Chemistry & Biology, Volume *20*

Supplemental Information

Engineering the Substrate Specificity of the DhbE Adenylation Domain by Yeast Cell Surface Display

Keya Zhang, Kathryn M. Nelson, Karan Bhuripanyo, Kimberly D. Grimes, Bo Zhao, Courtney C. Aldrich, and Jun Yin

Inventory of the Supplemental Information

Supplemental Figures

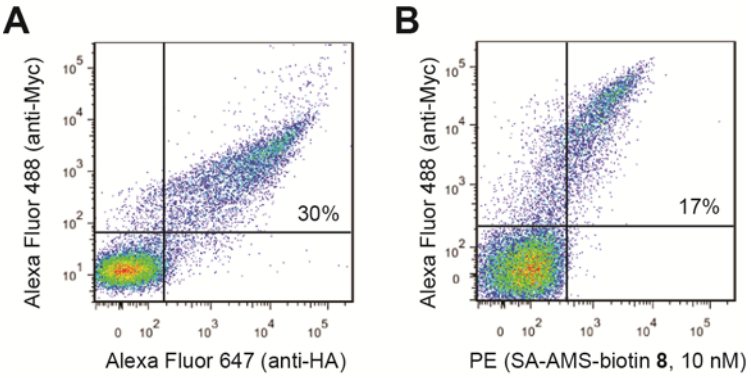


Figure S1. (Related to Figure 1) Flow cytometry analysis of wtDhbE displayed on the surface of yeast cells. (A) Yeast cells displaying DhbE were labeled with chicken anti-myc antibody and mouse anti-HA antibody followed by the detection with goat anti-chicken antibody-Alexa Fluor 488 and goat anti-mouse antibody-Alexa Fluor 647 conjugates. (B) Yeast cells displaying wtDhbE were labeled with chicken anti-myc antibody and biotin linked SA-AMS **8**, followed by the detection with goat anti-chicken antibody-Alexa Fluor 488 and streptavidin-PE conjugates. Percentages of double positive cells that were labeled with two fluorophores were shown on the flow cytometry plots.

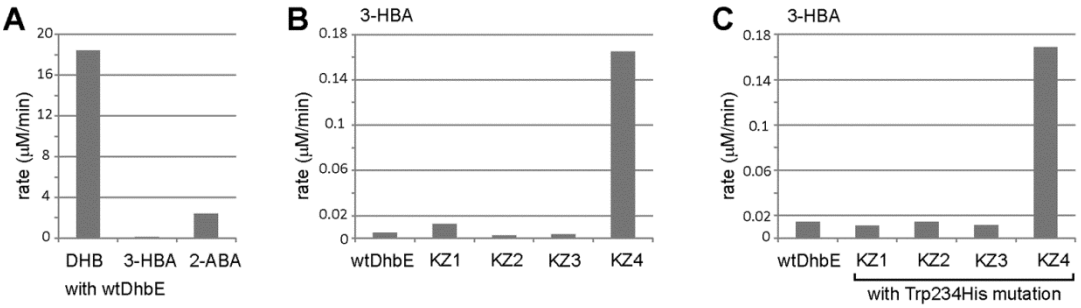


Figure S2. (Related to Figure 1) ATP-PPi exchange assay on wtDhbE and the mutants. (A) Activity of wtDhbE with DHB, 3-HBA and 2-HBA. (B) Comparison of the activities of wtDhbE and mutants KZ1-4 from yeast selection with 3-HBA as the substrate. (C) Comparison of the activities of wtDhbE and mutants KZ1-4 with the Trp234His mutation with 3-HBA as the substrate.

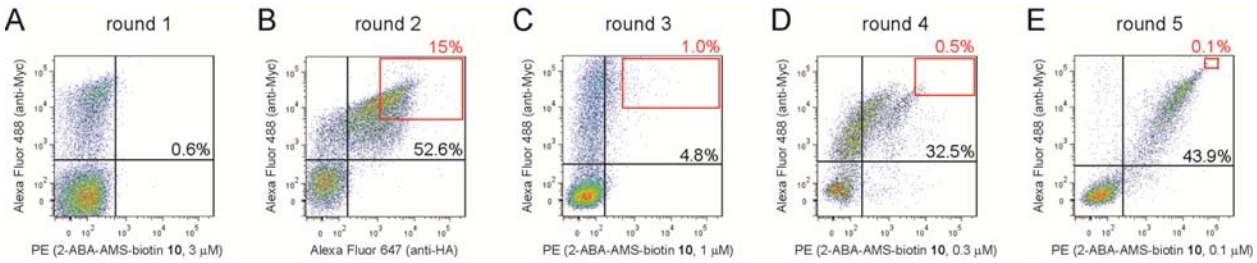


Figure S3. (Related to Figure 1) Sorting of the DhbE library displayed on the yeast cell surface by binding to biotin-linked 2-ABA-AMS **10**. (A) Sorting with magnetic beads coated with streptavidin. Binding of cells to **10** and an anti-Myc antibody was analyzed by flow cytometry. (B) FACS sorting by antibody binding to the HA and Myc tags flanking DhbE on the cell surface. (C) – (E), FACS sorting by binding to **10** and an anti-Myc antibody. Percentages of doubly labeled cells were shown in the flow cytometry plots. Red frames in the plots represent the sorting gates used for the collection of yeast cells.

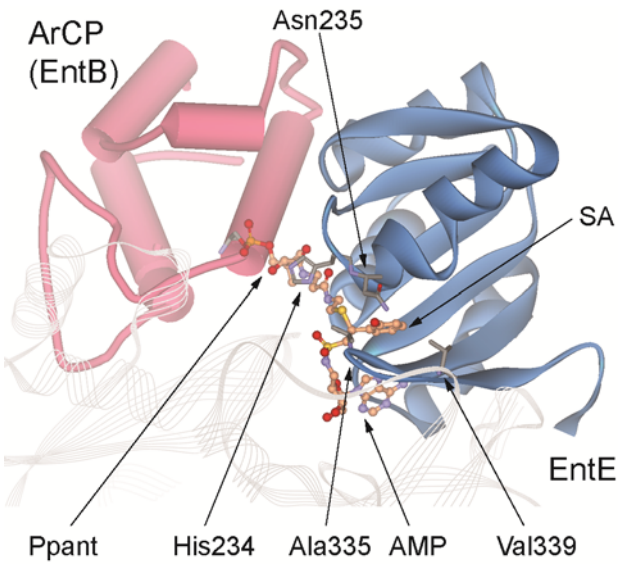


Figure S4. (Related to Figures 1 and 2) Crystal structure of EntE in complex with EntB PCP (PDB ID 3RG2) (Sundlov, et al., 2012). The core domain of EntE for substrate and ATP binding is shown as blue ribbons. The rest of the EntE structure is shown in grey ribbons. The ArCP domain of EntB is shown in red tubes. His234, Asn235, Ala335 and Val339 in EntE correspond to the residues randomized in the DhbE library.

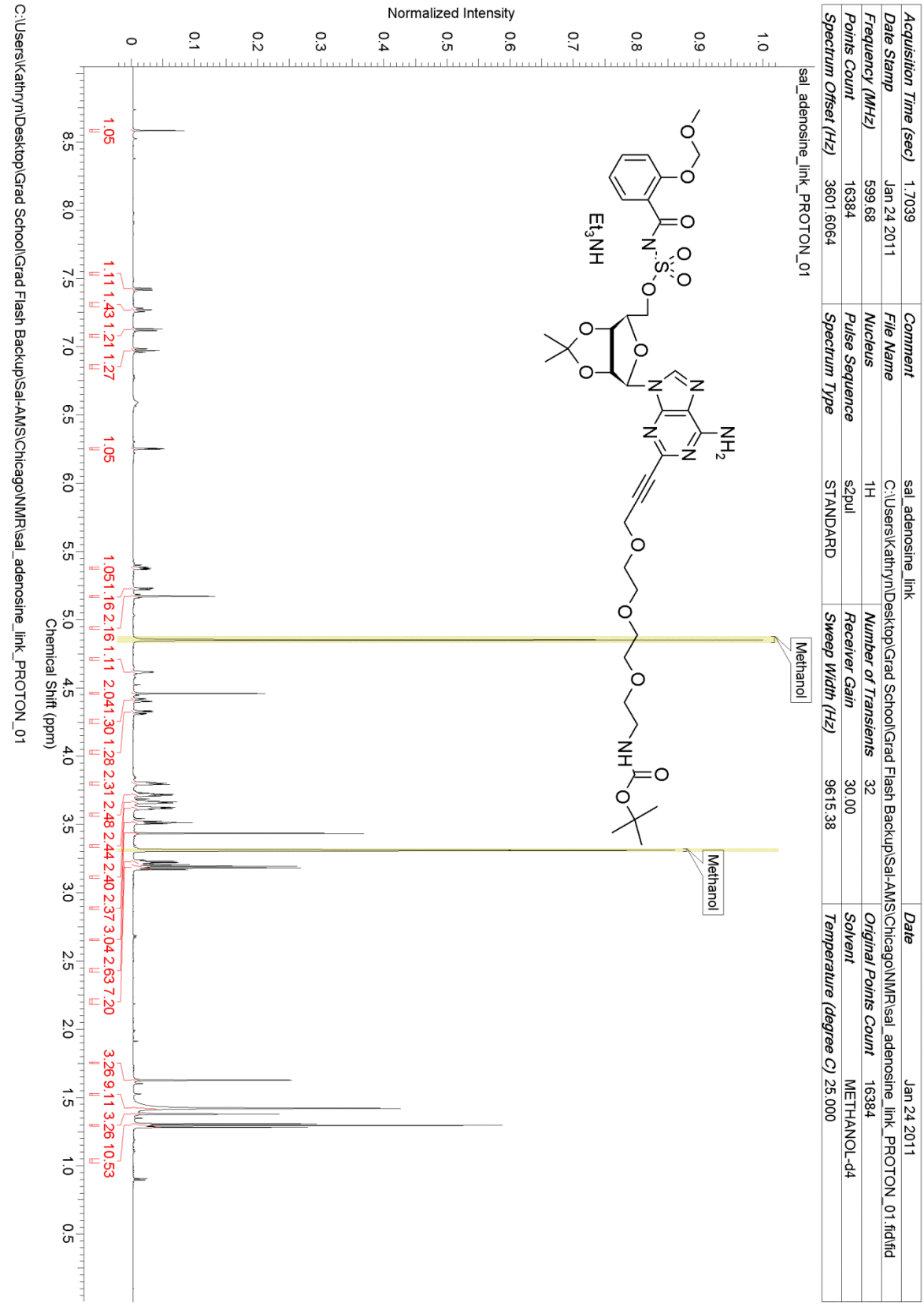
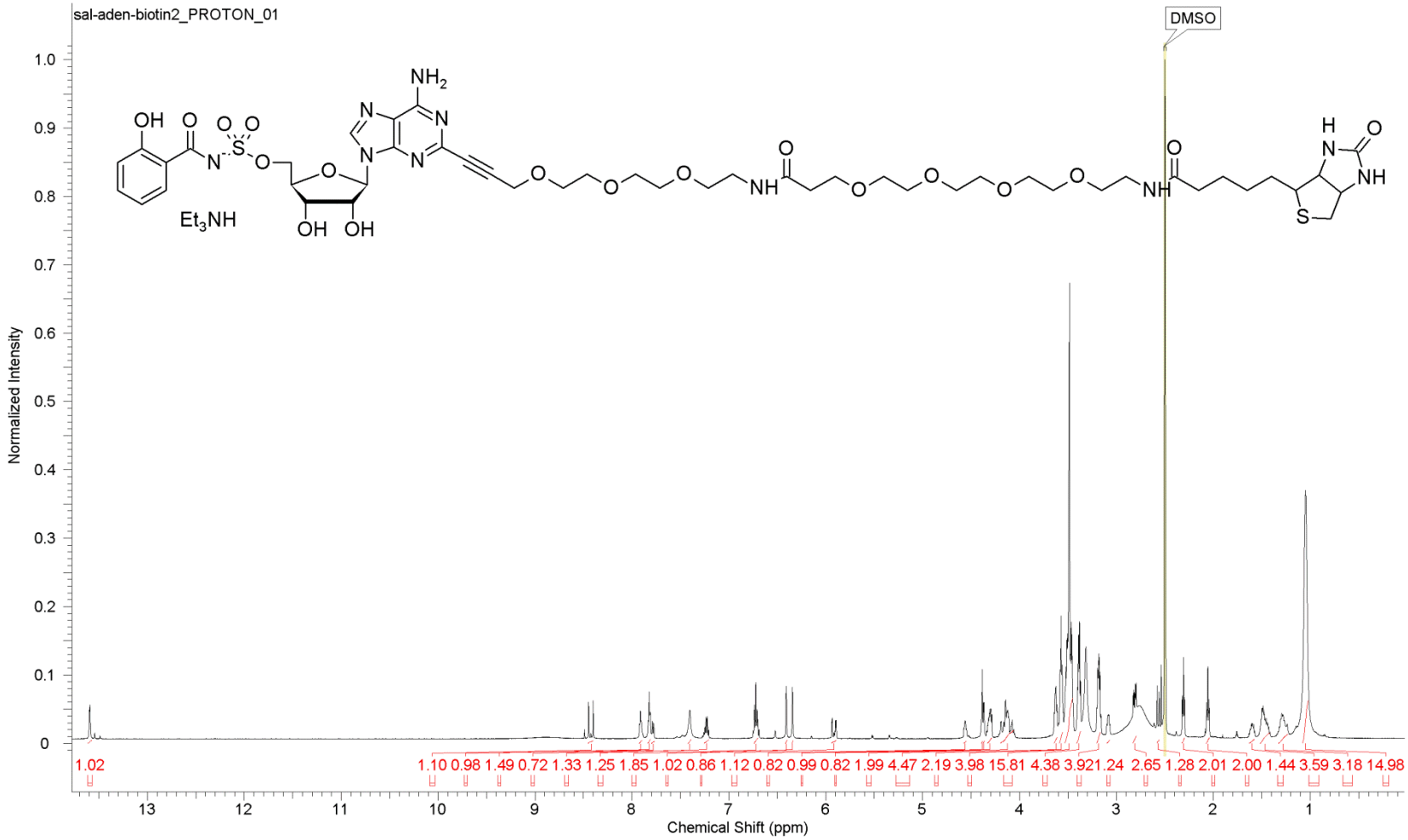


