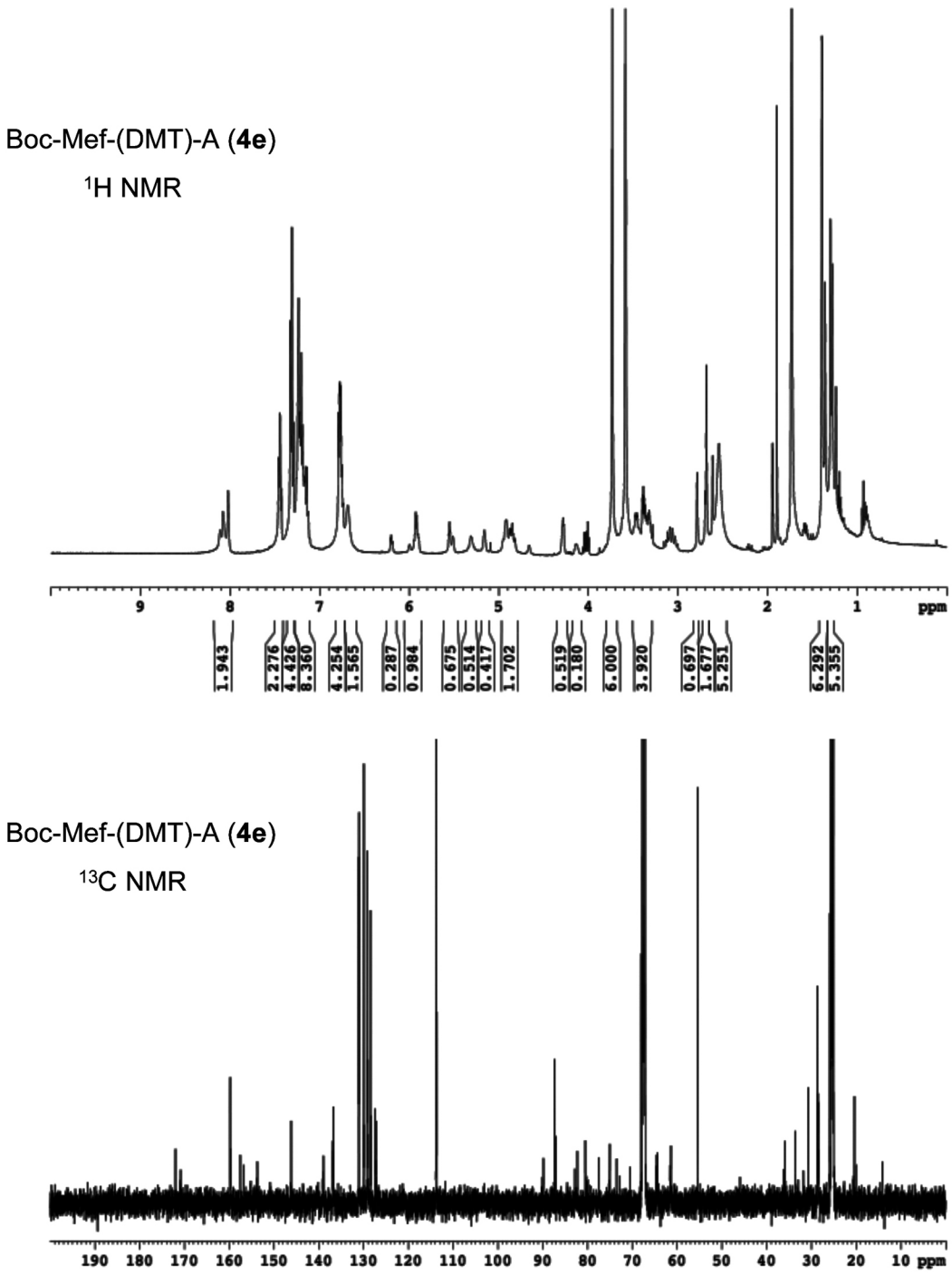


Figure S4. ^{1}H and ^{13}C NMR Characterization of Boc-Mef-(DMT)-A **4e**.

