Electronic Supplementary Material (ESI) for Chemical Science. This journal is © The Royal Society of Chemistry 2018

Electronic Supporting Information for

A DNA-conjugated small molecule catalyst enzyme mimic for site-selective ester hydrolysis

By Moira L. Flanagan, A. Emilia Arguello, Drew E. Colman, Jiyeon Kim, Jesse N. Krejci, Shimu Liu, Yueyu Yao, Yu Zhang, David J. Gorin*

Smith College, Department of Chemistry, Northampton MA 01063

^{*}dgorin@smith.edu

Table of Contents

Materials and Methods	
General Information	S3
Synthesis of DCats	S3
Fluorogenic Assays for Ester Hydrolysis	S6
Quantification of [umbelliferone]	\$6
Control Experiments: DCat structure	S6
Data Analysis	S8
Michaelis-Menten Kinetic Analysis	S8
First-order Rate Law Parameters	S8
Kinetics of imidazole-catalyzed hydrolysis of substrate 1	S10
Uncatalyzed Hydrolysis Rates	S10
Toehold-Switch Experiments:	
Synthesis of esters	
Synthesis of substrate 1	
Synthesis of substrate 2	
Two-step synthesis of substrate 3	
NMR data	S16
Supplemental References	\$20

Materials and Methods:

General Information:

Unless otherwise noted, all reagents were obtained commercially and used without further purification. RP-HPLC purification was performed with Acetonitrile/0.1 M TEAA as the mobile phase at a rate of 1 mL/minute at 40 °C using a reverse-phase Zorbax Eclipse XDB-C18 column (5 μ m, 4.6 x 250 mm). NMR spectra were referenced to chloroform or methanol and obtained with a Bruker 300 MHz NMR or Bruker 500 MHz NMR, as noted. DCat and DNA-construct mass were confirmed by ESI-MS using a Thermo Finnigan LGQ Deca XP Max mass spectrometer, while high-resolution mass spectral data to confirm mass of substrates 1 and 3 was obtained by the University of Massachusetts Mass Spectrometry Center.

Synthesis of DCats:

Cholic acid-binding DCats were synthesized from published aptamers.^{1,2} Oligomers with C₆-NH₂ modifications at the desired base locations (Fig S1b) for each DCat were ordered from TriLink BioTechnologies (Table S1).

For each DCat synthesis, stocks of 20 mM disuccinimidyl glutarate (DSG, 5.2 mg, 16 μ mol in 800 μ L DMSO) and 30 mM histamine (2.8 mg, 24.8 μ mol in 800 μ L DMSO) were prepared. 500 μ L of each stock were combined (new concentrations = 10 mM DSG and 15 mM histamine) and agitated for 10 minutes. A second solution containing DNA (20 nmol, 20 μ L from 1mM stock in PBS buffer) and 40 μ L 0.5 M TEA/HCl pH10 was prepared, and 240 μ L of the DSG/histamine solution was then added to DNA solution. The reaction was vortexed for 30 minutes and incubated at room temperature overnight. Elution through an Illustra NAP-5 column (GE Healthcare) and RP-HPLC provided pure DCat. Concentration was determined with a Nanodrop 8000 spectrometer, and mass was confirmed by ESI-MS (Table S2).

Figure S1: Scheme for synthesis of DCats

Construct	Aptamer Sequence
DCat1	5' GCA GGG T(dC-C6-NH ₂)A ATG GAA TTA ATG ATC AAT TGA CAG ACG CAA GTC TCC TGC 3'
DCat2	5' GCA GGG TCA ATG GAA TTA ATG ATC AAT TGA CAG A(dC-C6-NH ₂)G CAA GTC TCC TGC 3'
DCat3	5' GCA GGG TCA ATG GAA TTA ATG ATC AAT (dT-C6-NH₂)GA CAG ACG CAA GTC TCC TGC 3'
DCat4	5' GCA GGG TCA ATG GAA TTA ATG ATC AAT TGA CAG ACG CAA G(dT-C6-NH ₂)C TCC TGC 3'
DCat5	5' GCA GGG TCA ATG GAA TTA ATG ATC AAT TGA CAG ACG CAA GTC TCC (dT-C6-NH₂)GC 3'
DCat6	5' AGC GCC GAT TGA CCC AAA TCG TTC TGT ATG CAG AAG (dC-C6-NH₂)GC T 3'
DCat7	5' AGC GCC GA(dT-C6-NH₂) TGA CCC AAA TCG TTC TGT ATG CAG AAG CGC T 3'
DCat8	5' AGC GCC GAT TGA CCC AAA (dT-C6-NH₂)CG TTC TGT ATG CAG AAG CGC T 3'
DCat9	5' AGC GCC GAT TGA CCC AAA TCG T(dT-C6-NH₂)C TGT ATG CAG AAG CGC T 3'
DCat10	5' AGC GCC GAT TGA CCC AAA TCG TTC (dT-C6-NH ₂)GT ATG CAG AAG CGC T 3'