Figure S5. (Related to Figure 1) ^{1}H NMR of S3.

Figure S6. (Related to Figure 1) ^{13}C NMR of S3.

Acquisition Time (sec)	1.7039	*Comment*	sal-aden-biotin2	*Date*	Feb 4 2011	*Date Stamp*	Feb 4 2011
File Name	C:\Users\Kathryn\Desktop\Grad School\Grad Flash Backup\Sal-AMS\Chicago\NMR\sal-aden-biotin2_PROTON_01.fid\fid						
Frequency (MHz)	599.69	*Nucleus*	1H	*Number of Transients*	64	*Original Points Count*	16384
Points Count	16384	*Pulse Sequence*	s2pul	*Receiver Gain*	30.00	*Solvent*	DMSO-d6
Spectrum Offset (Hz)	3608.2822	*Spectrum Type*	STANDARD	*Sweep Width (Hz)*	9615.38	*Temperature (degree C)*	25.000



C:\Users\Kathryn\Desktop\Grad School\Grad Flash Backup\Sal-AMS\Chicago\NMR\sal-aden-biotin2_PROTON_01

Figure S7. (Related to Figure 1) ^{1}H NMR of 8.

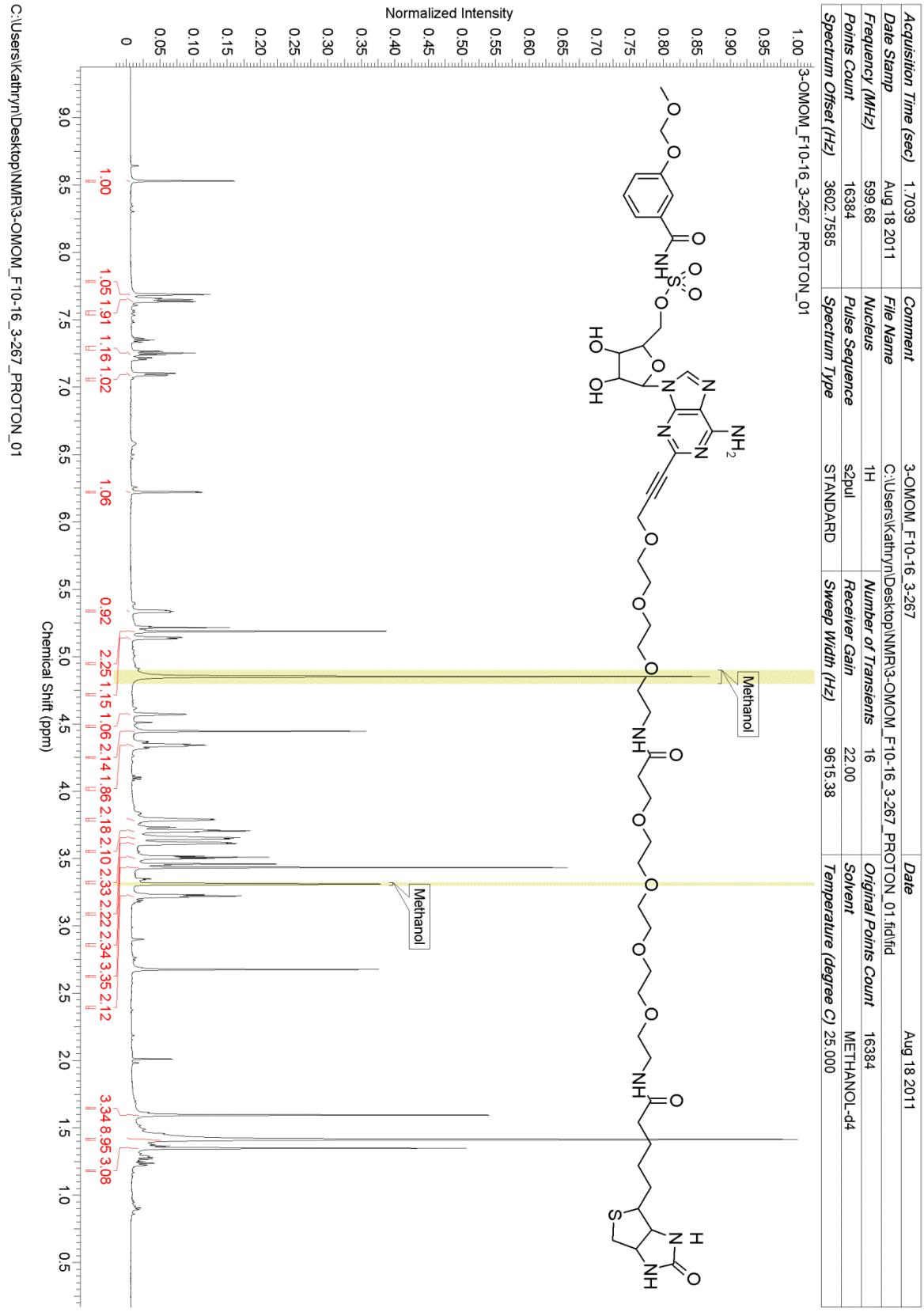


Figure S8. (Related to Figure 1) ^{1}H NMR of S7.

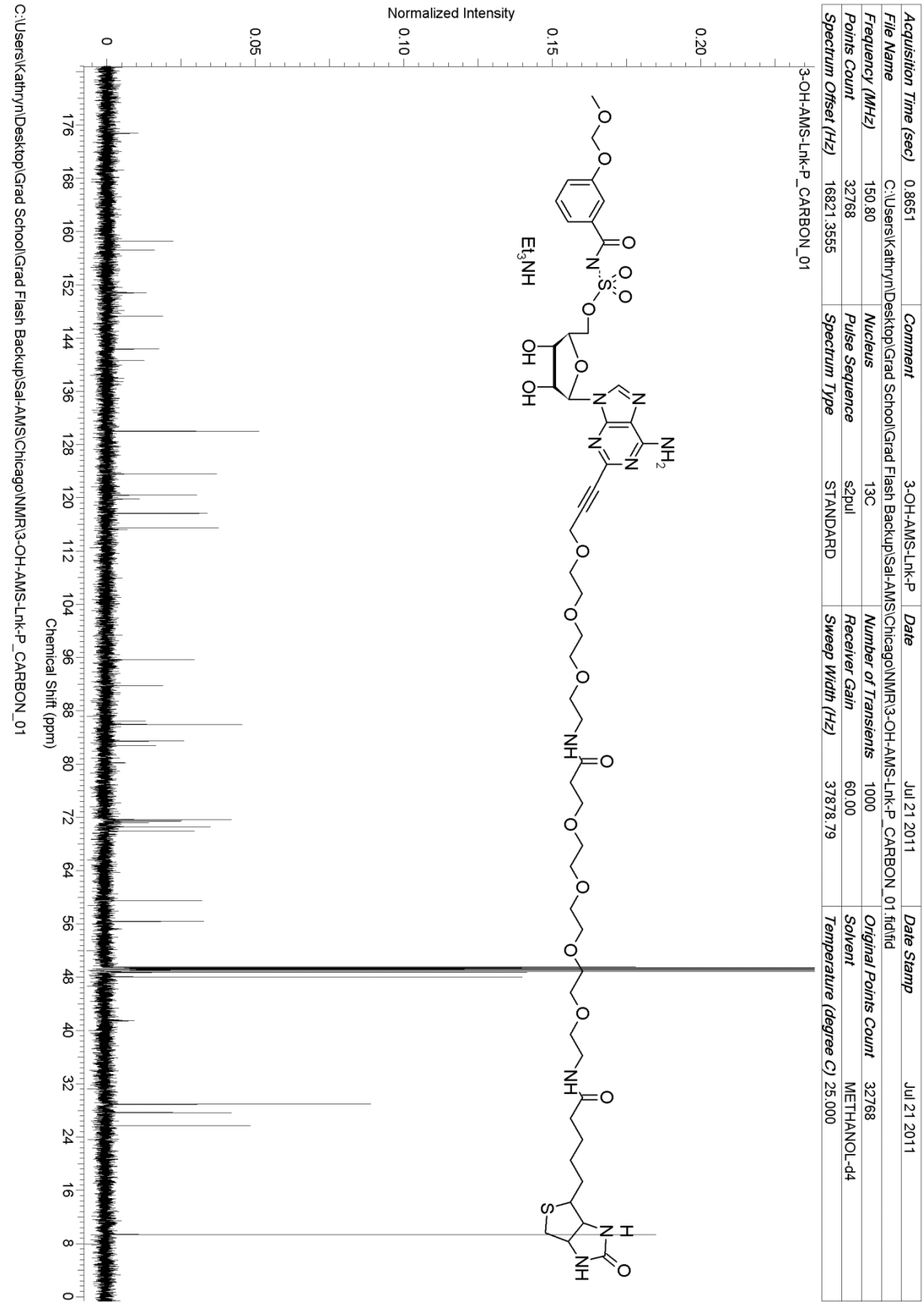


Figure S9. (Related to Figure 1) ^{13}C NMR of S7.

