SUPPORTING INFORMATION

Site-specific fluorescence labeling of RNA using bio-orthogonal reaction of trans-cyclooctene and tetrazine.

P.N. Asare-Okai, E. Agustin, D. Fabris and M. Royzen *

University at Albany, SUNY, Department of Chemistry
1400 Washington Avenue, Albany, NY 12222. Email: mroyzen@albany.edu

Table of contents:

| Materials and methods | Page: S2 |
| Synthetic procedures | S3-S8 |
| NMR and MS spectra | S9-S15 |

**Fig. S1** Denaturing gel analysis of the *in vitro* transcription experiments used to determine incorporation efficiency of 1

**Fig. S2** Tandem MS analysis of RNase T1 fragment ACAAUG

**Fig. S3** Tandem MS analysis of RNase T1 fragment ACAG

**Table S1** Summary of calculated and observed CID fragments for ACAAUG>p and ACAG>p.

**Fig. S4** Fluorescence spectra of 1 treated with 2

**Fig. S5** A 20 % denaturing gel analysis of *in vitro* transcription products from the 435 nt-long template

**Fig. S6** The Sequence of the sense strand DNA template that was inserted into a pUC57-Kan vector and used for the run off transcription

Electronic Supplementary Material (ESI) for ChemComm. This journal is © The Royal Society of Chemistry 2014
Materials and Methods

All chemicals were received from commercial sources and used without further purification. AG-1-X8 resin was purchased from Bio-Rad Laboratories, Inc. and was converted to bicarbonate form by elution with 1M NaHCO3. AG50W-X8 resin was also purchased from Bio-Rad Laboratories, Inc. and was used without any modifications. Chromatographic purifications were conducted using SiliaSphere™ spherical silica gel 5μm, 60 Å silica gel (Silicycle). Thin layer chromatography (TLC) was performed on SiliaPlate™ silica gel TLC plates (250 μm thickness) purchased from Silicycle. Preparative TLC was performed on SiliaPlate™ silica gel TLC plates (1000 μm thickness). HPLC purification was performed using Phenomenex Luna 5μ C18(2) semi-preparative column (250 x 10 mm) using a gradient of 10-40% CH3CN in 0.1 M TEAB, pH 7.5. 1H and 13C NMR spectroscopy was performed on a Bruker NMR at 400 (1H), 100 (13C) MHz and 162 (31P) MHz. All 13C NMR spectra were proton decoupled.

The in vitro transcribed RNA was extensively desalted through 3-5 cycles of buffer exchange using mass spectrometry grade ammonium acetate (Sigma-Aldrich, St. Louis MO) and 3 KDa centicon (EMD Millipore, Billerica MA). Forty picomoles of the desalted RNA was digested using 1 unit of RNAse T1 in 150 mM NH4Ac for 2 h. The samples were analyzed on a Thermo Fisher Scientific (West Palm Beach, CA) LTQ Orbitrap Velos Mass spectrometer in the negative ion mode, using quartz capillary emitters. To facilitate spray optimization, 10 % isopropyl alcohol was added to each sample prior to MS analysis. The percentage conversion to the cycloaddition product was calculated based on the respective peak intensities of the various species normalized to their charge state using the equation:

$$\frac{\Sigma (I_{ACAAUGp}/z_{ACAAUGp})}{\Sigma (I_{ACAAUGp}/z_{ACAAUGp}) + \Sigma (I_{ACAGp}/z_{ACAGp})} \times 100$$

In which ACAAUGp and ACAGp are the unreacted RNA fragments and ACAAUGp and ACAGp are the corresponding cycloaddition products. Ix is the peak intensity of each signal and zx is the respective charge state. Only the 12C monoisotopic peaks were used in these calculations.

Fluorescence experiment were performed in 1X PBS buffer pH 7.4 (137 mM NaCl, 2.7 mM KCl, 10 mM Na2HPO4 and 1.8 mM KH2PO4). Appropriate amounts of 1 and 2 were mixed and incubated for 1hr at rt. Fluorescence spectra was acquired using Fluorolog® fluorimeter (Horiba Jobin Yvon), model FL-1039. The Fluorescent gels were scanned using a Typhoon Trio scanner (GE Healthcare).