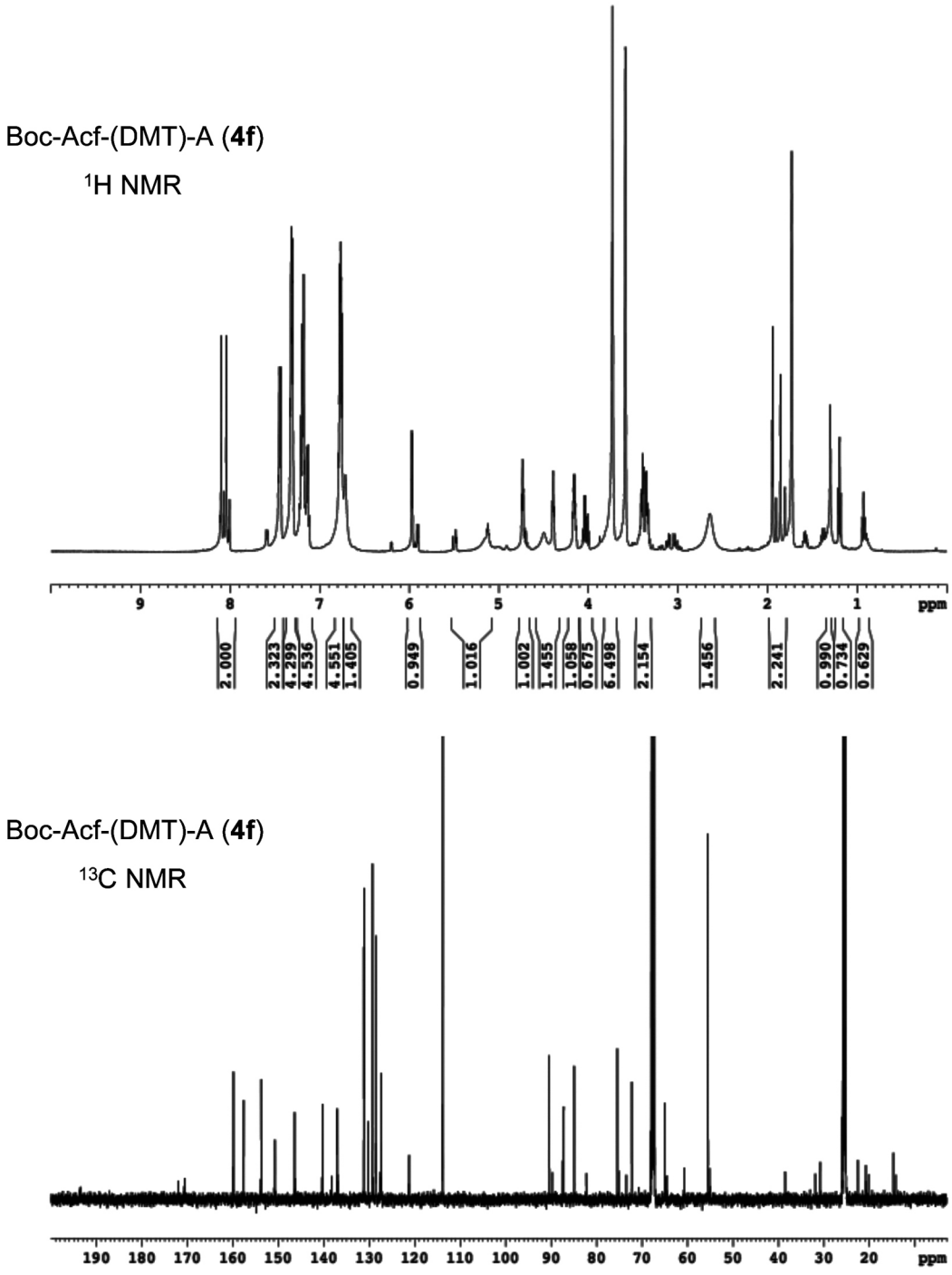


Figure S5. ^{1}H and ^{13}C NMR Characterization of Boc-Acf-(DMT)-A **4f**.

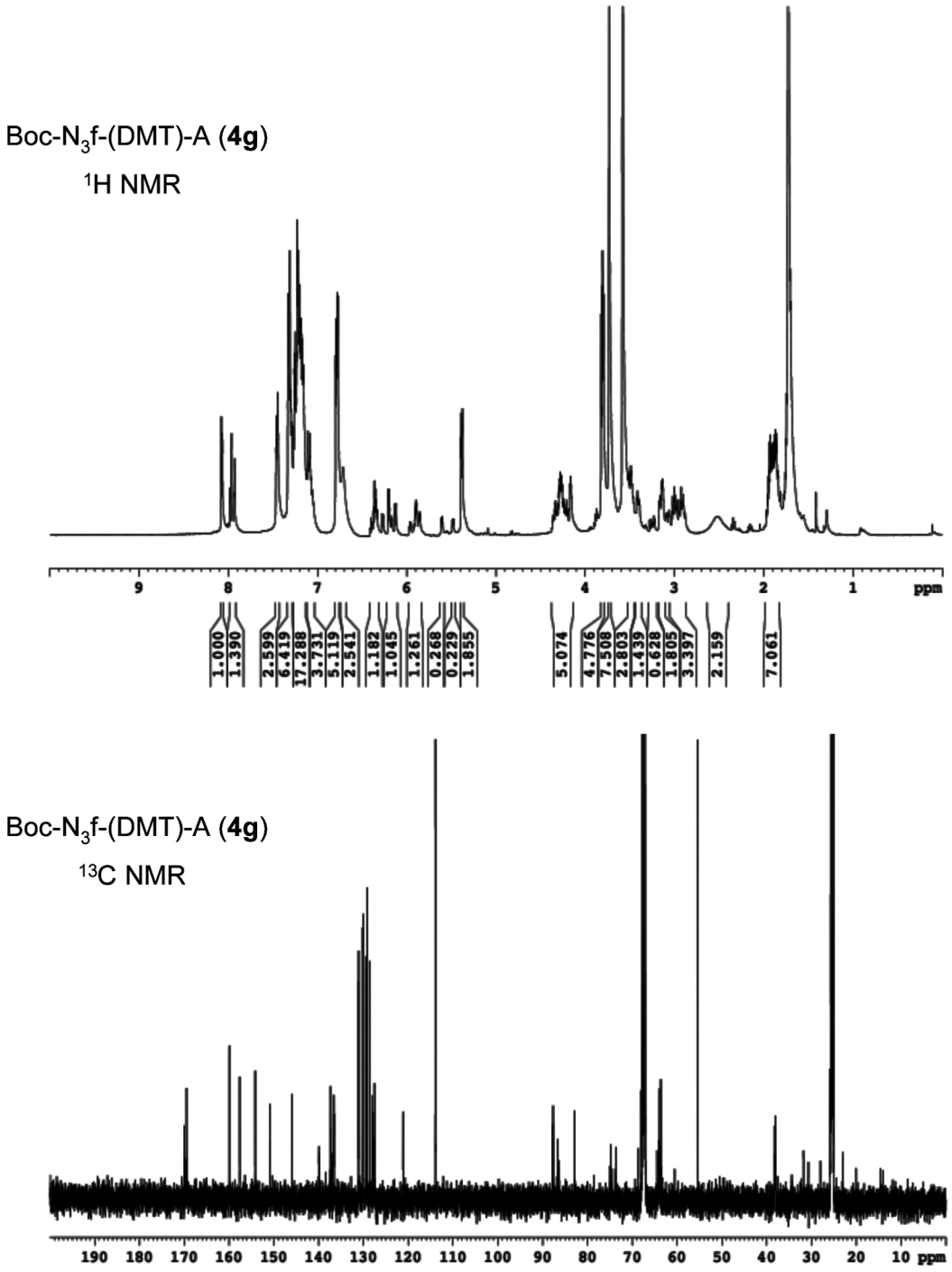


Figure S6. ^{1}H and ^{13}C NMR Characterization of Boc-N_3f-(DMT)-A **4g**.