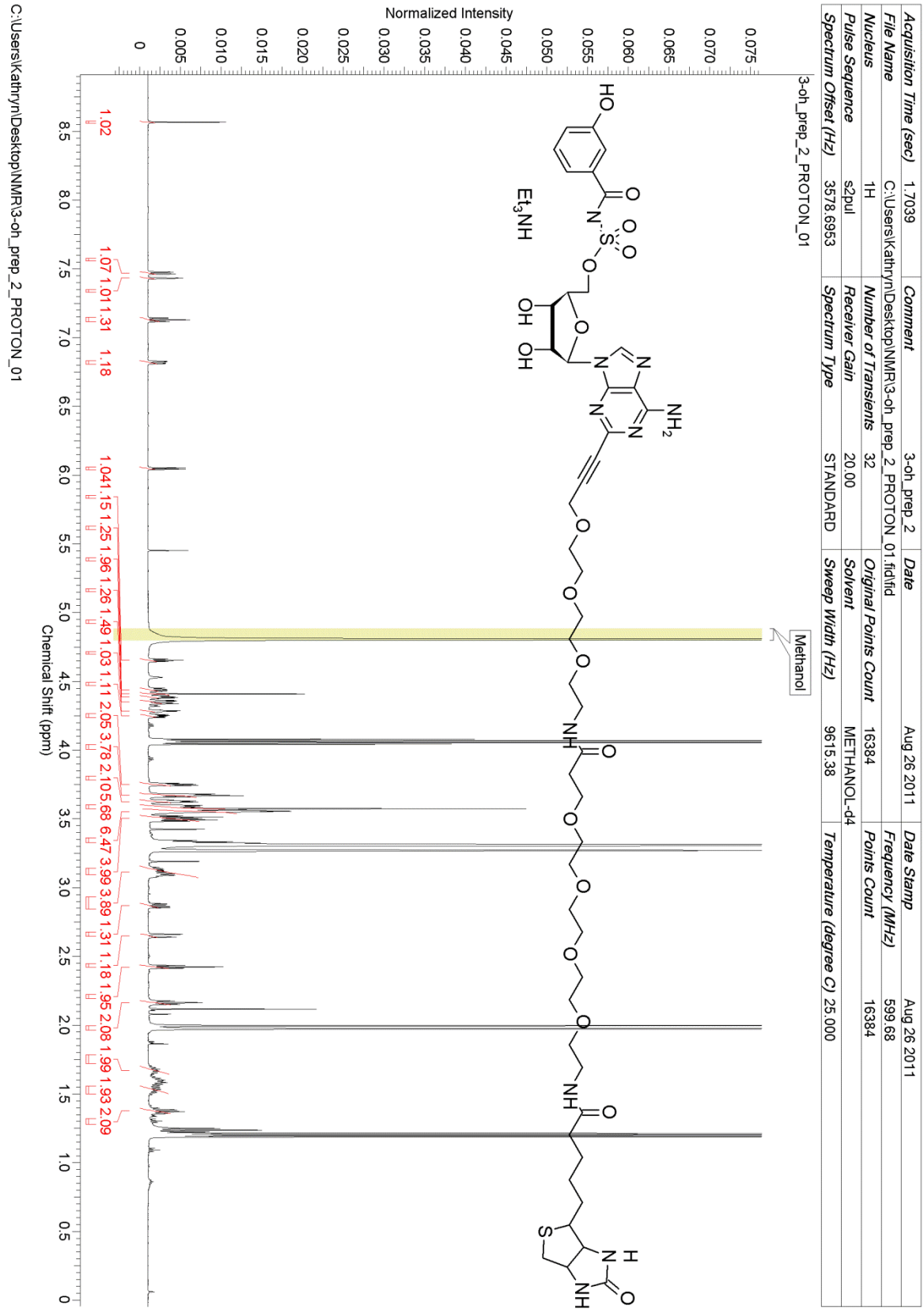


Figure S10. (Related to Figure 1) ^{1}H NMR of 9.

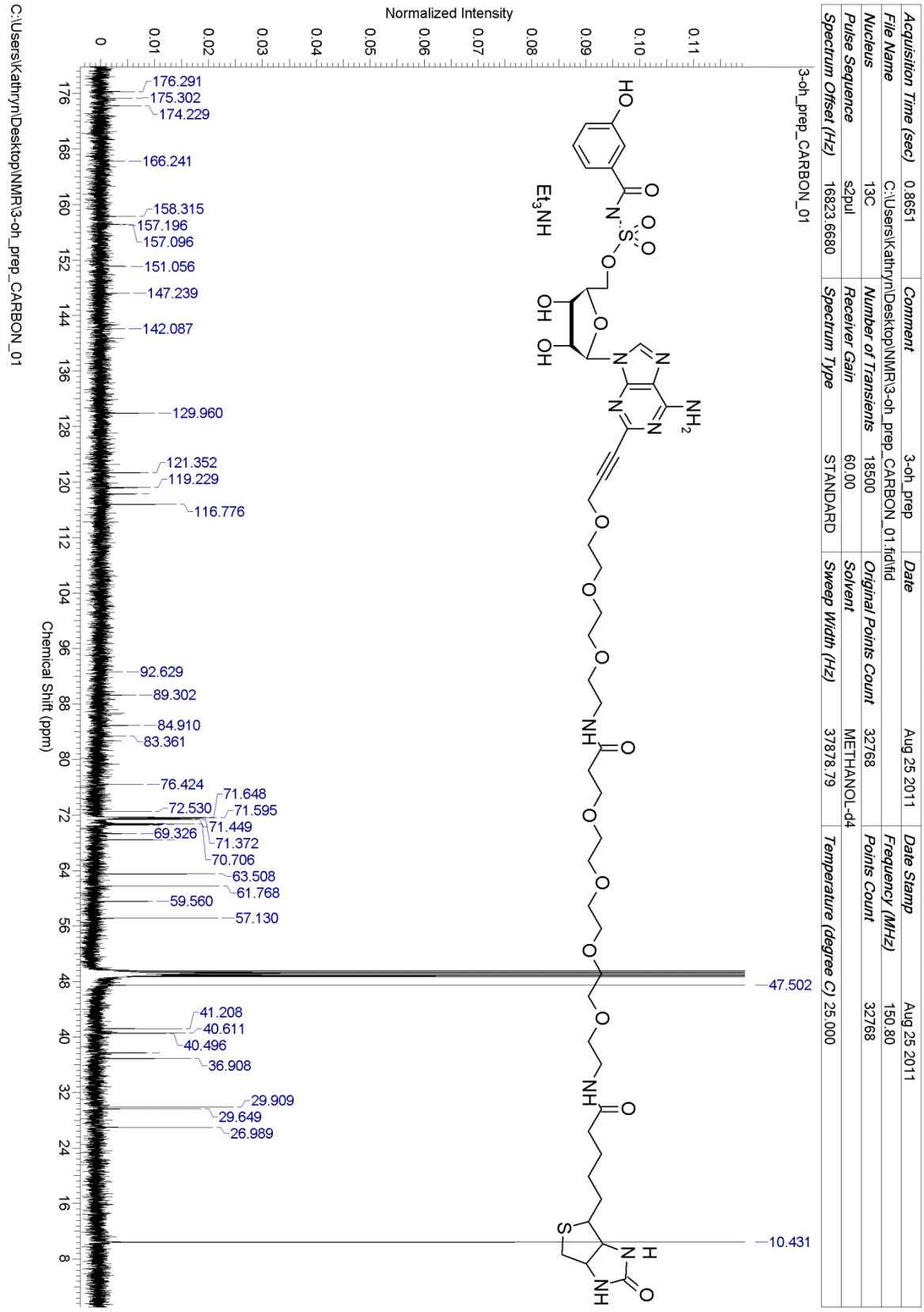


Figure S11. (Related to Figure 1) $^{/}$C NMR of 9.

Acquisition Time (sec)	1.7039	*Comment*	amino-ams-lnk	*Date*	Jul 22 2011	*Date Stamp*	Jul 22 2011
File Name	C:\Users\Kathryn\Desktop\NMR\20110722_amino-ams-lnk_01\amino-ams-lnk_PROTON_01.fid\fid					*Frequency (MHz)*	599.68
Nucleus	1H	*Number of Transients*	8	*Original Points Count*	16384	*Points Count*	16384
Pulse Sequence	s2pul	*Receiver Gain*	18.00	*Solvent*	METHANOL-d4	*Spectrum Offset (Hz)*	3601.5847
Spectrum Type	STANDARD	*Sweep Width (Hz)*	9615.38	*Temperature (degree C)*	25.000		