Table S1: Sequences of aptamers incorporated into the designated DCat. Amine-modified bases are in bold.

Compound	Ion Charge (Z)	Calculated m/z Observed m/			
	-16	949.18	949.27		
DCat1	-17	893.29	893.40		
	-18	843.60	843.80		
	-16	949.18	949.27		
DCat2	-17	893.29	893.33		
	-18	843.60	843.73		
	-16	948.24	948.10		
DCat3	-17	892.40	892.21		
	-18	842.77	842.65		
	-16	948.24	948.33		
DCat4	-17	892.40	893.13		
	-18	842.77	843.07		
	-13	971.63	971.48		
DCat5	-14	902.16	901.94		
	-15	841.95	841.81		
	-13	971.63	971.48		
DCat6	-14	902.16	901.94		
	-15	841.95	841.74		
	-13	972.79	972.56		
DCat7	-14	903.23	903.02		
	-15	842.95	842.68		
	-13	971.63	972.53		
DCat8	-14	902.16	902.80		

	-15	841.95	842.60
	-13	971.63	972.47
DCat9	-14	902.16	903.07
	-15	841.95	843.07
	-13	971.63	971.73
DCat10	-14	902.16	902.33
	-15	841.95	842.20

 Table S2: MS characterization of DCat constructs

Fluorogenic Assays for ester hydrolysis:

Unless otherwise noted, reactions were performed in 40 μ L volumes with 5 μ M DCat and 10 μ M substrate compound in buffer (20 mM Tris-HCl pH 7.4, 100 μ M NaCl, 5 mM KCl, 2 mM MgCl₂, 1 mM CaCl₂) and 10% (v/v) DMSO. Before the addition of substrate, reaction mixtures were heated for 10 minutes at 95°C and cooled to 37°C at a rate of 0.1°C/s in a MJ Research PTC-2000 Peltier Thermal Cycler. Hydrolysis reactions were run at 37°C. All experiments were repeated at least twice. To avoid confusion and clutter, graphs with many experimental conditions show only one representative experiment for each set of conditions.

Quantification of [umbelliferone]

Umbelliferone fluorescence was monitored (λ_{ex} =320 nm and λ_{em} =460 nm) for 36-120 hours after the addition of substrate to the reaction with a Molecular Devices SpectraMax i3x Multi-Mode Detection Platform. Umbelliferone concentration is linearly proportional to fluorescence, (Fig S2) enabling conversation of fluorescence units to product concentration. A standard ladder of umbelliferone solutions was included for every experiment for standardization.

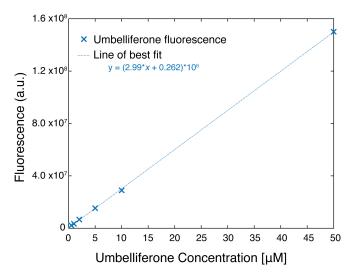


Figure S2: Umbelliferone fluorescence is linear with concentration shown in a representative calibration curve.

Control experiments: DCat structure

To ensure that the entire DCat construct is necessary for catalytic activity, we compared **DCat1**-catalyzed hydrolysis of **1** to control reactions where the DCat is replaced with a known starting material or side product from the DCat synthesis (Fig. S3). Additionally, a hydrolysis reaction was run in the presence of

both amine-modified Ap1 and free imidazole (Fig S4). No rate enhancement was observed, confirming that the covalent link between Ap1 and imidazole is essential. Catalytic activity is only observed with **DCat1**.

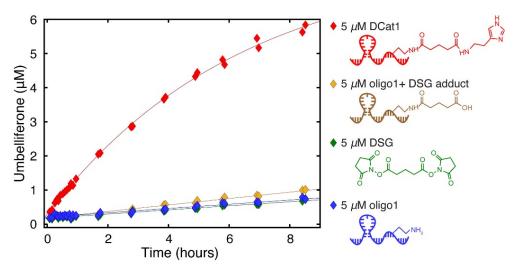


Figure S3: Control experiments for starting materials and side-products that might be present following the synthesis of **DCat1**. Reactions run in standard conditions.