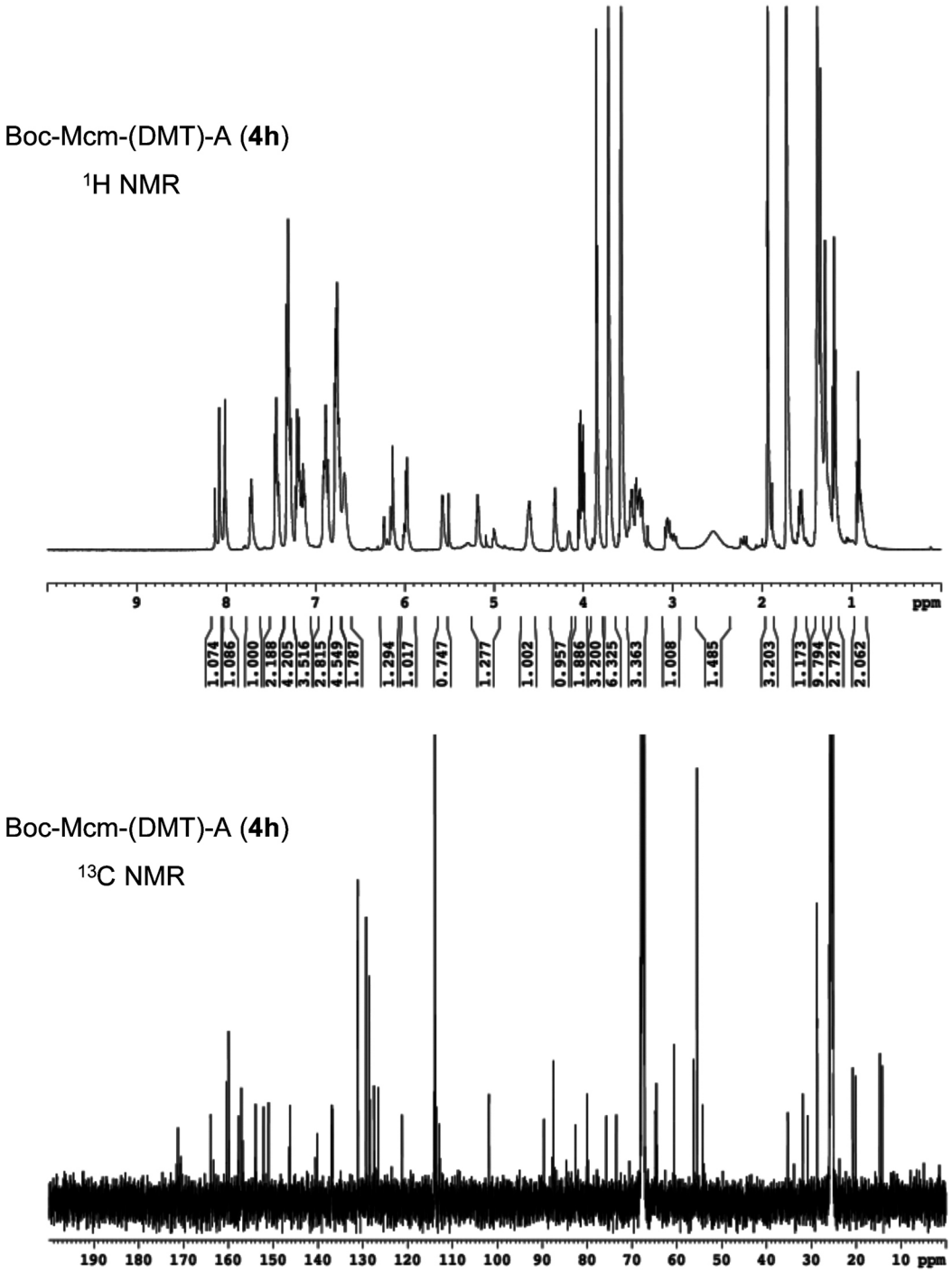


Figure S7. ^{1}H and ^{13}C NMR Characterization of Boc-Mcm-(DMT)-A **4h**.

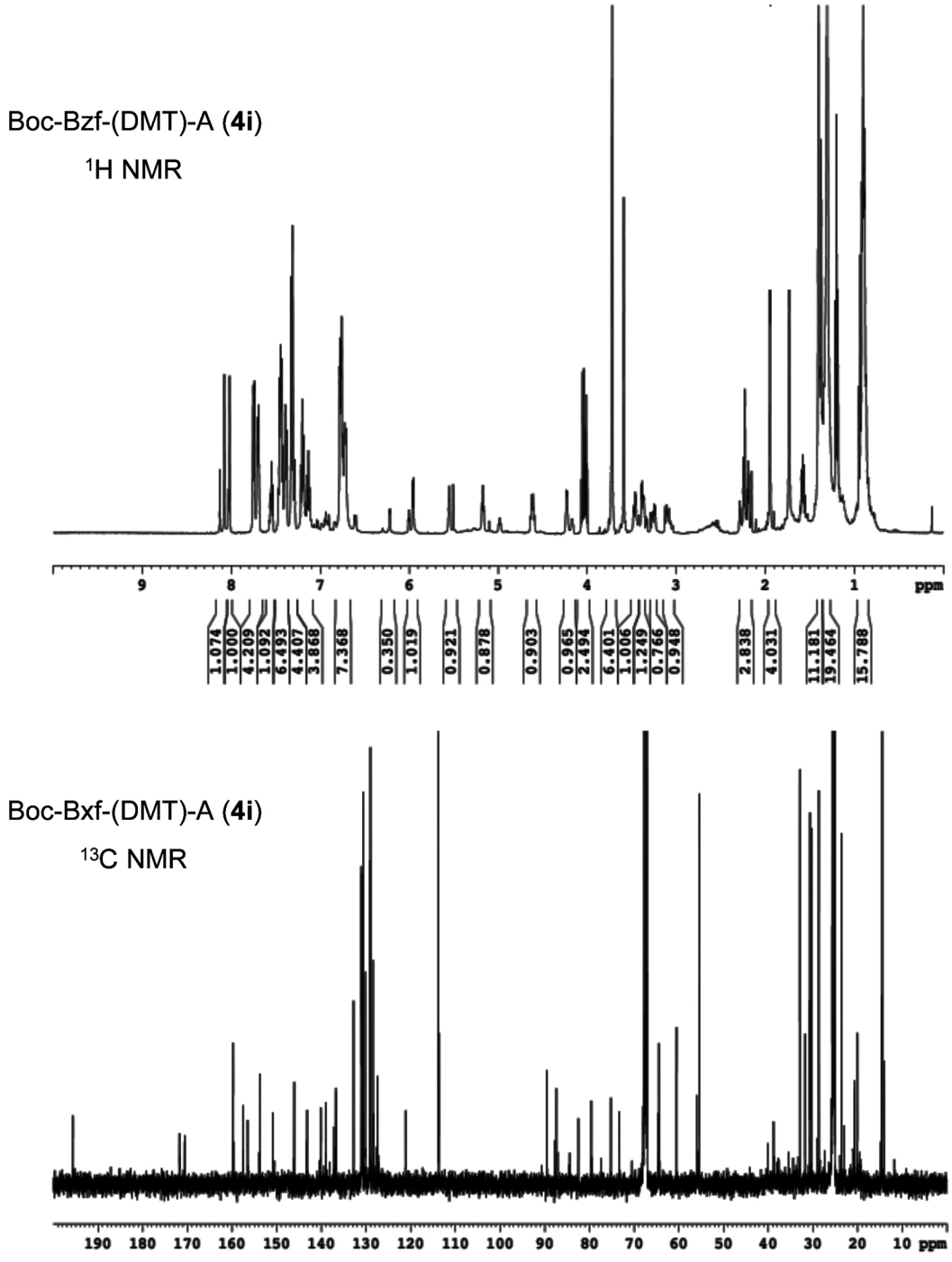


Figure S8. ^{1}H and ^{13}C NMR Characterization of Boc-Bzf-(DMT)-A **4i**.

HPLC Analysis of Donor Deprotection. Trifluoroacetic acid (1 ml) and TIPSH (10 mg, 13 μl, 65 μmol) were added to **3a** (10 mg, 13 μmol) and stirred for 12 h. Solvent was removed under reduced pressure followed by precipitation with diethyl ether. HPLC analysis of the deprotection reaction is shown in Figure S9. The assignment of the 2' and 3' acetylated forms of Phe-A is based on the known thermodynamic preference for 3' acylation of adenosine.[S1] The HPLC method (Solvents A and B defined in main text) had the following solvent gradient (Gradient 3): 0 min 2% B, 60 min 100% B, 65 min 1% B. Monitored at 215 nm and 260 nm.