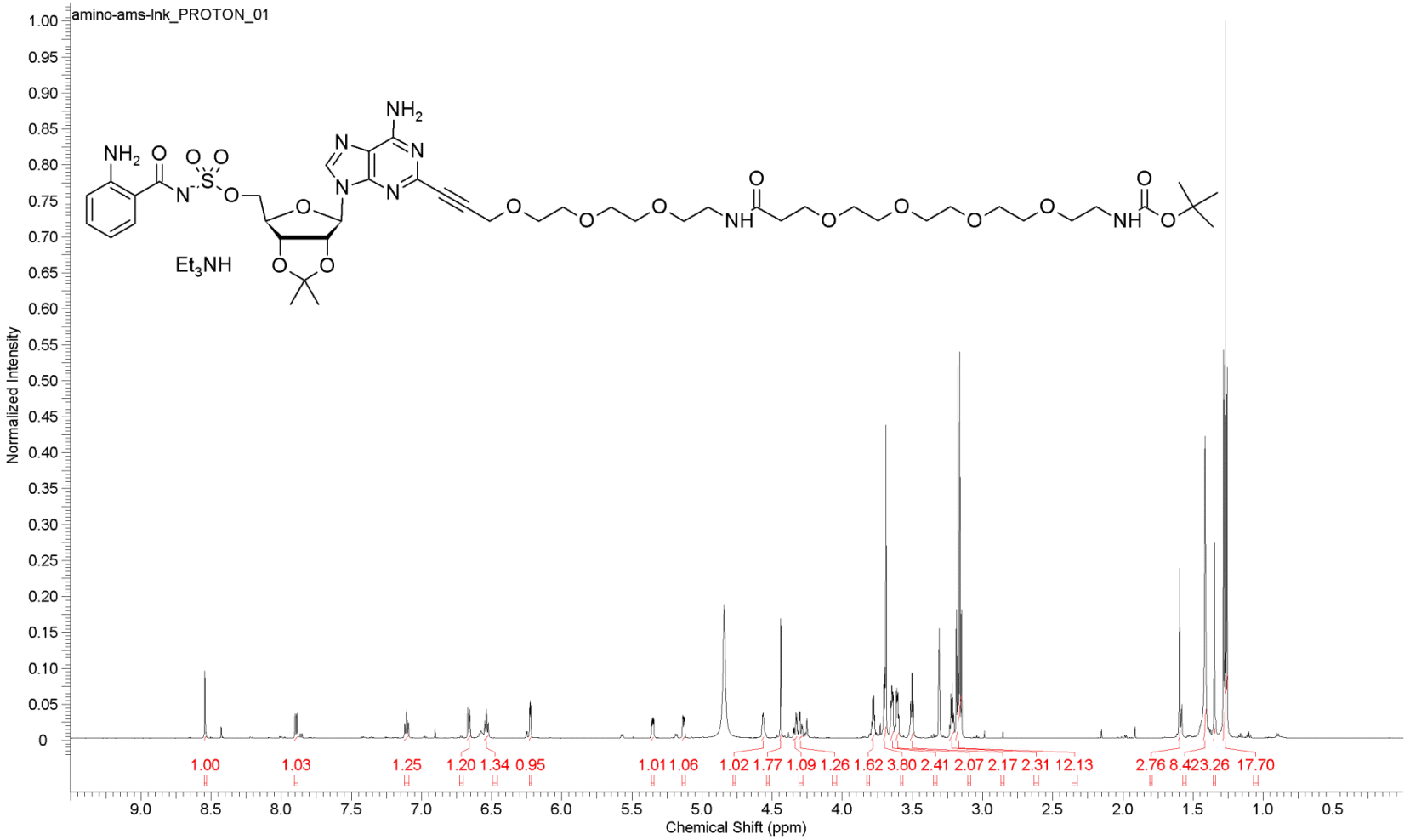


Figure S12. (Related to Figure 1) ^{1}H NMR of S9.

Acquisition Time (sec)	1.7039	*Comment*	amino-ams-lnk	*Date*	Jul 22 2011	*Date Stamp*	Jul 22 2011
File Name	C:\Users\Kathryn\Desktop\NMR\20110722_amino-ams-lnk_01\amino-ams-lnk_PROTON_01.fid\fid					*Frequency (MHz)*	599.68
Nucleus	1H	*Number of Transients*	8	*Original Points Count*	16384	*Points Count*	16384
Pulse Sequence	s2pul	*Receiver Gain*	18.00	*Solvent*	METHANOL-d4	*Spectrum Offset (Hz)*	3601.5847
Spectrum Type	STANDARD	*Sweep Width (Hz)*	9615.38	*Temperature (degree C)*	25.000		

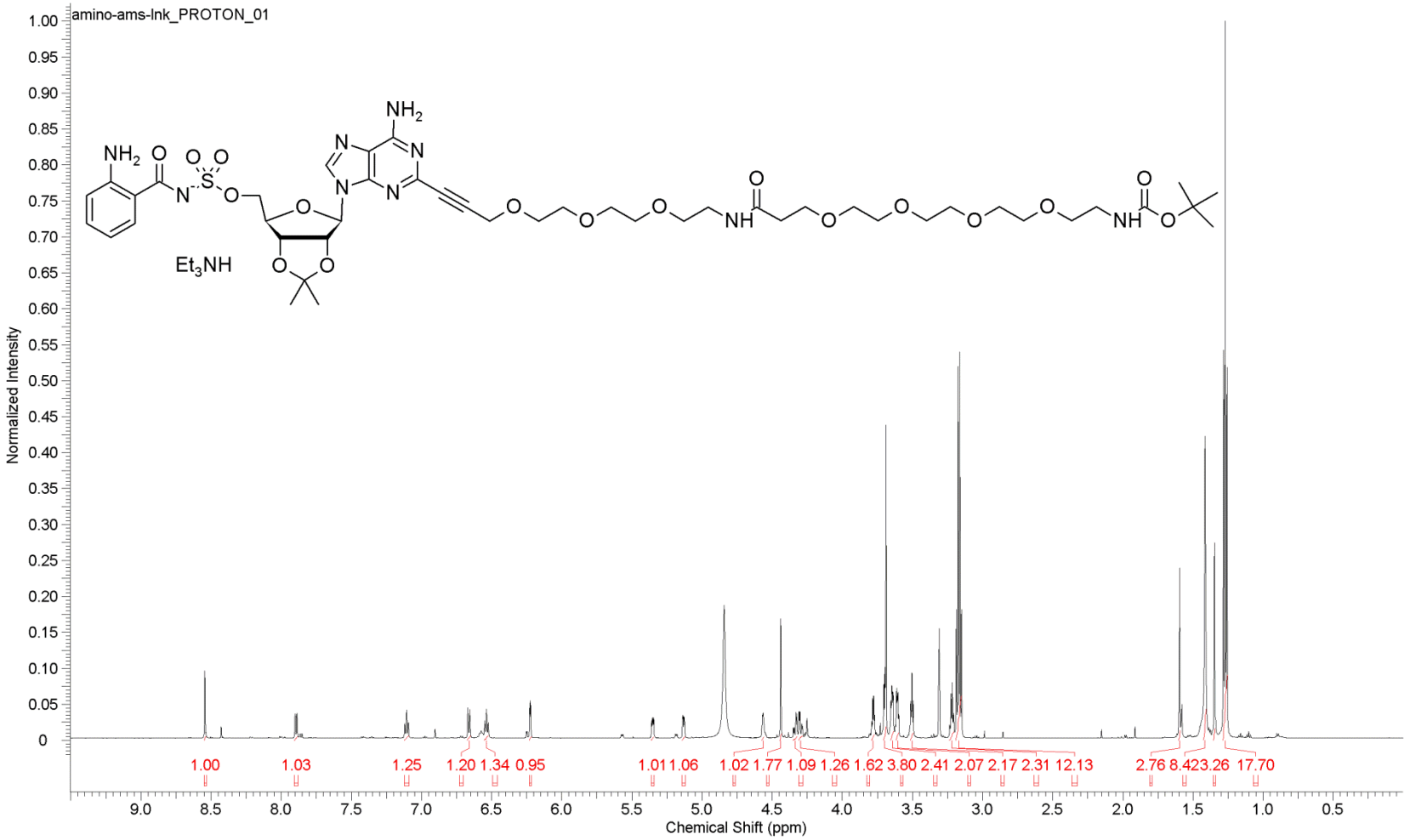


Figure S13. (Related to Figure 1) ^{13}C NMR of S9.

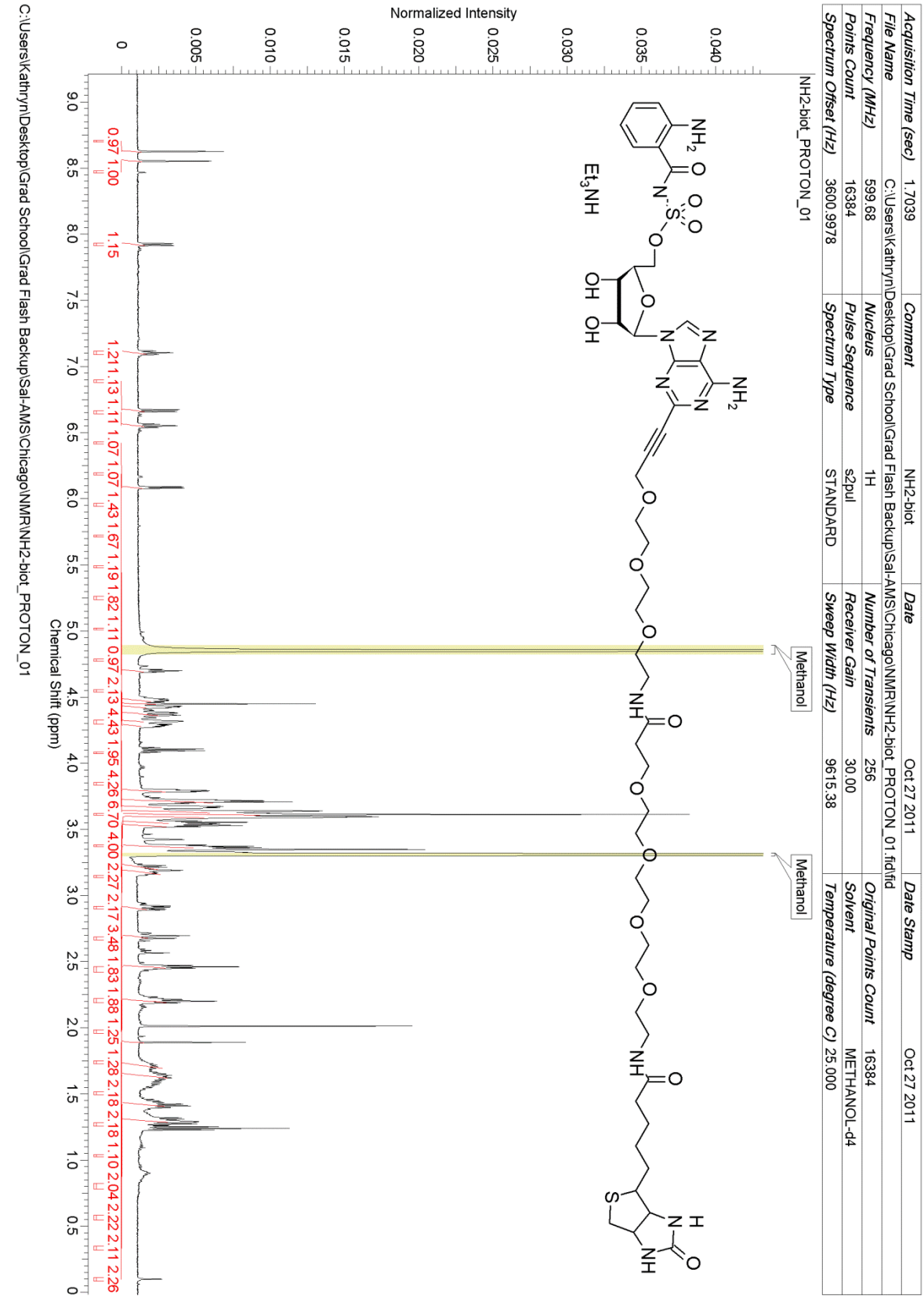


Figure S14. (Related to Figure 1) ^{1}H NMR of 10.

Supplemental Tables

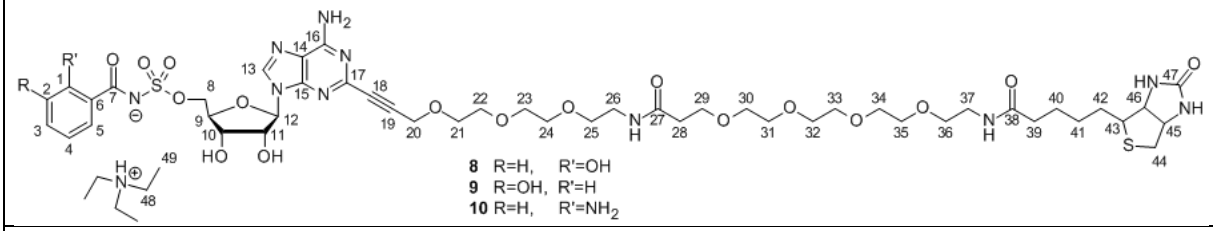
Table S1. (Related to Table 1) Kinetics of the ATP-PPi exchange reaction catalyzed by DhbE and mutants from yeast selection.