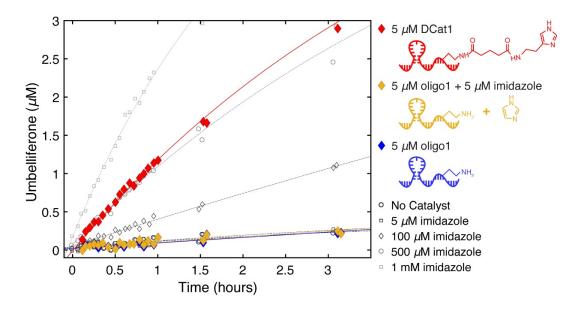


Figure S4: Control experiments of uncorrelated and untethered DNA and imidazole. Reactions run in standard conditions.

Data Analysis:

Michaelis-Menten Kinetic Analysis:

The Michaelis-Menten parameters (Fig 5) were determined by fitting the 5 uM **DCat1** data in Figs 5b and S5b to the following, Michaelis-Menten equation:

$$v_0([S]) = \frac{k_{cat}[E_0][S]}{K_m + [S]}$$
 (1)

Figures 5b and S5b contain data the correlates the initial reaction rate, v_0 , to the substrate concentration. The initial reaction rate for a given substrate concentration was found by extrapolating the slope of the linear region (0-1 hours) of substrate hydrolysis (Fig S5a). The slope is equal to the initial reaction rate, which is the maximum rate for that given set of conditions.

First-order Rate Law Parameters:

The first-order rate constants were determined by finding the initial reaction rate for a given set of reaction conditions, as shown in Fig S5a. The effective first order rate constant was found by the relationship: $k_{obs} = v_0/[substrate]$. In figures 2, 4, S3, S4 and S8, we show first-order rate law best fit curves to the data, but these lines are only the aid the viewer and are not the fits used to determine rate constants. The curves are fits to the following equation, and the parameters for each dataset can be found in table S3:

$$y = A(1 - e^{-kt}) + B \tag{2}$$

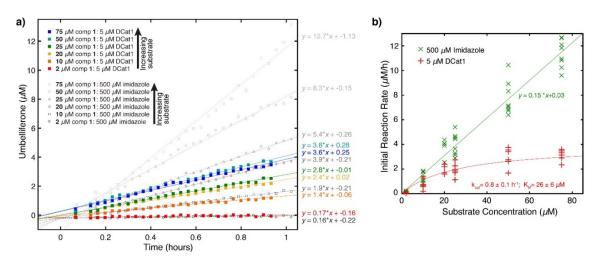


Figure S5: Analysis of DCat1 hydrolysis of substrate 1 to determine Michaelis-Menten parameters. a) Demonstration of how initial reaction rates were determined for one experiment repetition. The early time data (linear region) is fit to a line, allowing variation in the y-intercept due to background fluorescence. The slope is recorded as the initial reaction rate for each substrate level. b) the initial reaction rates found from n=4 (DCat 1) and n=6 (imidazole) experiments for both imidazole and DCat1 catalysts used to characterize the kinetics of each reaction.