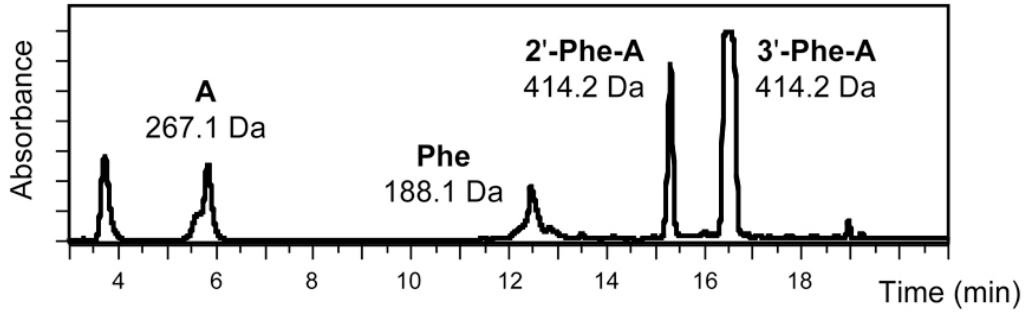


Figure S9. HPLC Analysis of Donor **4a** purity after TFA deprotection.

HPLC Analysis of Donor Hydrolysis. The amino acid analog donor (**4a–h**) was suspected to be hydrolytically unstable and was tested in mock LysAlaAcm reactions without enzyme present and then analyzed by HPLC. The mock LysAlaAcm reactions were performed as described in the LysAlaAcm Ligation Assay section of the main text; however, the AaT solution was replaced with water. Phe-A solutions were analyzed at 30 min, 1 h, and 4 h. Immediately following the reaction, the sample was diluted to 1200 µL and injected onto the C18 HPLC column using Gradient 1. As seen in Figure S10, Phe-A is completely hydrolyzed after 4 h. Additionally, 1 mM A and 1 mM Phe were injected on the HPLC to serve as standards to verify retention times. Note: The relative concentrations of 2' and 3' Phe-A would be different in an

actual reaction where the 3' Phe-A is preferentially consumed by AaT in the transfer reaction.[S2] The change in 2':3' ratio between 0 h and 30 min probably reflects the higher pH of the 30 min data.[S1]

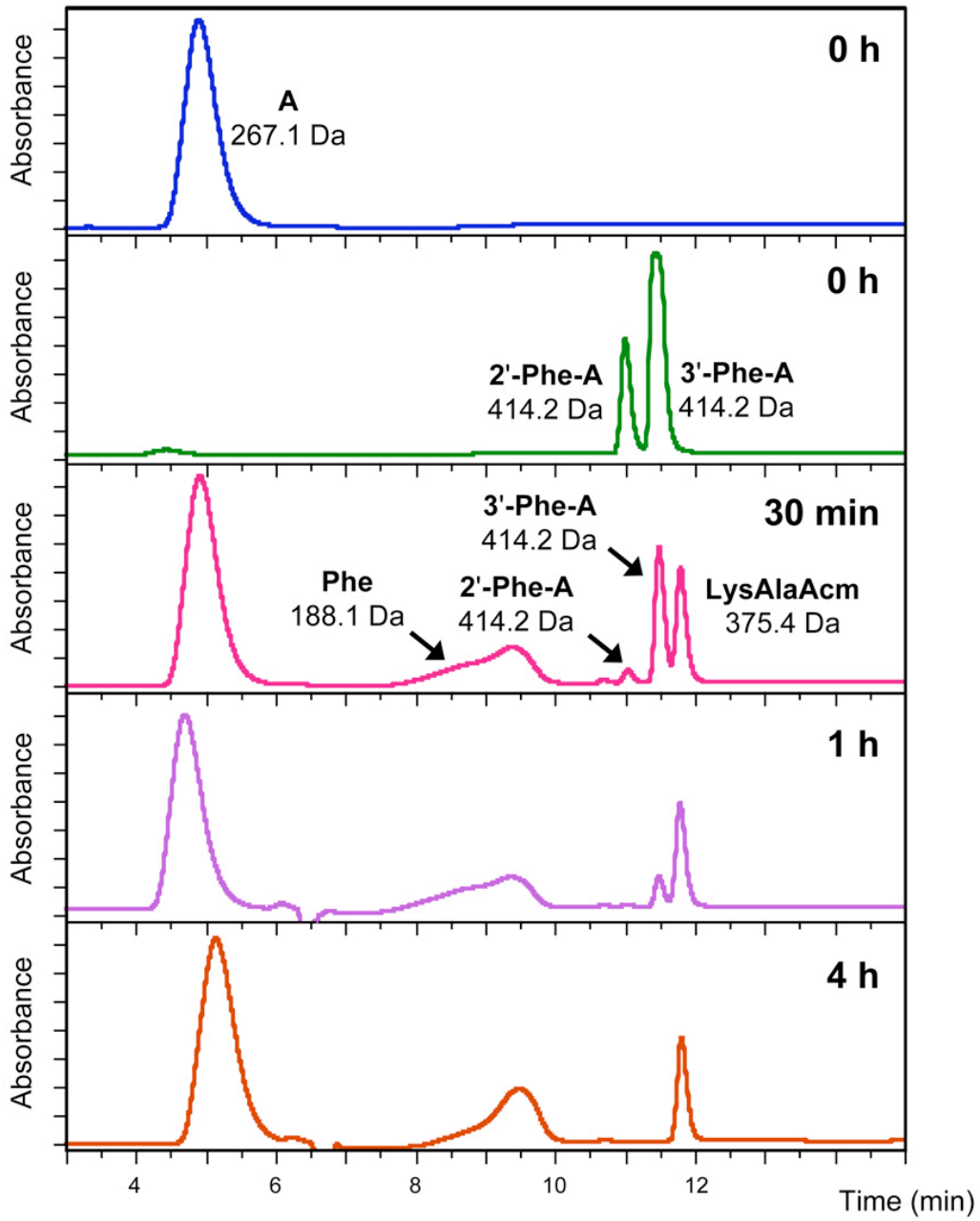


Figure S10. HPLC Analysis of Donor **4a** Hydrolysis in Mock Transferase Reactions. Slight differences in retention time were observed, compound identities confirmed by MALDI MS.