Enzyme	Substrate	K_m (μM)	k_{cat} (min^{-1})	k_{cat}/K_m ($min^{-1}\mu M^{-1}$)
wtDhbE	DHB	2.5±0.4	8.0±0.2	3200
	SA	20±2	2.3±0.1	110
KZ4	3-HBA	4.0±1	0.36±0.02	90
KZ4 (Trp234His)	DHB	4.6±0.8	0.52±0.02	110
	SA	52±6	0.84±0.03	16
	3-HBA	9.2±0.3	0.36±0.01	39
KZ12 (Trp234His)	DHB	37±6	0.28±0.02	7.6
	SA	39±5	0.23±0.01	5.9
	2-ABA	8.7±1	0.04±0.01	4.6

Table S2. (Related to Table 2) DNA oligonucleotides used in this study.

Name	Sequence
Jun287	5'- CAGGCGGCCGGTACCCTGAAAGGATTTACGCCTTGG
Jun290	5'- GGCGGCCGGGAATTCTGGAGATTCTCTTTCGTGGCTGCTTAGGCTT TTGAATGGTGGTACCCTGAAAGGATTTACGCCTTGG
Jun292	5'- GCCAAGCT ACTAGTTTTTTTAAATCCTGCAAGAAGC
Jun326	5'- GGC GGA TCG CAT ATG GCA AAC ACT GGC AAA AAG AAC G
Keya1	5'- CAGGCGGCCGGTACCCTGAAAGGATTTACGCCTTGG
Keya2	5'- GGATCCACTAGTTTTTTTAAATCCTGCAAGAAGC
Keya3	5'- GATGACAAGCCGCGGCTGAAAGGATTTACGCCTTGG
Keya4	5'- GCGCTGCCGATGGCANNKNNKTATCCGCTGAGCTCACC
Keya5	5'- TGCCATCGGCAGCGC
Keya6	5'- GCAGGTGTTCGGAATGNNKGAGGGTCTCNNKAATTATACGAGATTGG
Keya7	5'- CATTCCGAACACCTGC
Keya8	5'- GCGCTGCCGATGGCACACTCTTATCCGCTGAGCTCACC
Keya9	5'- GCGCTGCCGATGGCACACGCGTATCCGCTGAGCTCACC
Keya10	5'- GCGCTGCCGATGGCACACAATTATCCGCTGAGCTCACC
Keya11	5'- GCGCTGCCGATGGCACACCAGTATCCGCTGAGCTCACC

Table S3. (Related to Figure 1) ^{1}H and ^{13}C NMR assignments for compounds **8–10**.



δ_H and δ_C (ppm), mult. (*J* in Hz)

Position	8	9		10	
	δ_H	δ_H	δ_C	δ_H	δ_C (gHMBC)
1	OH = 13.59, s	7.43, s	116.8		149.59
2	7.79, d (7.8)		158.3	6.67, d (8.4)	116.6
3	7.91, t (5.4)	6.82, dd (7.2, 2.4)	118.3	7.10, t (8.4)	137.3
4	7.21–7.26, m	7.13, t (7.8)	130.0	6.55, t (7.2)	131.2
5	7.82, d (7.2)	7.47, d (7.8)	119.2	7.93, d (7.8)	119.2
6			142.1		
7			174.2		172.6
8	4.29–4.32, m	4.25, q (4.2)	63.5	4.43, dd (8.0, 3.6)	68.8
8'	4.55–4.58, m	4.44, q (4.2)		4.48, dd (7.8, 4.8)	
9	5.90, d (3.6)	4.38–4.39, m	84.9	4.29, dd (7.8, 4.8)	84.8
10	6.34, s	4.34, d (2.4)	72.5	4.32, d (2.4)	70.2
11	6.41, s	4.65, t (5.4)	76.4	4.70, t (6.0)	76.3
12	6.69–6.75, m	6.05, (5.4)	92.6	6.09, d (5.4)	104.6
13	8.45, s	8.57, s	147.2	8.63, s	140.6
14			121.4		117.8
15			151.1		149.63
16			157.2		154.1
17			157.1		153.4
18			83.4		81.1
19			89.3		99.9
20	4.38, s	4.41, s	59.6	4.45, s	57.1
21	3.37–3.40, m	3.48–3.52, m	70.5	3.52–3.56, dt (15.0, 6.0)	
22	3.37–3.40, m	3.48–3.52, m	70.6	3.52–3.56, dt (15.0, 6.0)	70.07
23	3.46–3.52, m	3.55–3.57, m	71.7	3.60–3.64, m	66.7
24	3.46–3.52, m	3.55–3.57, m	71.59	3.60–3.64, m	70.0
25	3.55–3.58, m	3.62–3.64, m	71.65	3.66–3.68, m	70.06
26	3.16–3.20, m	3.09–3.14, m	41.2	3.15–3.24, m	41.4
27			175.3		178.8
28	2.31, t (6.6)	2.42, t (6.0)	36.9	2.46, t (6.0)	35.2
29	3.46–3.52, m	3.55–3.57, m	69.3	3.60–3.64, m	60.1
30	3.46–3.52, m	3.58–3.60, m	71.7	3.60–3.64, m	61.9
31	3.46–3.52, m	3.48–3.52, m	71.56	3.60–3.64, m	69.84
32	3.46–3.52, m	3.58–3.60, m	71.4	3.60–3.64, m	69.75
33	3.55–3.58, m	3.58–3.60 m	71.4	3.60–3.64, m	69.3
34	4.06–4.21, m	3.66–3.68, m	70.80	3.70–3.72, m	69.1
35	4.06–4.21, m	3.66–3.68, m	70.78	3.70–3.72, m	69.0
36	3.61–3.64, m	3.74–3.76, m	70.7	3.79–3.80, m	67.0
37	3.16–3.20, m	3.09–3.14, m	40.6	3.15–3.24, m	39.1
38			176.3		174.7
39	2.06, t (7.2)	2.16, t (7.2)	37.7	2.20, t (7.8)	36.3
40	1.42–1.50, m	1.53–1.61, m	29.6	1.60–1.66, m	28.0
41	1.24–1.34, m	1.36–1.38, m	29.9	1.40–1.46, m	28.3
42	1.57–1.63, m	1.67–1.69, m	27.0	1.68–1.73, m	25.4
43	3.07–3.10, m	3.19, s	57.1	3.42, br s	55.5
44	2.57, d (12.0)	2.65, d (12.6)	40.5	2.68, q (12.6)	39.6
44'	2.81, dd (12.6, 5.4)	2.87, dd (12.6, 4.8)		2.91, dd (13.2, 4.8)	
45	4.06–4.21, m	4.29, d (3.0)	61.8	4.37, dd (10.8, 3.0)	58.6
46	4.06–4.21, m	4.36, d (2.4)	68.4	4.37, dd (10.8, 3.0)	60.0
47			166.2		158.2
48			47.5	3.35–3.38, m	46.5
49			10.4	1.27–1.32, m	9.8

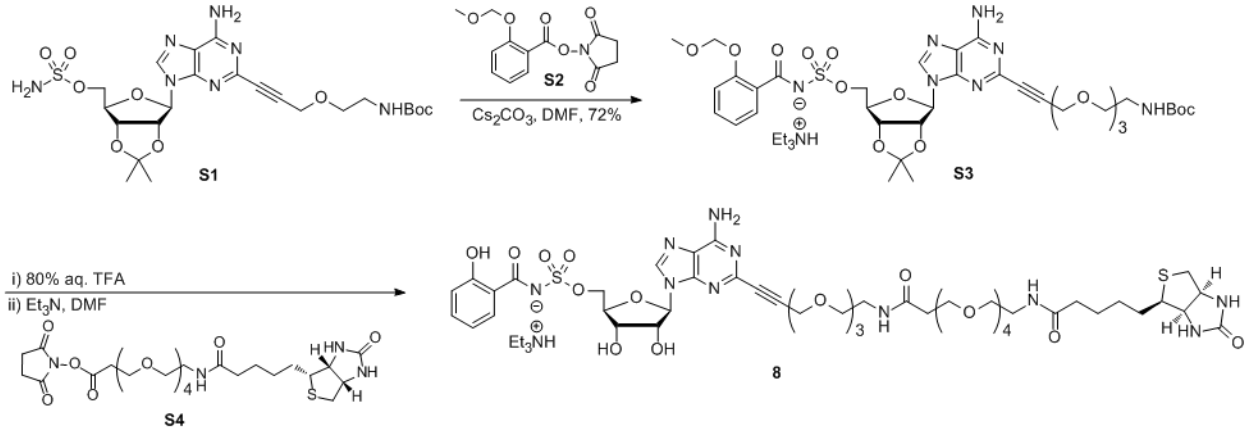
Supplemental Experimental Procedures

Chemical Synthesis (Related to Figure 1)

General methods: All reactions were performed under an inert atmosphere of dry Argon in oven-dried (150 °C) glassware. ^{1}H and ^{13}C NMR experiments were recorded on a Varian 600 MHz spectrometer. Proton chemical shifts are reported in ppm from an internal standard of residual chloroform (7.26 ppm), methanol (3.31 ppm), or dimethyl sulfoxide (2.50 ppm). Carbon chemical shifts are reported using an internal standard of residual chloroform (77.0 ppm), methanol (49.1 ppm), or dimethyl sulfoxide (39.5 ppm). Proton chemical data are reported as follows: chemical shift, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, p = pentet, m = multiplet, br = broad), coupling constant, and integration. High-resolution mass spectra were acquired on an Agilent TOF II TOF/MS instrument equipped with either an ESI or APCI interface. Semi-preparative HPLC was performed on a Phenomenex Gemini 10u C18 110A (250 × 10.0 mm) column operating at 4.0 mL/min with detection at 254 nm. Flash chromatography was performed on an ISCO Combiflash Companion® purification system with prepacked silica gel cartridges and the indicated solvent system.

Materials: All commercial reagents (Sigma-Aldrich, Fisher, Fluka, Strem) were used as provided. NHS-dPEG®$_4$-biotin was purchased from Quanta BioDesign, Ltd. (Powell, OH, USA) and 2-iodoadenosine from Tokyo Chemical Industry Co. (Portland, OR, USA). Sulfamoyl chloride was prepared by the method of Heacock without recrystallization.(Heacock, et al., 1996) Compound **S1** was prepared according to literature procedure and all spectral data agree with reported values.(Grimes and Aldrich, 2011) An anhydrous solvent dispensing system (J. C. Meyer) using 2 packed columns of neutral alumina was used for drying THF, DMF and DCM and the solvents were dispensed under Argon. Anhydrous DME (Sigma-Aldrich) was used as provided.

Scheme S1. Synthesis of probe **8**.