Experiment	Para	meter	E	xperiment	Pa	ramete	r
Figure 2c	Α	В	k	Figure 2d	Α	В	k
No Catalyst	1.17	0.08	0.06	5 μM DCat1, n=9	7.16	7.03	0.15
5 μM imidazole	0.86	0.06	0.10	5 μM imidazole, n=11	-73.21	-73.26	0.00
100 μM imidazole	3.81	0.06	0.08	500 μM imidazole, n=11	5.91	5.89	0.20
500 μM imidazole	5.42	0.08	0.22				
1 mM imidazole	6.05	0.09	0.39				
5 μM DCat1	6.46	0.13	0.18				
5 μM DCat2	4.22	0.09	0.09				
5 μM DCat4	2.91	0.08	0.08				
5 μM DCat7	3.08	0.09	0.10				
5 μM DCat10	5.00	0.12	0.15				
Figure 4c	Α	В	k	Figure 4d	Α	В	k
DCat1 + 1	7.07	6.66	0.09	DCat1 + 1	7.07	6.66	0.09
imidazole + 1	9.35	9.14	0.01	imidazole + 1	9.35	9.14	0.01
100x imidazole + 1	9.36	9.01	0.11	100x imidazole + 1	9.36	9.01	0.11
DCat1 + 2	10.67	10.40	0.01	Dcat1+ 3	13.09	11.37	0.01
imidazole + 2	4.46	4.19	0.04	imidazole + 3	4.53	3.06	0.03
100x imidazole + 2	7.67	7.88	0.42	100x imidazole +3	6.65	5.01	0.19
Figure S3	Α	В	k				
5 μM DCat1 (red)	7.48	0.29	0.16				
5 μM oligo1+DSG adduct (gold)	16.67	0.18	0.01				
5 μM DSG (green)	14.08	0.17	0.00				
5 μM oligo1 (blue)	7.83	0.20	0.01				
Figure S8a	Α	В	k	Figure S8b	Α	В	k
DCat 1, n=2	6.25	0.18	0.12	DCat7, n=2	3.57	0.10	0.05
DCat 1 + random DNA n=2	6.07	0.21	0.13	DCat7 + comp1 + rescue1, n=4	3.35	0.10	0.06
DCat 1 + comp1 n=4	5.22	0.11	0.01	DCat7 + comp1, n=4	3.55	0.10	0.05
DCat 1 + comp1 + rescue1, n=4	5.90	-0.16	0.11	5 μM imidazole, n=11	4.37	0.05	0.02
DCat 1 + comp1 + rescue1, n=2	5.92	-0.80	0.10	500 μM imidazole, n=11	5.62	-0.09	0.22

Table S3: Observed first order fit parameters for figures 2, 4, S3, and S8

Kinetics of imidazole-catalyzed hydrolysis of substrate **1**:

To establish that imidazole-catalyzed hydrolysis of substrate ${\bf 1}$ follows second-order kinetics, we checked whether reaction rate increases linearly with both substrate and catalyst (Fig S6), as is indeed the case. If the reaction follows the second order rate law, v=k[catalyst][substrate], the two experiments shown in fig. S6 should appear as pseudo first-order, where the slope of the line of best fit in fig S6a is equal to $m=k[substrate\ {\bf 1}]=0.0022\ \mu Mh^{-1}$ and the slope in fig S6b is equal to $m=k[imidazole]=0.15\ \mu Mh^{-1}$. That is, the experiment in fig S6a gives $k=0.0022\ \mu Mh^{-1}\ /10\ \mu M=2.2\times 10^{-4}h^{-1}$ while fig S6b gives $k=0.15\ \mu Mh^{-1}/500\ \mu M=3.0\times 10^{-4}h^{-1}$; these values are not significantly different. Therefore, we conclude that the imidazole-catalyzed hydrolysis of substrate ${\bf 1}$ follows: $v=3\times 10^{-4}h^{-1}[imidazole][substrate\ {\bf 1}]$.

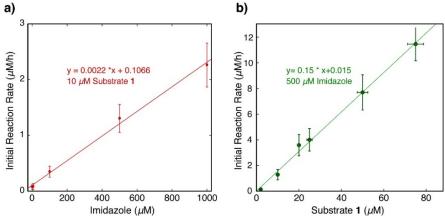


Figure S6: The imidazole-catalyzed hydrolysis of substrate 1 follows second order kinetics. a) The initial rate increases linearly with increasing concentration of imidazole catalyst. b) The initial rate increases linearly with increasing concentration of substrate 1.

Uncatalyzed Hydrolysis Rates:

To account for background (uncatalyzed) hydrolysis of each ester substrate (1, 2, 3) under the reaction conditions, we determined k_{back} , the hydrolysis reaction rate constant for each substrate in the absence of imidazole and/or DCat (Figure S7). These reactions were monitored for 6 h. The slope of the line of best fit is the initial reaction rate and the rate constant is determined by dividing the initial reaction rate by the initial substrate concentration.

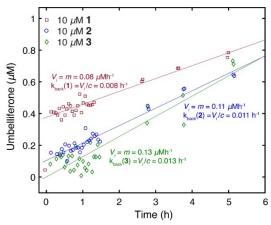


Figure S7: Determination of background hydrolysis rate constant, k_{back}, for each ester substrate.