Aminoacyl Transferase Expression and Purification. *E. coli* AaT was expressed from the pEG6 plasmid in *E. coli* BL21-Gold (DE3) cells using a procedure adapted from Graciet *et al*.[S3] *E. coli* were grown in a primary culture of 5 mL LB at 37 °C to OD_{600} of 0.5 and then were rediluted into a secondary culture of 500 mL LB and grown to OD_{600} of 0.6. AaT expression was induced using 0.1 mM isopropyl β-D-thiogalactoside and cells were grown at 25 °C for ~16 h. Cells were pelleted at 6, 000 RPM using a GS3 rotor and Sorvall RC-5 centrifuge. Cell pellets were resuspended in the Ni-NTA binding buffer (50 mM Tris, 10 mM imidazole, 300 mM KCl, and 5 mM β-mercaptoethanol, pH 8.0) and included protease inhibitor cocktail, 1 mM PMSF, and 10 units/mL DNAse1–Grade II. Following resuspension, the cells were lysed using sonication. Soluble proteins were collected *via* centrifugation at 13, 200 RPM for 15 min. Collected soluble protein was gently shaken for 1 h at ambient temperature with Ni-NTA resin. The resin was prepared by rinsing with Ni-NTA binding buffer and then washed with four volumes of Ni-NTA wash buffer (50 mM Tris, 50 mM imidazole, 300 mM KCl, and 5 mM β-mercaptoethanol, pH 8.0). The proteins were eluted with elution buffer (50 mM Tris, 250 mM imidazole, 300 mM KCl, and 5 mM β-mercaptoethanol, pH 8.0). Pure elution fractions of *E. coli* AaT were dialyzed overnight in transferase buffer (50 mM Tris, 30% glycerol, 120 mM $(NH_4)_2SO_4$, 5 mM β-mercaptoethanol, pH 8.0). The dialyzed enzymes were stored at - 80°C. Protein concentrations were determined using the Bradford assay and a bovine serum albumin standard curve according to the manufacturer's instructions.[S4]

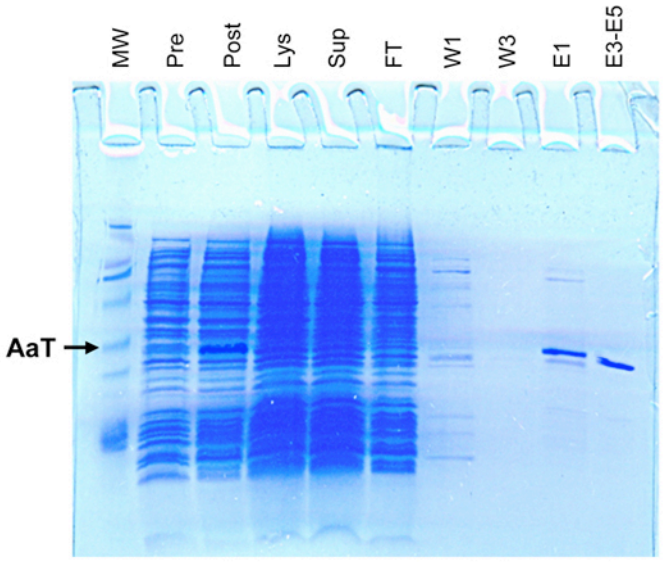


Figure S11. PAGE Gel Analysis of AaT Expression and Purification. Lanes (left to right): 1) Molecular weight markers (Masses in kDa: 16, 25, 32, 47, 80, 100, 210); 2) Preinduction; 3) Postinduction; 4) Crude cell lysate; 5) Supernatant after centrifugation; 6) Ni-NTA column flow-through; 7) Column wash 1; 8) Column wash 3; 9) Elution fraction 1; 10) Combined elution fractions 3-5 after overnight dialysis.

AaT Ligation Reaction Analysis by HPLC. All HPLC analyses of LysAlaAcm ligations were monitored as described in the main text (Gradient 1). Peptide retention time monitored at 325 nm (Acm absorption) and MALDI MS analyses, confirming peptide identity are in Table S1.

Table S1. HPLC Analysis of LysAlaAcm Transfer Reactions.

Peptide	Retention Time (min)	$(M + H)^+$ Mass Calcd, Found (m/z)
LysAlaAcm	11.8	375.2, 374.9
PheLysAlaAcm	13.0	522.3, 522.1
LeuLysAlaAcm	12.7	488.3, 487.3
AzfLysAlaAcm	13.6	563.3, 563.3 (- N_2 537.3, 537.3)
NapLysAlaAcm	14.1	572.3, 572.1
MefLysAlaAcm	13.1	536.3, 536.3
AcfLysAlaAcm	NA	NA
N_3fLysAlaAcm	NA	NA
McmLysAlaAcm	NA	NA
BzfLysAlaAcm	14.6	626.3, 626.1

Calibration of NapLysAlaAcm Fluorometric Transfer Analysis. In order to monitor conversion of LysAlaAcm to NapLysAlaAcm (**5d**) in real time, we designed a fluorometric assay based on intramolecular quenching of aminocoumarin fluorescence by the naphthyl ring. Equimolar solutions of LysAlaAcm and NapLysAlaAcm, as well as mixtures of the two solutions, were prepared and their fluorescence compared. (Fig. S12A) The fluorescence emission of NapLysAlaAcm at 390 nm (λ_{ex} = 325 nm) was 72% lower than LysAlaAcm, and no non-linear effects were exhibited in the mixtures. (Fig. S12B) Therefore, the mole fraction of NapLysAlaAcm could be linearly related to the measured fluorescence of a reaction solution ($F_{\mathbf{5d}}$), and the concentration of **5d** determined as (1 - ($F_{\mathbf{5d}}/F_{Control}$))•1.39•[LysAlaAcm], where $F_{Control}$ is the fluorescence reading of the reaction with water added instead of donor **4d**. This calibration was used to determine the time-dependent concentration of the NapLysAlaAcm product as described in the main text.