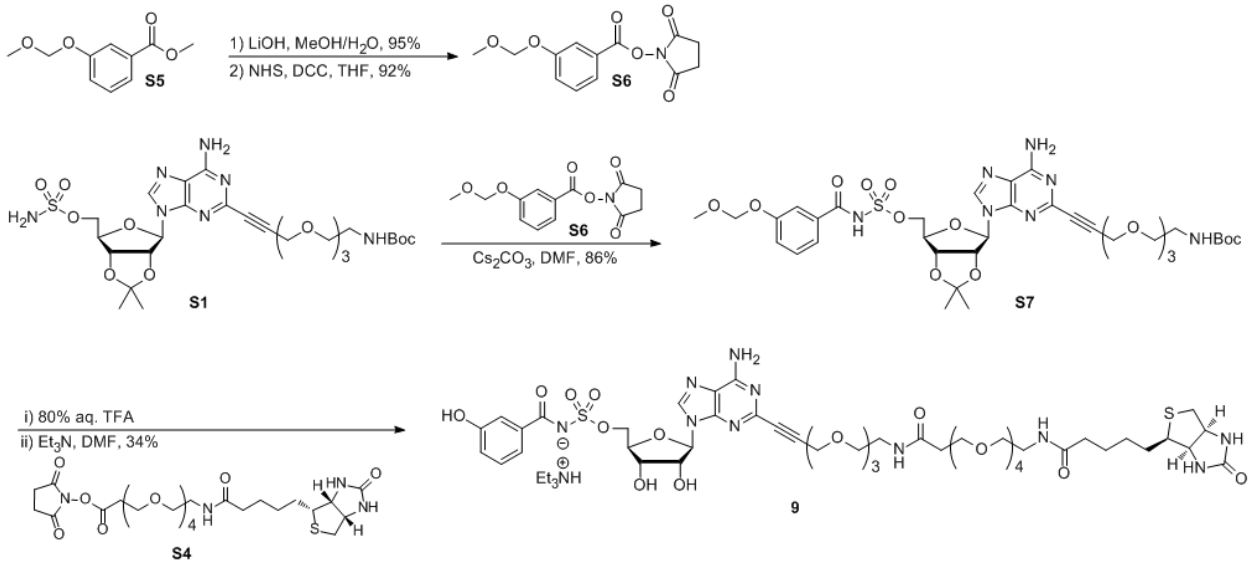


2-[3-(2-{2-[2-(*tert*-Butoxycarbonylamino)ethoxy]ethoxy}ethoxy)prop-1-ynyl]-2′,3′-*O*-isopropylidene-5′-*O*-{*N*-[2-(methoxymethoxy)benzoyl]sulfamoyl}adenosine triethylammonium salt (S3). To a solution of **S1(Grimes and Aldrich, 2011)** (67 mg, 0.10 mmol, 1.0 equiv) in DMF (5 mL) at 0 ºC was added **S2(Somu, et al., 2006)** (42 mg, 0.15 mmol, 1.5 equiv) and Cs_2CO_3 (98 mg, 0.30 mmol, 3.0 equiv) and the reaction stirred at 23 °C for 16 h. The reaction mixture was filtered and the filtrate was concentrated. Purification by flash chromatography (15% MeOH–EtOAc with 1% Et_3N) afforded the title compound (67.5 mg, 72%) as a yellow oil: R_f = 0.18 (10% MeOH–EtOAc); ^{1}H NMR (600 MHz, CD_3OD) δ 1.20 (t, J = 7.2 Hz, 9H), 1.28 (s, 3H), 1.32 (s, 9H), 1.53 (s, 3H), 3.09 (q, J = 7.2 Hz, 6H), 3.12 (q, J = 6.0

Hz, 2H), 3.34 (s, 3H), 3.41 (t, J = 5.4 Hz, 2H), 3.51–3.53 (m, 2H), 3.56–3.57 (m, 2H), 3.61–3.63 (m, 2H), 3.69–3.71 (m, 2H), 4.22 (dd, J = 11.4, 3.0 Hz, 1H), 4.31 (dd, J = 11.4, 3.0 Hz, 1H), 4.36 (s, 2H), 4.51–4.52 (m, 1H), 5.07 (s, 2H), 5.13 (dd, J = 6.0, 1.8 Hz, 1H), 5.28 (dd, J = 6.0. 3.6 Hz, 1H), 6.12 (d, J = 3.6 Hz, 1H), 6.87 (t, J = 7.8 Hz, 1H), 7.02 (d, J = 8.4 Hz, 1H), 7.17 (t, J = 7.8 Hz, 1H), 7.32 (d, J = 7.8 Hz, 1H), 8.48 (s, 1H); ^{13}C NMR (150 MHz, CD_3OD) δ 9.4, 25.9, 27.9, 29.1, 48.0, 56.8, 59.6, 70.1, 70.6, 71.2, 71.4, 71.6, 71.7, 82.8, 83.6, 85.9, 86.0, 86.4, 91.8, 96.5, 96.8, 115.5, 117.1, 122.7, 129.8, 131.2, 142.3, 147.3, 155.8, 157.3 (missing 6 carbons); MS (ESI–) calcd for $C_{36}H_{48}N_7O_{14}S$ $[M - H]^-$ 834.30, found 834.27.

2-[3-(2-{2-[2-({dPEG®$_4$-biotyl}amino)ethoxy]ethoxy}ethoxy)prop-1-ynyl]-5′-*O*-{*N*-[2-(hydroxy)benzoyl]sulfamoyl}adenosine triethylammonium salt (8). To **S2** (20 mg, 0.021 mmol, 1.0 equiv) was added aqueous 80% TFA (2.0 mL). The mixture was stirred at 23 °C for 4 h, then concentrated under reduced pressure to remove all traces of TFA. To the resulting crude residue in DMF (1 mL) was added triethylamine (0.10 mL, 0.72 mmol, 34 equiv) and **S4** (19.0 mg, 0.032 mmol, 1.5 equiv). The mixture was stirred at 23 °C for 16 h and then concentrated under reduced pressure. Purification by semi-preparative reverse-phase HPLC using a Phenomenex Gemini 10u C18 110Å (250 × 10.0 mm) column and a gradient from 5 to 65% MeCN–10 mM aqueous triethylammonium bicarbonate over 20 min followed by 65% MeCN for 3 min. The retention time of the product was 16 min and the appropriate fractions were pooled and lyophilized to afford the title compound as white solid: ^{1}H NMR (600 MHz, CD_3OD) δ 1.24–1.34 (m, 2H), 1.42–1.50 (m, 2H), 1.57–1.63 (m, 2H), 2.06 (t, J = 7.2 Hz, 2H), 2.31 (t, J = 6.6 Hz, 2H), 2.57 (d, J = 12.0 Hz, 1H), 2.81 (dd, J = 12.6, 5.4 Hz, 1H), 3.07–3.10 (m 1H), 3.16–3.20 (m, 4H), 3.37–3.40 (m, 4H), 3.46–3.52 (m, 12H), 3.55–3.58 (m, 4H), 3.61–3.64 (m, 2H), 4.06–4.21 (m, 5H), 4.29–4.32 (m, 1H), 4.38 (s, 2H), 4.55–4.58 (m, 1H), 5.90 (d, J = 3.6 Hz, 1H), 6.34 (s, 1H), 6.41 (s, 1H), 6.69–6.75 (m, 2H), 7.21–7.26 (m, 1H), 7.40 (br s, 1H), 7.79 (d, J = 7.8 Hz, 1H), 7.82 (d, J = 7.2 Hz, 1H), 7.91 (t, J = 5.4 Hz, 1H), 8.45 (s, 1H), 13.59 (br s, 1H); HRMS (ESI–) calcd for $C_{47}H_{67}N_{10}O_{18}S_2$ $[M - H]^-$ 1123.4082, found 1123.4048 (error 3.0 ppm).

Scheme S2. Synthesis of probe **9**.



***N*-Hydroxysuccinimidyl 3-(methoxymethyloxy)benzoate (S6).** A solution of LiOH (328 mg, 7.65 mmol, 3.0 equiv) in MeOH (27 mL) and water (3 mL) was added to methyl 3-(methoxymethyloxy)-benzoate **S5**(Winkle and Ronald, 1982) (500 mg, 2.55 mmol, 1.0 equiv), and the reaction mixture was refluxed for 2 h. The reaction mixture was concentrated, the residue was dissolved in H_2O (25 mL), and the pH was adjusted to 2 and then extracted with EtOAc (3 × 50 mL). The combined organic extracts were washed with saturated aqueous NaCl (50 mL), and then concentrated to afford 3-(methoxymethyloxy)- benzoic acid (440 mg, 95%), which was directly carried onto the next step:

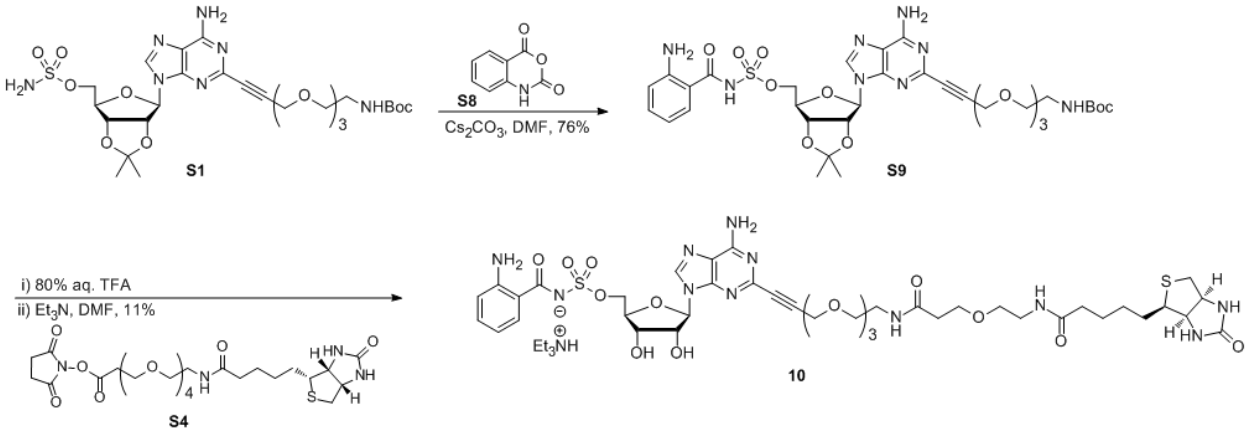
To a solution of the crude product (440 mg, 2.41 mmol, 1.0 equiv) from above in dry THF (25 mL) at 0 °C was added *N*-hydroxysuccinimide (306 mg, 2.66 mmol, 1.1 equiv) and DCC (498 mg, 2.41 mmol, 1.0 equiv). The resulting mixture was stirred for 30 min at 0 °C and then 16 h at rt. The reaction mixture was filtered to remove the DCU precipitate, and the filtrate was concentrated under reduced pressure. Purification by flash chromatography (4:1 EtOAc/hexanes) afforded the title compound (620 mg, 92%) as a viscous, colorless oil.