Toehold-Switch Experiments:

DNA sequences used in Figure 3 are compiled in Table S4. The experiment shown in Figure 3 of the manuscript was also performed with **DCat7** as a control (Figure S8). Notably, **DCat7** was constructed with aptamer sequence **Ap2**, which has a different primary sequence than **Ap1**, which was used in **DCat1**. Therefore, **comp1** is not complementary to the DNA in **DCat7** (Table S1 and S4) and should not inhibit **DCat7** activity. Indeed, we see similar catalytic behavior of **DCat7** regardless of the inclusion of **comp1** or **rescue1**.

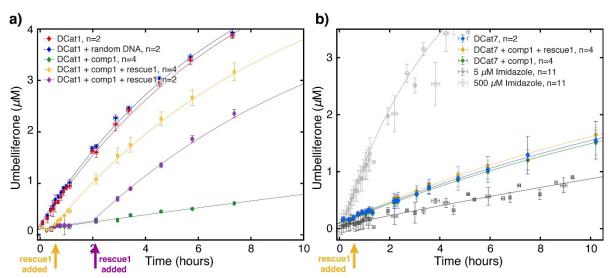


Figure S8: Control of DCat activity using toehold exchange shows catalytic behavior depends on aptamer structure: a) Experimental demonstration of toehold displacement rescue of DCat activity. The data shows product formation from hydrolysis of compound **1** with DCat1 under a series of conditions. DCat activity is suppressed by comp1 (green, yellow, purple diamonds) and then rescued at 35 minutes (yellow diamonds) or 130 minutes (purple diamonds) by the addition of rescue1 DNA, while "random" DNA does not suppress DCat activity (blue diamonds). Error bars represent 1 standard deviation, n as indicated. b) similar experiment with DCat7 in place of DCat1. Oligo rescue1, added at 35 minutes, does not change hydrolysis rate. Error bars represent 1 standard deviation, n as indicated.

oligo	DNA Sequence
comp1	5' GCA GGA GAC TTG CGT CTG TCA ATT GAT CAT TAA TTC CAT TGA CCC TGC GCG CAT GCA T 3'
rescue1	5' A TGC ATG CGC GCA GGG TCA ATG GAA TTA ATG ATC AAT TGA CAG ACG CAA GTC TCC TGC 3'
random	5' AGC GCT TCT GCA TAC AGA ACG ATT TGG GTC AAT CGG CGC T'

Table S4: Sequences of oligonucleotides used in experiments in Fig 3 and Fig S8. Toehold is in bold.

Synthesis of esters:

Under a nitrogen atmosphere, cholic acid (200 mg, 0.49 mmol, 1 equiv.), umbelliferone (0.49mmol, 79 mg, 1 equivalent), DMAP (12 mg, 0.10 mmol, 0.2 equiv.), and DCC (111 mg, 0.42 mmol, 1.1 equiv.) were added to a 10 mL round bottom flask. Anhydrous CH_2Cl_2 (2.5 mL, 39 mmol) was added and the resulting mixture was stirred overnight at room temperature. The precipitate was removed by celite filtration and concentrated. After column chromatography (2-4% MeOH gradient in CH_2Cl_2), **1** was isolated as a white solid (14 mg, 0.025 mmol, 5.2% yield). 1 H-NMR (500 MHz, CDCl₃): δ 7.72 (d, 1 H, J = 9.5 Hz), 7.51 (d, 1 H, J = 8.0 Hz), 7.13 (d, 1 H, J = 2.0 Hz), 7.07 (dd, 1 H, J = 8.5, 2.0 Hz), 6.42 (d, 1 H, J = 9.5 Hz), 4.03 (s, 1 H), 3.88 (s, 1 H), 3.49 (m, 1 H), 2.69 (m, 1 H), 2.56 (m, 1 H), 2.24 (m, 2 H), 2.01-1.01 (multiplet, 26 H), 0.92 (s, 3 H), 0.74 (s, 3 H). 13 C-NMR (125 MHz, CDCl₃): δ 12.6, 17.4, 22.5, 23.2, 26.7, 27.5, 28.4, 30.5, 30.7, 31.3, 34.6, 34.7, 35.1, 35.2, 39.6, 39.7, 41.4, 42.0, 46.5, 47.0, 68.4, 72.0, 72.9, 110.5, 116.0, 116.6, 118.4, 128.5, 142.9, 153.3, 154.7, 160.4, 172.0. HRMS (ESI) m/z: Calculated for $C_{33}H_{45}O_7^+$: 553.3160, found: 553.3161.