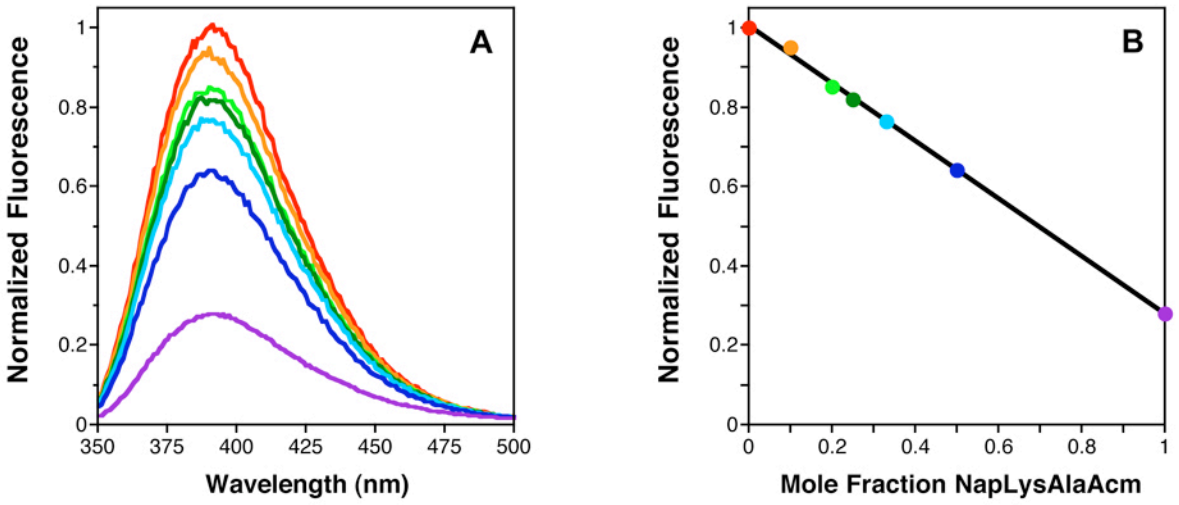


Figure S12. Fluorescence Emission of NapLysAlaAcm Solutions. (A) Fluorescence emission of mixtures of LysAlaAcm and NapLysAlaAcm (**5d**); $\chi_{\mathbf{5d}}$ = 0.00 (red), 0.10 (orange), 0.20 (light green), 0.25 (green), 0.33 (light blue), 0.50 (blue), 1.00 (purple). (B) Linear fit of normalized fluorescence vs. mole fraction NapLysAlaAcm (y = 1 – 0.72x, R = 0.999)

Michaelis-Menten Kinetic Analysis of Phe Transfer. The reaction scheme for the catalysis of *N*-terminal aminoacylation by AaT is shown in Equation S1. Previously, it has been shown that the sequential bisubstrate AaT reaction can be treated as pseudo-first order when carried out with saturating concentrations of one substrate (acceptor peptide).[S5] Therefore, we characterized the reaction in terms of the Michaelis-Menten kinetic scheme in Equation S2.

$$\text{Phe-A + LysAlaAcm + AaT} \rightleftharpoons \text{Phe-A•LysAlaAcm•AaT} \longrightarrow \text{PheLysAlaAcm + A + AaT} \quad \text{(S1)}$$

$$\text{Phe-A + LysAlaAcm•AaT} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \text{Phe-A•LysAlaAcm•AaT} \xrightarrow{k_{cat}} \text{PheLysAlaAcm + A + AaT} \quad \text{(S2)}$$

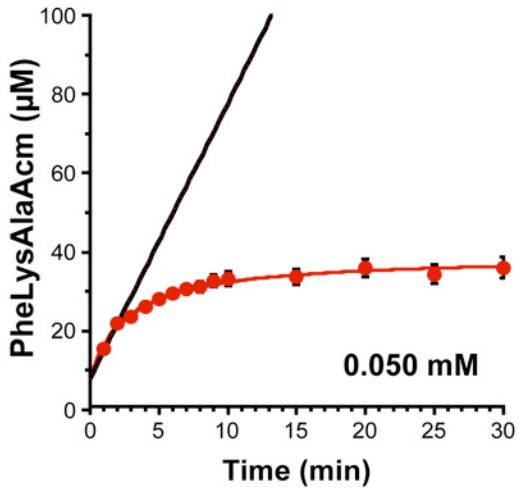
To determine initial velocities, we fit the primary kinetic data to the hyperbolic Equation S3, which, though phenomenological, produces a high quality fit ($R > 0.99$ for all 5 data sets). Data from the first two minutes were fit to a linear expression (Eqn. S4). We employ a non-zero intercept for improved quality of fit; this can be justified in terms of a rapid burst phase not captured in the 1 min resolution of this assay. Note that this burst phase seems to be extremely rapid, even the 15 s intervals of our direct fluorometric assay are not sufficient to capture this phase. (See Fig. 3A in the main text) The results of these linear fits are shown in Figure S13.

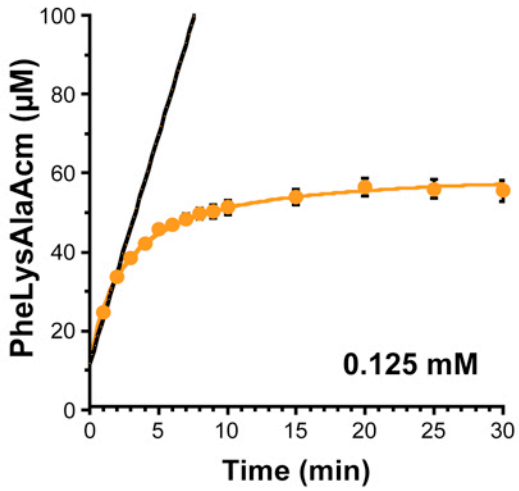
$$[\text{PheLysAlaAcm}] = c_1 t/(c_2 + t) \quad \text{or} \quad [\mathbf{5a}] = c_1 t/(c_2 + t) \quad \text{(S3)}$$

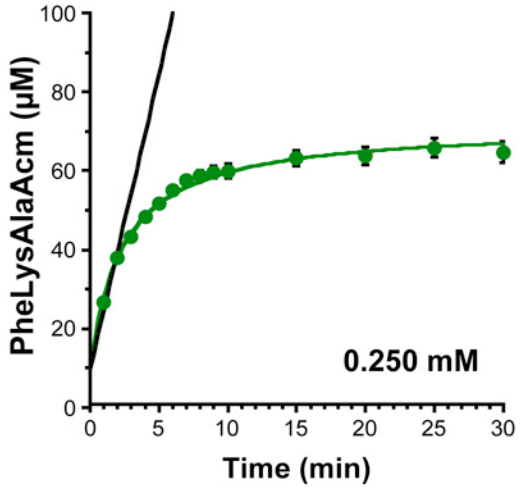
$$[\text{PheLysAlaAcm}] = [\text{PheLysAlaAcm}]_0 + V_0 t \quad \text{or} \quad [\mathbf{5a}] = [\mathbf{5a}]_0 + V_0 t \quad \text{(S4)}$$

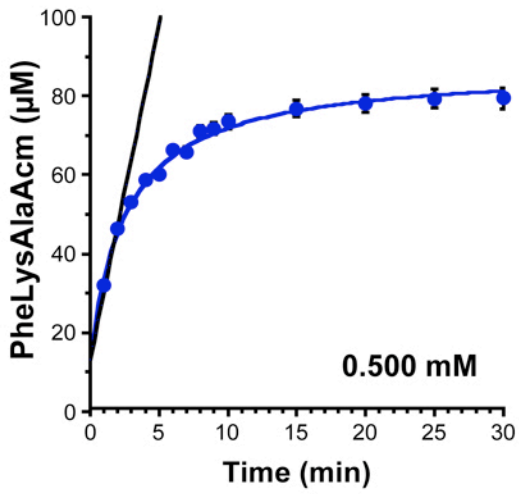
The initial velocities were then plotted as a function of Phe-A donor concentration and fit to Equation S5 to obtain k_{cat} and K_M values, where $K_M = (k_{-1} + k_{cat})/k_1$. (See Fig. 3B in the main text)