IVT was performed using the mirVana™ miRNA Probe Construction Kit, using manufacturer described protocols:
http://www.lifetechnologies.com/order/catalog/product/AM1550 The samples were processed with 2 units of DNase1 to degrade the DNA template, denatured in 7M urea in the presence of 25 mM EDTA and resolved on a 20% denaturing polyacrylamide gel. The Fluorescent gels were scanned using a Typhoon Trio scanner (GE Healthcare). After that, the gel was stained with SYPBPR Green and imaged on a Molecular Imager® Gel Doc™ XR System (Biorad).
5-[3-(trifluoroacetamido)propynyl]cytidine (3)

5-Iodocytidine (0.443 g, 1.20 mmol), N-propargyl trifluoroacetamide (0.544 g, 3.60 mmol), Cul (45.0 mg, 0.240 mmol), and Pd(PPh₃)₄ (0.138 g, 0.120 mmol) were placed in a 50 mL round bottom flask and dissolved in anhydrous and thoroughly degassed DMF (6 mL). The round bottom flask was placed under vacuum and refilled with N₂. Triethylamine (0.334 mL, 2.40 mmol) was added to the reaction mixture which was stirred at room temperature under nitrogen for 18 h. The reaction mixture was diluted with a 1:1 solution of MeOH:CH₂Cl₂ (18 mL) and AG-1-X8 resin (bicarbonate form, 3 g) was added. After stirring for 30 min, the crude product was filtered and the resin was once rinsed with a 1:1 solution of MeOH:CH₂Cl₂ (18 mL). The solvents were removed under high vacuum and the title product was obtained as a light yellow-brown solid by flash chromatography (5% MeOH in CH₂Cl₂). Yield = 0.376 g (85%).

¹H NMR (CD₃OD, 400 MHz) δ 8.42 (s, 1H), 5.85 (as, 1H), 4.31 (s, 1H), 4.15-4.14 (m, 2H), 4.05-4.03 (m, 1H), 3.92 (dd, J₁ = 12.4 Hz, J₂ = 2.0 Hz, 1H), 3.76 (dd, J₁ = 12.3 Hz, J₂ = 2.2 Hz, 1H). ¹³C NMR (CD₃OD, 100 MHz) δ 166.54, 159.11, 158.74, 157.03, 146.71, 121.75, 118.90, 116.05, 113.20, 112.10, 92.56, 91.03, 85.95, 76.52, 75.68, 70.51, 61.70, 31.15

The spectra is consistent with the one previously reported by:

Compound 3 (0.830 g, 2.11 mmol) was dissolved in anhydrous MeOH (20 mL). Palladium hydroxide on carbon matrix (20% wt.) (0.166 g) and triethylsilane (2.46 g, 21.1 mmol) were added to the reaction mixture. After stirring for 24 hours at room temperature, the reaction mixture was filtered through celite and the filtrate was concentrated under reduced pressure. The title product was obtained as a yellow solid by preparative TLC, using 20% MeOH in CH\textsubscript{2}Cl\textsubscript{2} as a mobile phase. Yield = 0.472 g (56%).

\[ \text{H NMR (CD\textsubscript{3}OD, 400 MHz)} \delta 7.98 (s, 1H), 5.86 (app s, 1H), 4.15 (app s, 2H), 4.02 (app s, 1H), 3.91 (d, \textit{J} = 12.5 Hz, 1H), 3.77 (d, \textit{J} = 12.3 Hz, 1H), 3.35-3.31 (m, 2H), 2.37 (t, \textit{J} = 9.3 Hz, 2H), 1.79 (quin, \textit{J} = 7.3 Hz, 2H). \]

\[ \text{\textsuperscript{13}C NMR (CD\textsubscript{3}OD, 100 MHz)} \delta 165.12, 157.83, 157.46, 156.72, 139.34, 117.50, 114.66, 105.68, 90.74, 84.31, 74.74, 69.11, 60.36, 38.58, 26.62, 23.90 \]

The spectra is consistent with the one previously reported by:

5-(3-aminopropyl)cytidine (4)