Under a nitrogen atmosphere, umbelliferone (195 mg, 1.2 mmol, 1 equiv) was added to 1.5 mL anhydrous diethylether in a 10 mL round bottom flask, the solution is cooled to 0°C and triethylamine (180 uL, 1.3 mmol, 1.1 equivalent) was added. A solution of acetyl chloride (100 μ L, 1.4 mmol, 1.2 equivalents) in 1 mL diethyl ether was added dropwise and the reaction was stirred overnight. After column chromatography (9:1, hexanes/ethyl acetate eluent), **2** was isolated as a white solid. The spectral data were consistent with reported values. H-NMR (300 MHz, CDCl₃): δ 7.72 (d, 1 H, J = 8.7 Hz), 7.52 (d, 1 H, J = 7.5 Hz), 7.15 (d, 1 H, J = 2.1 Hz), 7.08 (dd, 1 H, J = 9.0, 2.1), 6.43 (d, 1 H, J = 9.9), 2.37 (s, 3 H)

Two-step synthesis of substrate 3:

Figure S9: first step in synthesis of substrate 3

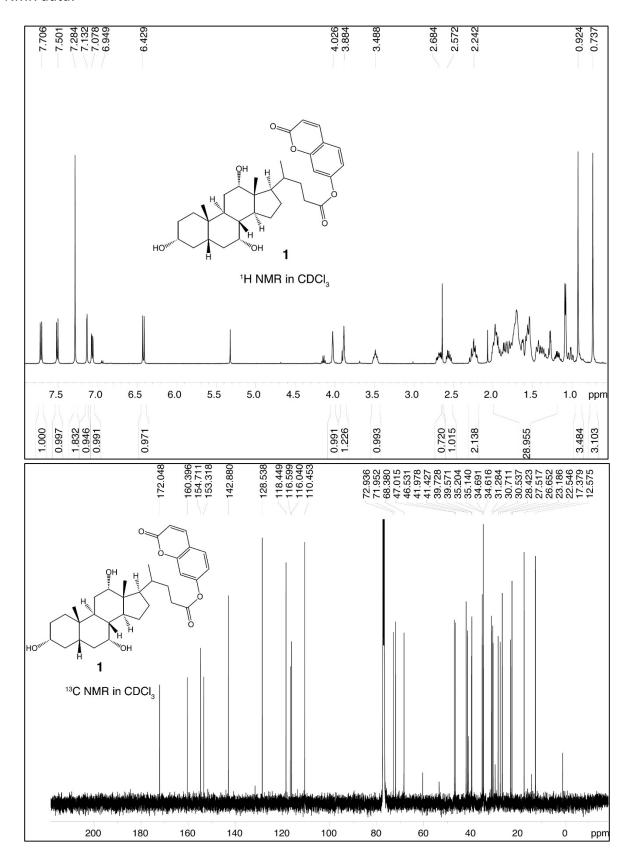
To a 25 mL round bottom flask under N_2 equipped with magnetic stir bar was added CH_2CI_2 (3.6 mL), 5-hexynoic acid (0.200 mL, 1.78 mmol, 1 equiv), EDC-HCl (0.376 g, 1.96 mmol, 1.1 equiv), and umbelliferone (0.376 g, 2.31 mmol, 1.3 equiv). Then DMAP (0.022 g, 0.18 mmol, 0.1 equiv) was added. The reaction mixture was stirred overnight at room temperature and the solvent was then removed *en vacuo*. After column chromatography (4:1 hexanes/ethyl acetate eluent), **4** was isolated as a white solid (235.5 mg, 52% yield), which was carried forward to the next step. 1 H NMR (CDCl₃, 500 MHz): δ 7.71 (d, 1 H, J = 9.5 Hz), 7.51 (d, 1 H, J = 8.5 Hz), 7.15 (d, 1 H, J = 2), 7.08 (dd, 1 H, J = 8.5, 2.0 Hz), 6.42 (d, 1 H, J = 9.5 Hz), 2.79 (t, 2 H, J = 7.5 Hz), 2.39 (td, 2 H, J = 7.0, 3.0 Hz), 2.05 (t, 1 H, J = 2.7 Hz), 2.01 (quin, 2 H, J = 7.2 Hz). 13 C NMR (CDCl₃, 125 MHz) δ 17.8, 23.3, 32.9, 69.6, 82.8, 110.4, 116.1, 116.7, 118.4, 128.5, 142.8, 153.2, 154.7, 160.3, 170.9.