2-[3-(2-{2-[2-(*tert*-Butoxycarbonylamino)ethoxy]ethoxy}ethoxy)prop-1-ynyl]-2′,3′-*O*-isopropylidene-5′-*O*-{*N*-[3-(methoxymethoxy)benzoyl]sulfamoyl}adenosine (S7). To a solution of **S1(Grimes and Aldrich, 2011)** (16 mg, 0.024 mmol, 1.0 equiv) in DMF (2.5 mL) was added **S6(Somu, et al., 2006)** (10 mg, 0.036 mmol, 1.5 equiv) and Cs_2CO_3 (23 mg, 0.071 mmol, 3.0 equiv) and the reaction was stirred at 23 °C for 22 h. The reaction mixture was filtered and concentrated to a pale yellow film. Purification by flash chromatography (linear gradient 0–15% MeOH–EtOAc) gave the title compound (17.2 mg, 86%) as a colorless oil: R_f = 0.36 (10% MeOH–EtOAc); ^{1}H NMR (600 MHz, CD_3OD) δ 1.35, (s, 3H), 1.41 (s, 9H), 1.60 (s, 3H), 3.22 (t, J = 5.4 Hz, 2H), 3.43 (s, 3H), 3.51 (t, J = 5.4 Hz, 2H), 3.61–3.62 (m, 2H), 3.64–3.65 (m, 2H), 3.70–3.71 (m, 2H), 3.78–3.79 (m, 2H), 4.34 (t, J = 4.8 Hz, 2H), 4.45 (s, 2H), 4.57 (s, 1H), 5.14 (d, J = 5.4 Hz, 1H), 5.19 (s, 2H), 5.34 (q, J = 3.6 Hz, 1H), 6.22 (d, J = 2.4 Hz, 1H), 7.10 (d, J = 8.4 Hz, 1H), 7.25 (t, J = 7.2 Hz, 1H), 7.64 (d, J = 7.8 Hz, 1H), 7.69 (s, 1H), 8.53 (s, 1H); ^{13}C NMR (150 MHz, CD_3OD) δ 25.7, 27.7, 28.9, 41.4, 56.4, 59.5, 70.0, 70.6, 71.2, 71.4, 71.5, 71.7, 80.1, 82.8, 83.5, 86.0, 86.5, 91.8, 95.7, 115.5, 117.7, 119.8, 120.4, 123.6, 130.0, 140.6, 142.3, 147.3, 150.8, 157.2, 158.5, 174.7 (missing 2 carbons); HRMS (ESI−) calculated for $C_{36}H_{48}N_7O_{14}S$ $[M - H]^-$ 834.2985, found 834.2965 (error 2.4 ppm).

2-[3-(2-{2-[2-({dPEG®$_4$-biotyl}amino)ethoxy]ethoxy}ethoxy)prop-1-ynyl]-5′-*O*-{*N*-[3-(hydroxy)benzoyl]sulfamoyl}adenosine triethylammonium salt (9). To **S5** (17.2 mg, 0.021 mmol, 1.0 equiv) was added 80% aqueous TFA (2 mL). The mixture was stirred at 23 °C for 4 h, then concentrated under reduced pressure to remove all traces of TFA. To the resulting crude residue in DMF (1 mL) was added triethylamine (0.100 mL, 0.717 mmol, 34 equiv) and **S4** (18.2 mg, 0.031 mmol, 1.5 equiv). The mixture was stirred at 23 °C for 13 h, then concentrated under reduced pressure. Purification by semi-preparative reverse-phase HPLC using a Phenomenex Gemini 10u C18 110A (250 × 10.0 mm) column and a gradient of 5 to 35% MeCN–10 mM aqueous triethylammonium bicarbonate over 20 minutes, followed by 35% for 8 minutes. The retention time of the product was 18.5 minutes and the appropriate fractions were pooled and lyophilized to afford the title compound (8 mg, 34%) as an off-white solid: ^{1}H NMR (600 MHz, CD_3OD) δ 1.36–1.38 (m, 2H), 1.53–1.61 (m, 2H), 1.67–1.69 (m, 2H), 2.16 (t, J = 7.2 Hz, 2H), 2.42 (t, J = 6.0 Hz, 2H), 2.65 (d, J = 12.6 Hz, 1H), 2.87 (dd, J = 12.6, 4.8 Hz, 1H), 3.09–3.14 (m, 4H), 3.19 (s, 1H), 3.48–3.52 (m, 4H), 3.55–3.57 (m, 6H), 3.58–3.60 (m, 6H), 3.62–3.64 (m, 2H), 3.66–3.68 (m, 4H), 3.74–3.76 (m, 2H), 4.25 (q, J = 4.2 Hz, 1H), 4.29 (d, J = 3.0 Hz, 1H), 4.34 (d, J = 2.4 Hz, 1H), 4.36 (d, J = 3.0 Hz, 1H) 4.38–4.39 (m, 1H), 4.41 (s, 2H), 4.44 (q, J = 4.2 Hz, 1H), 4.65 (t, J = 5.4 Hz, 1H), 6.05 (d, J = 5.4 Hz, 1H), 6.82 (dd, J = 7.8, 2.4 Hz, 1H), 7.13 (t, J = 7.8 Hz, 1H), 7.43 (s, 1H), 7.47 (d, J = 7.8 Hz, 1H), 8.57 (s, 1H); ^{13}C NMR (150 MHz, CD_3OD) δ 10.4, 27.0, 29.6, 29.9, 36.9, 37.7, 40.5, 40.6, 41.2, 47.5, 57.1, 59.6, 61.8, 63.5, 68.4, 69.3, 70.5, 70.6, 70.7, 70.78, 70.80, 71.37, 71.45, 71.56, 71.60, 71.65, 71.69, 71.73, 72.5, 76.4, 83.4, 84.9, 89.3, 92.6, 116.8, 118.3, 119.2, 121.4, 130.0,

142.1, 147.2, 151.1, 157.1, 157.2, 158.3, 166.2, 174.2, 175.3, 176.3; HRMS (ESI−) calculated for $C_{47}H_{67}N_{10}O_{18}S_2$ $[M - H]^-$ 1123.4082, found 1123.4126 (error 3.9 ppm).

Scheme S3. Synthesis of probe **10**.



2-[3-(2-{2-[2-(*tert*-Butoxycarbonylamino)ethoxy]ethoxy}ethoxy)prop-1-ynyl]-2′,3′-*O*-isopropylidene-5′-*O*-{*N*-[2-(amino)benzoyl]sulfamoyl}adenosine triethylammonium salt (S9). To a solution of **S1(Grimes and Aldrich, 2011)** (16 mg, 0.02 mmol, 1.0 equiv) in DMF (3 mL) at 23 ºC was added **S8** (8 mg, 0.05 mmol, 2.5 equiv) and Cs_2CO_3 (23 mg, 0.07 mmol, 3.5 equiv) and the reaction stirred for 16 h. The reaction mixture was filtered and the filtrate concentrated. Purification by flash chromatography (linear gradient 0–20% MeOH–EtOAc with 1% Et_3N) afforded the title compound (12.1 mg, 76%) as a yellow oil: R_f = 0.20 (10% MeOH/EtOAc); ^{1}H NMR (600 MHz, CD_3OD) δ 1.27 (t, J = 7.2 Hz, 9H), 1.35 (s, 3H), 1.41 (s, 9H), 1.60 (s, 3H), 3.17 (q, J = 7.8 Hz, 6H), 3.22 (q, J = 6.0 Hz, 2H), 3.51 (t, J = 6.0 Hz, 2H), 3.60–3.61 (m, 2H), 3.64–3.65 (m, 2H), 3.69–3.70 (m, 2H), 3.78 (q, J = 4.2 Hz, 2H), 4.30 (dd, J = 10.8, 3.6 Hz, 1H), 4.34 (dd, J = 10.8, 3.6 Hz, 1H), 4.44 (s, 2H), 4.57 (d, J = 2.4 Hz, 1H), 5.13 (dd, J = 6.0, 1.8 Hz, 1H), 5.35 (dd, J = 6.0. 3.0 Hz, 1H), 6.23 (d, J = 3.6 Hz, 1H), 6.54 (t, J = 7.8 Hz, 1H), 6.66 (d, J = 8.4 Hz, 1H), 7.11 (t, J = 7.2 Hz, 1H), 7.89 (d, J = 7.8 Hz, 1H), 8.55 (s, 1H); ^{13}C NMR (150 MHz, CD_3OD) δ 9.4, 25.7, 27.7, 28.9, 41.4, 48.0, 55.3, 59.5, 69.9, 70.6, 71.2, 71.4, 71.5, 71.6, 80.2, 82.8, 83.5, 85.9, 86.0, 86.4, 91.8, 115.5, 116.9, 118.1, 119.8, 120.2, 132.7, 133.2, 142.4, 147.3, 150.8, 151.4, 157.2, 176.9.

2-[3-(2-{2-[2-({dPEG®$_4$-biotyl}amino)ethoxy]ethoxy}ethoxy)prop-1-ynyl]-5′-*O*-{*N*-[2-(amino)benzoyl]sulfamoyl}adenosine triethylammonium salt (10). To **S9** (12.1 mg, 0.015 mmol, 1.0 equiv) was added 80% aqueous TFA (2 mL) and the mixture stirred at 23 ºC for 2.5 h, then concentrated under reduced pressure to remove all traces of TFA. The crude residue was dissolved in DMF (2 mL) and triethylamine (0.006 mL, 0.046 mmol, 3.0 equiv) and **S4** (13 mg, 0.023 mmol, 1.5 equiv) were added. The mixture was stirred at 23 ºC for 16 h, then concentrated to a yellow oil. Purification by semi-preparative reverse-phase HPLC using a Phenomenex Gemini 10u C18 110Å (250 × 10.0 mm) column and a gradient of 5 to 35% MeCN–10 mM aqueous triethylammonium bicarbonate over 20 min followed by 35% MeCN for 3 min. The retention time of the product was 19.8 minutes and the appropriate fractions were pooled and lyophilized to afford the title compound (2.1 mg, 11%) as an off-white solid: ^{1}H NMR (600 MHz, CD_3OD) δ 1.27–1.32 (m, 9H), 1.40–1.46 (m, 2H), 1.60–1.66 (m, 2H), 1.68–1.73 (m, 2H), 2.20 (t, J = 7.8 Hz, 2H), 2.46 (t, J = 6.0 Hz, 2H), 2.68 (q, J = 12.6 Hz, 1H), 2.91 (dd, J = 13.2, 4.8 Hz, 1H), 3.15–3.24 (m, 4H), 3.35–3.38 (m, 6H), 3.42 (br s, 1H), 3.52–3.56 (dt, J = 15.0, 6.0 Hz, 4H), 3.60–3.64 (m, 14H), 3.66–3.68 (m, 2H), 3.70–3.72 (m, 4H), 3.79–3.80 (m, 2H), 4.29 (dd, J = 7.8, 4.8 Hz,

1H), 4.32 (d, J = 2.4 Hz, 1H), 4.37 (dd, J = 10.8, 3.0 Hz, 2H), 4.43 (dd, J = 8.0, 3.6 Hz, 1H), 4.45 (s, 2H), 4.48 (dd, J = 7.8, 4.8 Hz, 1H), 4.70 (t, J = 6.0 Hz, 1H), 6.09 (d, J = 5.4 Hz, 1H), 6.55 (t, J = 7.2 Hz, 1H), 6.67 (d, J = 8.4 Hz, 1H), 7.10 (t, J = 8.4 Hz, 1H), 7.93 (d, J = 7.8 Hz, 1H), 8.63 (s, 1H); ^{13}C NMR (150 MHz, CD_3OD, assigned by HMBC and HSQC) δ 9.8, 25.4, 28.0, 28.3, 35.2, 36.3, 39.1, 39.6, 41.4, 46.5, 55.5, 57.1, 58.6, 60.0, 60.1, 61.9, 66.7, 67.0, 68.8, 69.0, 69.1, 69.3, 69.75, 69.84, 70.0, 70.06, 70.07, 70.2, 76.3, 81.1, 84.8, 99.9, 104.6, 116.6, 117.8, 119.2, 131.2, 137.3, 140.6, 149.59, 149.63, 153.4, 154.1, 158.2, 172.6, 174.7, 178.8 (missing 2 carbons; 1 ethoxy linker C and 1 aromatic C); HRMS (ESI−) calculated for $C_{47}H_{68}N_{11}O_{17}S_2$ $[M - H]^-$ 1122.4242, found 1122.4344 (9.1 ppm error).