Compound 3b (0.482 g, 1.22 mmol) was dissolved in DI H$_2$O (8.5 mL). After complete dissolution, aqueous ammonium hydroxide (NH$_4$OH) (8.5 mL) was added to the reaction mixture. The reaction was stirred for 3 hours at room temperature and then concentrated down to give the crude product as a yellow-orange residue. The crude product was redissolved in DI H$_2$O (10 mL) and AG50W-X8 resin (2.5 g) was added to the solution. The mixture was stirred for 15 min and filtered over a bed of AG50W-X8 resin (2.5 g). The resin was washed with DI H$_2$O and the product was then eluted off of the resin with DI H$_2$O/conc. NH$_4$OH, 4:1 (100 mL). The title product was obtained as a light tan solid after removal of the solvents under reduced pressure. Yield = 0.473 g (98%).

$^1$H NMR (D$_2$O, 400 MHz) δ 7.64 (s, 1H), 5.76 (d, $J = 3.3$ Hz, 1H), 4.15 (t, $J = 5.0$ Hz, 1H), 4.09 (t, $J = 5.9$ Hz, 1H), 4.01-3.98 (m, 1H), 3.83 (dd, $J_1 = 12.6$ Hz, $J_2 = 2.5$ Hz, 1H), 3.69 (dd, $J_1 = 12.5$ Hz, $J_2 = 3.5$ Hz, 1H), 2.55 (t, $J = 7.3$ Hz, 2H), 2.23 (t, $J = 7.7$ Hz, 2H), 1.55 (quin, $J = 7.2$ Hz, 2H). $^{13}$C NMR (D$_2$O, 100 MHz) δ 165.15, 157.08, 138.65, 107.21, 90.18, 83.54, 74.05, 68.84, 60.11, 39.28, 27.54, 23.72.

HRMS (CI) m/z: calcd. for C$_{12}$H$_{21}$N$_4$O$_5$ [M+1] 301.1512; found 301.1489.
2-((E)-cyclooct-4-enoxo)acetic acid (38 mg, 0.303 mmol), N-hydroxysuccinimide (35 mg, 0.303 mmol), and N,N'-Dicyclohexylcarbodiimide (62 mg, 0.303 mmol) were dissolved in anhydrous CH₂Cl₂ (2 mL) and stirred at rt under N₂ atmosphere for 12 h. The precipitate was filtered and the eluent was concentrated under reduced pressure. The concentrate was combined with 4 (90 g, 0.303 mmol) and dissolved in anhydrous DMF (5 mL). Triethyl amine (123 μL, 0.909 mmol) was added and the reaction mixture was stirred at 30 °C for 18 hours. The solvent was removed under high vacuum and the title product was obtained as a white powder by preparative column chromatography using 10%MeOH in EtOAc as a mobile phase. Yield = 40 mg (39%).

1H NMR (CD₃OD, 400 MHz) δ 7.99 (s, 1H), 5.89 (d, J = 1.9 Hz, 1H), 5.60-5.45 (m, 2H), 4.21-4.17 (m, 2H), 4.04-4.03 (m, 1H), 3.96-3.87 (m, 3H), 3.77 (dd, J₁ = 11.2 Hz, J₂ = 2.6 Hz, 1H), 3.64 (dd, J₁ = 9.75 Hz, J₂ = 4.25 Hz, 1H), 3.35-3.31 (m, 2H), 2.39-2.29 (m, 4H), 2.21-2.17 (m, 2H), 2.02-1.99 (m, 1H), 1.86-1.77 (m, 5H), 1.59-1.51 (m, 1H), 1.28-1.19 (m, 1H).

13C NMR (CD₃OD, 100 MHz) δ 173.70, 171.61, 164.72, 139.25, 135.08, 131.03, 105.91, 90.58, 84.34, 75.95, 74.71, 69.07, 67.72, 60.33, 39.78, 37.78, 33.85, 31.96, 31.95, 29.23, 27.46, 23.94.