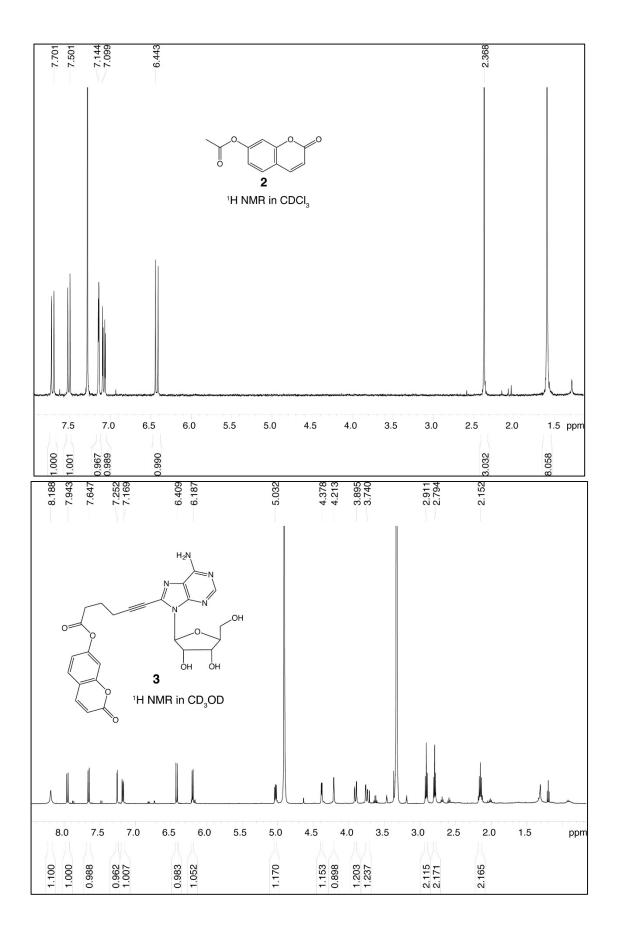
Figure S10: Final step in synthesis of substrate **3**.

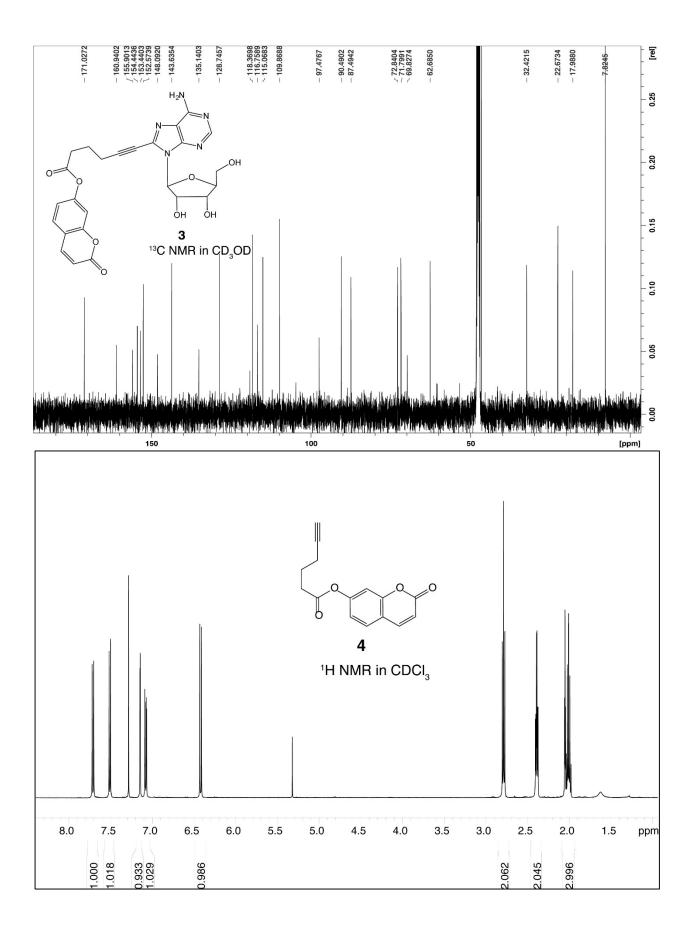
This procedure was adapted from Lambertucci et al.⁵ To a 25 mL round bottom flask under N_2 equipped with magnetic stir bar was added 8-bromoadenosine (0.050 g, 0.15 mmol, 1.0 equiv), DMF (4.8 mL), and Et_3N (0.65 mL, 4.60 mmol, 32 equiv). $PdCl_2(PPh_3)_2$ (0.005 g, 0.008 mmol, 0.05 equiv) and CuI (0.0003 g, 0.0016 mmol, 0.013 equiv) were added, followed by **4** (0.111 g, 0.043 mmol, 3.0 equiv). The reaction mixture was stirred at room temperature for 60 h, and the DMF was then removed by rotary evaporation. The adenosine ester (**3**, 15.6 mg, 20% yield) was obtained as a white solid after two rounds of column