Biochemical Methods (Related to Figures 3, 4 and 5, and Tables 1 and 2)

Reagents: Unless otherwise indicated, all reagents were obtained from commercial sources and used without further purification. All solutions and buffers for yeast selection were sterilized by either autoclaving or filtration. Kits for isolating DNA plasmids were from Qiagen. Restriction endonucleases, and Ni-NTA agarose resin for protein purification were from New England Biolabs. Taq DNA polymerase was from Promega. MALDI-TOF spectra were acquired with a Voyager DE PRO MALDI mass spectrometer (PerSpective Biosystems, Framingham, MA). All FACS screening experiments were conducted on FACSAria sorter (BD Biosciences). Oligonucleotides were supplied by Integrated DNA Technologies and their sequences are listed in Table S1.

Strains and vectors: XL1-Blue cells were purchased from Agilent Technologies (Santa Clara, California, USA). pET15b, pET21b and pET28a plasmids for protein expression were from Novagen. pCTCON2 plasmid and the yeast strain EBY100 were obtained from Professor K. Dane Wittrup of Massachusetts Institute of Technology.

Cloning and expression of DhbE and the ArCP domain of DhbB. DhbE gene was amplified by polymerase chain reaction (PCR) from the genomic DNA of *Bacillus subtilis* by primers Jun287 and Jun292. The amplified PCR fragment was digested by restriction enzymes KpnI and SpeI and cloned into a pET21b plasmid digested with the same set of restriction enzymes. The gene encoding the ArCP domain of DhbB was amplified by primers Jun290 and Jun326. The amplified PCR fragment was digested with NdeI and SpeI and cloned into the pET21 plasmid. The pET expression plasmids were transformed into BL21(DE3)pLysS chemical competent cells (Invitrogen) and plated on the LB-agar plates with appropriate antibiotics.

Protein expression and purification followed the protocol provided by the vendor of the pET expression system (Novagen) and the Ni-NTA agarose resin (Qiagen). Briefly, the pET plasmids of DhbE and DhbB were transformed into BL21(DE3) cells and the transformants were grown at 37 °C in 1 L LB supplemented with 100 μg/mL ampicillin, to an OD_{600} around 0.6-0.8. The culture was induced for protein expression by the addition of IPTG to a final concentration of 1 mM, and allowed to shake for another 16 hours at 15 °C. Cells were pelleted by centrifugation at 5,000 round per minute (rpm) for 10 min. The cell pellet was resuspended in 10 mL lysis buffer (50 mM Tris, pH 8.0, 500 mM NaCl and 5 mM imidazole) and lysed by passing through a French Press (Thermo Scientific). Insoluble cell debris in the cell lysate were precipitated by centrifugation at 12,000 rpm for 20 min. The clear lysate was collected and 1 mL Ni-NTA (QIAgen) agarose was added to capture the expressed proteins with the 6×His-tag. The mixture was gently rocked at 4 °C for 2 hours, and then transferred into a gravity column. The Ni-NTA agarose beads were washed once with 15 mL lysis buffer and twice with 15 mL wash buffer (50 mM Tris, pH 8.0, 500 mM NaCl, 20 mM imidazole) each time. Protein bound to the Ni-NTA agarose beads was eluted with 5 mL elution buffer (50 mM Tris, pH 8.0, 500 mM NaCl, 250 mM imidazole) and dialyzed against storage buffer (25 mM Tris, pH 8.0, 150 mM NaCl and 0.5 mM DTT). The purified protein was analyzed by SDS PAGE gel electrophoresis on a 4-15% Tris polyacrylamide gel (Biorad) to verify its size and purity.

To prepare the holo ArCP domain of DhbB, Sfp phosphopantetheinyl transferase was used to covalently transfer the phosphopantetheinyl group from coenzyme A (CoASH) onto a conserved Ser residue in apo ArCP (Quadri, et al., 1998; Yin, et al., 2006). A typical reaction for holo ArCP formation contained 50 mM HEPES, pH 7.5, 10 mM $MgCl_2$, 5mM DTT, 100-200 μM ArCP domain of DhbB, 150-250 μM CoASH and 5 μM Sfp. The reaction was completed within 30 min of incubation at 30 °C. Successful modification of ArCP by Ppant was confirmed by MALDI mass spectroscopy. Holo-ArCP of DhbB was used as the substrate of DhbE in the PPi release assays.

Cloning of DhbE into pCTCON2 vector and construction of the DhbE library. DhbE gene was PCR-amplified from pET-DhbE with PCR primers Keya2 and Keya3. The amplified DNA fragment was double-digested with SacII and SpeI, and ligated with pCTCON2 plasmid digested with the same restriction enzymes.

Overlapping extension PCR was used to construct a library of DhbE genes with random mutations at codons for residues His234, Asn235, Ala333 and Val337. PCR primers Keya4 and Keya6 had NNK codons substituting the codons for the residues to be randomized in the library. These primers paired with Keya7 and Keya2 respectively to amplify fragments of DhbE gene with randomized codons at residue positions 234, 235, 333 and 337. The 5' end of DhbE gene was amplified with PCR primers Keya3 and Keya5. The three DhbE gene fragments were then assembled by an overlapping extension PCR with primers Keya2 and Keya3. The full length DhbE gene with randomized mutations was digested with SacII and SpeI and ligated into the pCTCON2 vector. Transformation of the plasmid library into SS320 competent cells afforded a library of 2.0×10^7 in size, large enough to cover all the mutants with four randomized residues. Transformed SS320 cells were plated on LB-ampicillin plates (LB agar supplemented with 100 μg/mL ampicillin) and allowed to grow at 37 °C overnight. Colonies appeared on the plate were scraped and the library DNA was extracted from the scraped cells with the Plasmid Maxiprep Kit (QIAgen).

Cloning and expression of the selected DhbE mutants. Genes of DhbE mutants were amplified by PCR from their respective pCTCON2 constructs with primers Keya1 and Keya2. The amplified PCR fragments were double-digested with KpnI and SpeI, and ligated into pET21b vector. pET constructs encoding these mutants were individually transformed into BL21(DE3) cells, and the proteins were expressed in the same way as wtDhbE. To introduce the Trp234His mutation in the selected DhbE mutants, primers Keya8-11 that encode the respective mutations were each paired with Keya2. Overlapping extension PCR was used to assemble the full length DhbE gene with the Trp234His mutation. The PCR products were double-digested with KpnI and SpeI and ligated into the pET plasmid.

Kinetic characterization of DhbE and mutants by ATP-PPi exchange assay. A typical reaction for ATP-PPi exchange contained 75 mM Tris, pH=7.5, 10 mM $MgCl_2$, 2.5 mM DTT, 5 mM ATP, 1 mM ^{32}P sodium pyrophosphate (2-5 Ci/mol, PerkinElmer), and variable amount of enzyme and aryl acid substrates in a total volume of 50 μL. To assay the kinetics of enzymes with DHB or SA as the substrates, 0.2 μM wtDhbE and the mutants were used in the reaction. For reactions with 3-HBA and 2-ABA as the substrates, 2 μM of enzyme was used. After 30 min incubation at room temperature, the reactions were quenched with the addition of 500 μL charcoal suspension (16 g/L activated charcoal, 100 mM sodium pyrophosphate and 3.5% perchloric acid). The mixture was centrifuged at 13,000 rpm for 1 min to pellet the charcoal. The supernatant was removed by aspiration. The charcoal pellet was washed with 1 mL wash buffer (100 mM sodium pyrophosphate and 3.5% perchloric acid) for three times to remove residual ^{32}PPi before the pellet was resuspended in 500 μL H_2O and added to 3.5 mL Ultima Gold scintillation liquid (PerkinElmer). The radioactivity of the charcoal was measured by a scintillation counter (Beckman Coulter LS 6500). Parallel assays with a series of substrate concentrations were performed in triplicate to obtain rates of the ATP-PPi exchange reaction at various substrate concentrations. The kinetic data were fitted to the Michaelis-Menten equation to derive the values of K_m and k_{cat} of the reaction catalyzed by wtDhbE and the mutants.

Kinetic characterization of wtDhbE and the mutants by PPi release assay. Kinetics of wtDhbE and mutants catalyzed substrate activation and transfer to the ArCP domain of DhbB were also measured by a PPi release assay. In this assay the formation of PPi in the reaction was coupled with the phosphorylation of a chromogenic substrate 2-amino-6-mercapto-7-methylpurine ribonucleoside (MESG) in the presence of inorganic phosphatase and purine nucleoside phosphorylase. Reactions to measure PPi release rates were set up following the instructions of the PPi assay kit (Invitrogen). A master mixture contained 50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 100 μM sodium azide, 5 mM ATP, 200 μM MESG, 1 U purine nucleoside phosphorylase, 0.03 U inorganic pyrophosphatase, 125 μM holo-ArCP of DhbB and 0.5 μM wtDhbE or mutants. The reaction mixture was first equilibrated at room temperature for 5 min to "exhaust" the phosphate residue in the system. An aliquot of 100 μL of the master mix was then combined with the aryl acid substrates and mixed thoroughly. Absorbance at 360 nm was measured on a UV-VIS spectrometer over a period of 10 min against a blank solution that contained MESG in the reaction mixture with no enzymes. Initial velocity of the PPi release was determined by measuring the slope of the linear portion of the absorbance curve. The assays were set up in duplicates at varying concentrations of aryl acid substrates. The kinetic data were then fit to the Michaelis-Menten equation to calculate the K_m and k_{cat} values of wtDhbE and mutants with various substrates.

References

Grimes, K.D., and Aldrich, C.C. (2011). A high-throughput screening fluorescence polarization assay for fatty acid adenylating enzymes in Mycobacterium tuberculosis. Analytical biochemistry 417, 264-273.

Heacock, D., Forsyth, C.J., Shiba, K., and Musier-Forsyth, K. (1996). Synthesis and Aminoacyl-tRNA Synthetase Inhibitory Activity of Prolyl Adenylate Analogs. Bioorg. Chem. 24, 273-289.

Quadri, L.E., Weinreb, P.H., Lei, M., Nakano, M.M., Zuber, P., and Walsh, C.T. (1998). Characterization of Sfp, a Bacillus subtilis phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. Biochemistry 37, 1585-1595.

Somu, R.V., Wilson, D.J., Bennett, E.M., Boshoff, H.I., Celia, L., Beck, B.J., Barry, C.E., 3rd, and Aldrich, C.C. (2006). Antitubercular nucleosides that inhibit siderophore biosynthesis: SAR of the glycosyl domain. J Med Chem 49, 7623-7635.

Sundlov, J.A., Shi, C., Wilson, D.J., Aldrich, C.C., and Gulick, A.M. (2012). Structural and Functional Investigation of the Intermolecular Interaction between NRPS Adenylation and Carrier Protein Domains. Chem. Biol. 19, 188-198.

Winkle, M.R., and Ronald, R.C. (1982). Regioselective metalation reactions of some substituted (methoxymethoxy)arenes. Journal of Organic Chemistry 47, 2101-2108.

Yin, J., Lin, A.J., Golan, D.E., and Walsh, C.T. (2006). Site-specific protein labeling by Sfp phosphopantetheinyl transferase. Nature Protocols 1, 280-285.