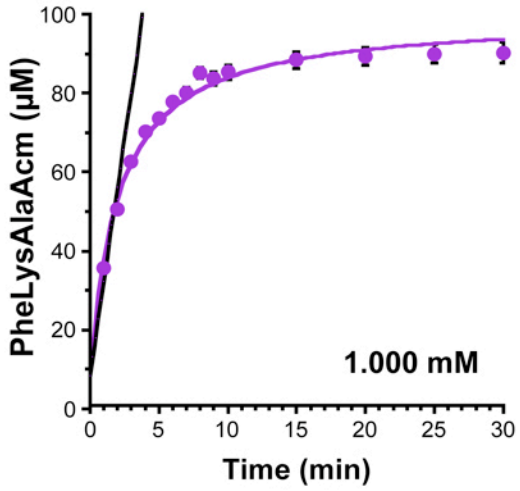
$$V_0 = k_{cat}[\text{AaT}][\text{Phe-A}]/(K_M + [\text{Phe-A}]) \quad \text{or} \quad V_0 = k_{cat}[\text{AaT}][\mathbf{4a}]/(K_M + [\mathbf{4a}]) \quad \text{(S5)}$$











[**4a**] (mM)	V_0 (μM•min^{-1})	[**5a**]$_0$ (μM)
0.050	7.34 ± 0.46	7.90 ± 1.39
0.125	11.08 ± 0.86	11.75 ± 3.39
0.250	14.17 ± 0.54	9.75 ± 2.11
0.500	17.00 ± 0.82	13.24 ± 2.83
1.000	20.70 ± 0.85	8.43 ± 4.27

Figure S13. Phe Transfer Reaction Kinetics. Product concentration as a function of time determined by integration of HPLC traces. Data are an average of 3 trials. Bars indicate standard error. Black lines indicate best fit to early timepoints using Eqn. S4. Bottom Right: Fits to kinetic data according to Eqn. S4.

Edman Degradation Analysis of α-Casein Modification. Phe was ligated to α-casein using **4a** as described in the main text with the following modifications. The substrate (1 mM) was 10 times more concentrated, and the total volume was 4 times greater. This reaction was run overnight; at the 4 h timepoint, an additional 50 μL dose of 25 mM **4a** donor was added to the reaction. The His_{10}-tagged AaT was separated from the reaction via nickel bead purification: 100 µL Ni-NTA resin was added at a ratio of 25 μL per 12.5 μM AaT, the beads were shaken with the reaction for 2 hours, then separated *via* centrifugation at 13, 200 RPM for 2 min. The supernatant was removed, diluted to 1 mL with Milli-Q water, and dialyzed against 1X PBS (Hyclone, Fisher) overnight at 4 °C to remove any residual **4a**. The protein was electroblotted onto PVDF membrane and sent to Johns Hopkins University Synthesis and Sequencing Facility (Baltimore, MD) for Edman Degradation analysis.

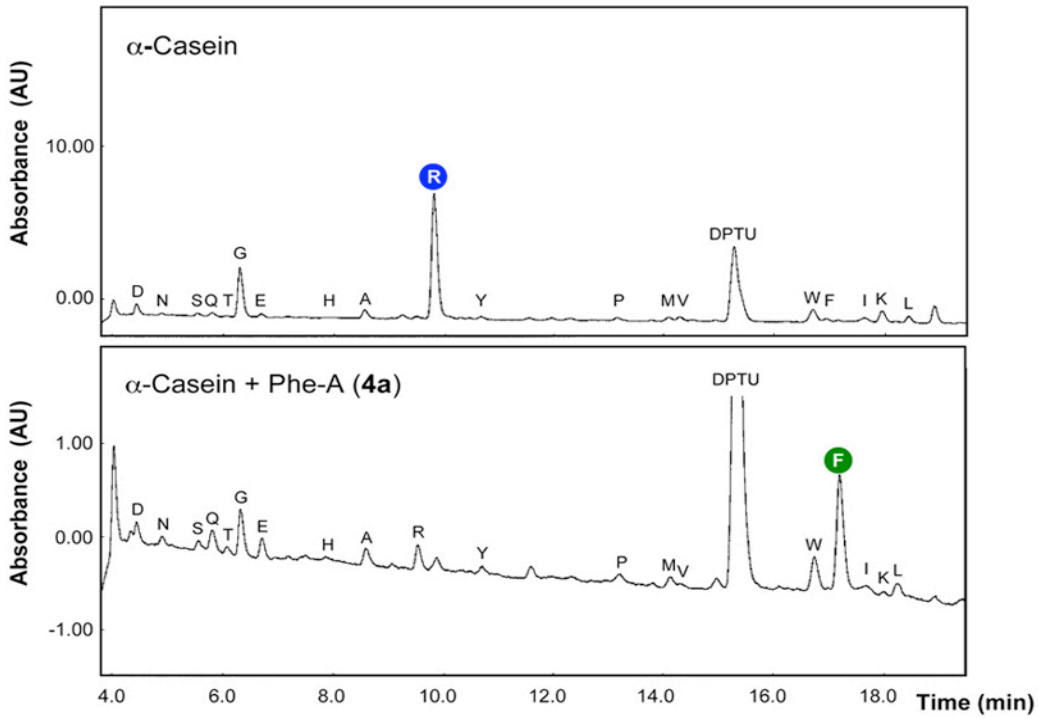


Figure S14. Edman Degradation Analysis of α-Casein Modification. Top: Unmodified α-casein. R peak area is 196 counts of total, F peak area is 44 counts. Bottom: Phe-modified α-casein. F peak area is 97 counts, R peak area is 53 counts.