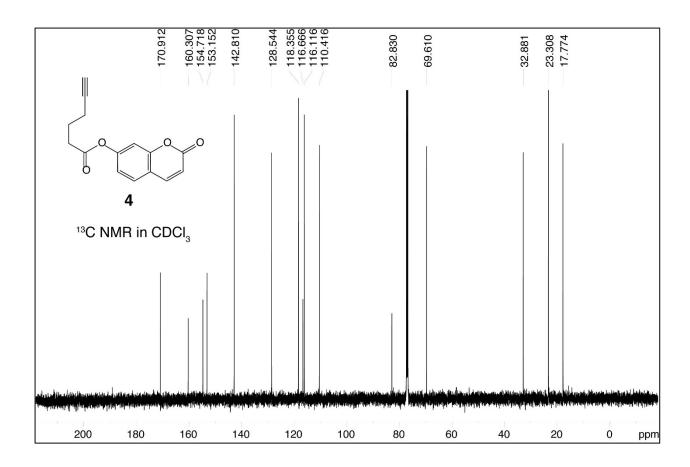
chromatography (19:1 CH₂Cl₂/MeOH). ¹H-NMR (500 MHz, CD₃OD): δ 8.19 (s, 1 H), 7.95 (d, 1 H, J = 9.5 Hz), 7.65 (d, 1 H, J = 8.0), 7.25 (d, 1 H, J = 2 Hz), 7.17 (dd, 1 H, J = 8.5 Hz, 2 Hz), 6.42 (d, 1 H, J = 9.5 Hz), 6.19 (d, 1 H, J = 7.5 Hz), 5.03 (dd, 1 H, J = 7.0 Hz, 5.5 Hz), 4.38 (dd, 1 H, J = 4.9 Hz, 1.5 Hz), 4.21 (d, 1 H, J = 1.5 Hz), 3.90 (dd, 1 H, J = 12.5 Hz, 6.5 Hz), 3.74 (t, 1 H, J = 13.5 Hz), 2.91 (t, 2 H, J = 7.5 Hz), 2.79 (t, 2 H, J = 7.0 Hz), 2.16 (quin, 2 H, J = 7.0 Hz). ¹³C-NMR (125 MHz, CD₃OD): δ 7.83, 17.99, 22.67, 32.42, 62.69, 69.83, 71.80, 72.84, 87.49, 90.49, 97.48, 109.87, 115.07, 116.76, 118.37, 128.75, 135.14, 143.64, 148.09, 152.57, 153.44, 154.44, 155.90, 160.94, 171.03. HRMS (ESI) m/z: Calculated for C₂₅H₂₄N₅O₈⁺: 522.1619, found: 522.1628.

NMR data:









References:

- 1 T. Kato, K. Yano, K. Ikebukuro and I. Karube, *Nucleic Acids Res.*, 2000, **28**, 1963–1968.
- 2 T. Kato, T. Takemura, K. Yano, K. Ikebukuro and I. Karube, *Biochim. Biophys. Acta BBA Gene Struct. Expr.*, 2000, **1493**, 12–18.
- 3 DCats are heated at 95 °C to melt all hybridization and tertiary structure, followed by a slow cool to 37 °C to encourage proper DNA folding and tertiary structure. For example: B. Högberg, T. Liedl and W. M. Shih, *J. Am. Chem. Soc.*, 2009, **131**, 9154–9155.
- 4 W. He, B.-L. Zhang, S.-Y. Zhou, X.-L. Sun and S.-Y. Zhang, *Synth. Commun.*, 2007, **37**, 361–367.
- 5 C. Lambertucci, S. Costanzi, S. Vittori, R. Volpini and G. Cristalli, *Nucleosides Nucleotides Nucleic Acids*, 2001, **20**, 1153–1157.