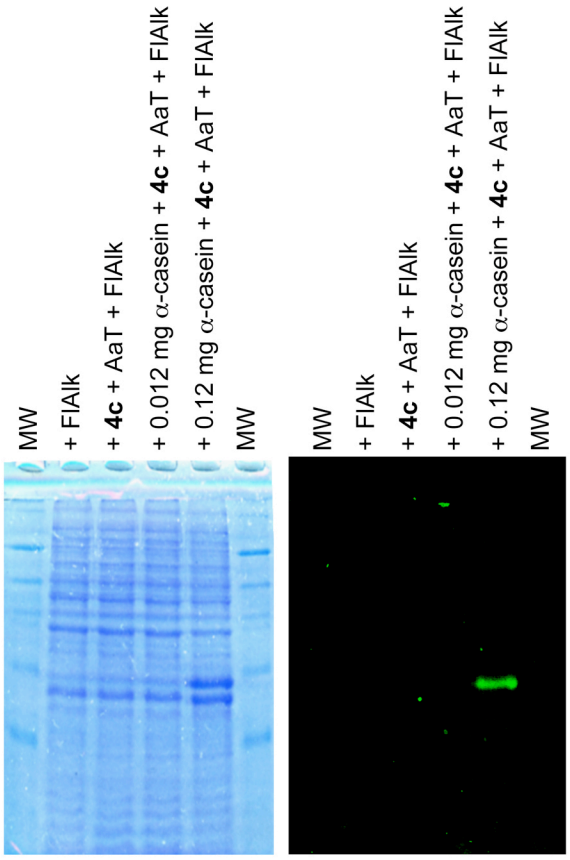


Figure S15. PAGE gel analysis of α-casein modification in cleared *E. coli* lysate. Left: Coomassie-stained gel. Right: Fluorescence image from 302 nm excitation. Lanes (left to right): 1) Molecular weight markers (Masses in kDa: 17, 25, 30, 46, 58, 80, 175); 2) Cleared lysate reacted with FlAlk; 3) Cleared lysate modified with **4c** then reacted with FlAlk; 4) Cleared lysate with α-casein (0.012 mg) modified with **4c** then reacted with FlAlk; 5) Cleared lysate with α-casein (0.12 mg) modified with **4c** then reacted with FlAlk; 6) Molecular weight markers.

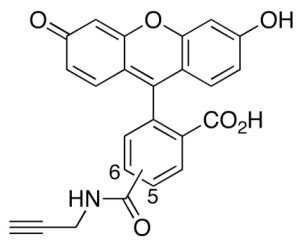
"Click" Reagent Synthesis. Fluorescein 5- and 6-proparglyamide (Fluorescein-alkyne, FlAlk) was synthesized as previously described.[S6] Briefly, propargylamine (8.7 μL, 136 μmol) was added to a solution of 5- and 6-Carboxyfluorescein *N*-hydroxysuccinimide ester (32 mg, 68 μmol) in tetrahydrofuran (5 mL). The reaction was stirred continuously and monitored by TLC. Solvent was removed by under reduced pressure. Purification by silica chromatography (R_f = 0.3, 20% MeOH in CH_2Cl_2) afforded fluorescein-alkyne as a red-orange solid.

Tris-(3-hydroxypropyltriazolylmethyl)amine (THPTA) was also synthesized as previously described.[S7] Briefly, to a solution of 3-bromo-propanol (2 g, 14.4 mmol) in dichloromethane (10 mL) was added Ac_2O (2.94 g, 28.8 mmol) and NEt_3 (2.91 g, 28.8 mmol). The reaction was stirred continuously and monitored by TLC. An aqueous solution of $NaHCO_3$ was added and the phases were separated. The organic layer was washed once more with $NaHCO_3$ and twice with brine. Solvent was removed under reduced pressure and 3-bromopropyl acetate was afforded as a colorless oil.

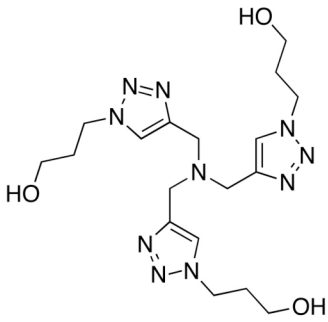
To a solution of 3-bromopropyl acetate (1.0 g, 6.0 mmol) in water (10 mL) was added sodium azide (780 mg, 12.1 mmol). The solution was stirred at 90 °C overnight. The solution was extracted three times with 15 mL CH_2Cl_2 and dried with $MgSO_4$. Solvent was removed under reduced pressure to afford 3-azidopropyl acetate as a pale yellow oil. To a solution of tripropargylamine (12.3 mg, 0.941 mmol) in tetrahydrofuran (6.5 mL) was added 3-azidopropyl acetate (0.672 g, 4.71 mmol) and Cu(I) acetate (5.8 mg, 5 mol%). The solution was refluxed overnight under inert atmosphere. Solvent was removed under reduced pressure. Silica chromatography purification (R_f = 0.5, 5-10% MeOH in CH_2Cl_2) afforded acetyl-protected THPTA as a yellow oil.

A solution of acetyl-protected THPTA (397 mg, 0.711 mmol) in 2.0 M ammonia in MeOH (10 mL) was continuously stirred overnight at 40 °C. Solvent was removed under reduced pressure. The solid was washed and filtered four times with acetonitrile and dried under reduced pressure to afford THPTA as a white solid.

For both FlAlk (25 % overall yield) and THPTA (98 % overall yield), intermediate and product identities were confirmed by comparison of ^{1}H NMR and LRMS (ESI) MS data to previous reports.[S6,S7]



Fluorescein 5- and 6-Proparglyamide
FlAlk



Tris(hydroxypropyltriazoylmethyl)amine
THPTA

Figure S16. Click Chemistry Reagents.

References

(S1) Taiji, M.; Yokoyama, S.; Miyazawa, T. *Biochemistry* **1983**, *22*, 3220-3225.
(S2) Watanabe, K.; Toh, Y.; Suto, K.; Shimizu, Y.; Oka, N.; Wada, T.; Tomita, K. *Nature* **2007**, *449*, 867-871.
(S3) Graciet, E.; Hu, R. G.; Piatkov, K.; Rhee, J. H.; Schwarz, E. M.; Varshavsky, A. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 3078-3083.
(S4) Bradford, M. M. *Anal. Biochem.* **1976**, *72*, 248-254.
(S5) Abramochkin, G.; Shrader, T. E. *J. Biol. Chem.* **1996**, *271*, 22901-22907.
(S6) Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G. *J. Am. Chem. Soc.* **2003**, *125*, 3192-3193.
(S7) Hong, V.; Presolski, S. I.; Ma, C.; Finn, M. G. *Angew. Chem., Int. Ed.* **2009**, *48*, 9879-9883.

Expanded References from Main Text

(13) Arduini, R. M.; Li, Z. F.; Rapoza, A.; Gronke, R.; Hess, D. M.; Wen, D. Y.; Miatkowski, K.; Coots, C.; Kaffashan, A.; Viseux, N.; Delaney, J.; Domon, B.; Young, C. N.; Boynton, R.; Chen, L. L.; Chen, L. Q.; Betzenhauser, M.; Miller, S.; Gill, A.; Pepinsky, R. B.; Hochman, P. S.; Baker, D. P. *Protein Exp. Purification* **2004**, *34*, 229-242.
(15) Baker, D. P.; Lin, E. Y.; Lin, K.; Pellegrini, M.; Petter, R. C.; Chen, L. L.; Arduini, R. M.; Brickelmaier, M.; Wen, D. Y.; Hess, D. M.; Chen, L. Q.; Grant, D.; Whitty, A.; Gill, A.; Lindner, D. J.; Pepinsky, R. B. *Bioconj. Chem.* **2006**, *17*, 179-188.