HRMS (CI) m/z: calcd. for C₂₂H₃₅N₄O₇ [M+1] 467.2506; found 467.2494.
5-[3-((E)-cyclooct-4-enoxyacetamido)propyl]cytidine-5′-triphosphate (1)

Dissolved compound 6 (83 mg, 0.18 mmol) and proton sponge (58 mg, 0.27 mmol) in trimethylphosphate (0.7 mL) and place in the ice bath. Slowly added POCl₃ (33.6 μL, 0.36 mmol) and stirred at 0 °C for 2 h. A solution of tributylamine (175 μL, 0.72 mmol) and tributylammonium pyrophosphate (642 mg, 1.17 mmol) in DMF (1.7 mL) was slowly added and the reaction mixture was stirred at 0 °C for 30 min. The reaction was quenched by addition of 0.5 M aqueous TEAB pH 7.5 (2.4 mL). The mixture was diluted with H₂O (5 mL) and subjected to anion exchange chromatography (DEAE Sephadex (GE Healthcare) with an elution gradient of 0->0.3->0.5->1 M TEAB. Each fraction was analyzed by ESI-MS. The title product was mainly found in the final fraction. Additional HPLC purification was carried out to obtain the product as a white foam. HPLC was performed using Phenomenex Luna 5u C18(2) semi-preparative column (250 x 10 mm) using a gradient of 10-40% CH₃CN in 0.1 M TEAB, pH 7.5.

Yield = 11 mg (5.6%).

¹H NMR (D₂O, 400 MHz) δ 7.70 (s, 1H), 5.89 (d, J = 4.3 Hz, 1H), 5.58-5.45 (m, 2H), 4.33 (t, J = 8.8 Hz, 1H), 4.26 (t, J = 9.8 Hz, 1H), 4.17 (d, J = 4.1 Hz, 3H), 3.85 (s, 2H), 3.60-3.54 (m, 1H), 3.29-3.21 (m, 2H), 3.10 (q, J = 15 Hz, 14H), 3.00-2.90 (m, 1H), 2.44-2.34 (m, 2H), 2.26-2.05 (m, 4H), 1.94 (d, J = 10.8 Hz, 1H), 1.84-1.66 (m, 4H), 1.62-1.47 (m, 2H), 1.18 (t, J = 7.7 Hz, 20H). ³¹P NMR (D₂O, 162 MHz) δ -10.67 (d, J = 19.4 Hz, 1P), -11.71 (d, J = 18.0 Hz, 1P), -23.20 (t, J = 19.7 Hz, 1P).

HRMS (Cl) m/z: calcd. for C₂₈H₅₅N₅O₁₆P₃ [M+Et₃N+1] 808.2622; found 808.2893.
5-(4-(6-methyl-1,2,4,5-tetrazin-3-yl)benzylcarbamoyl)-2-(3-hydroxy-6-oxo-6H-xanthen-9-yl)benzoic acid (2)

Dissolved compound 8 (108 mg, 0.332 mmol) and fluorescein-NHS ester (157 mg, 0.332 mmol) in DMF (10 mL). Added triethylamine (90 µL, 0.664 mmol) and stirred at rt overnight. The solvent was removed under reduced pressure and the title product was obtained as an orange powder by preparative TLC using a 9:1 mixture of CH$_2$Cl$_2$ and MeOH as mobile phase. Yield = 165 mg (89%).

$^1$H NMR (DMSO, 400 MHz) $\delta$ 9.50 (t, $J = 2.6$ Hz, 1H), 8.51 (s, 1H), 8.42 (d, $J = 8.2$ Hz, 2H), 8.23 (d, $J = 8.2$ Hz, 1H), 7.58 (d, $J = 8.2$ Hz, 2H), 7.34 (d, $J = 7.9$ Hz, 1H), 6.59 (d, $J = 2.5$ Hz, 2H), 6.56 (d, $J = 8.1$ Hz, 2H), 6.46 (dd, $J_1 = 8.1$ Hz, $J_2 = 2.5$ Hz, 2H), 4.61 (d, $J = 5.4$ Hz, 2H), 2.92 (s, 3H).

$^{13}$C NMR (DMSO, 100 MHz) $\delta$ 173.11, 168.23, 167.12, 165.00, 163.22, 161.71, 152.45, 144.16, 135.73, 133.98, 130.52, 129.33, 128.24, 127.53, 125.00, 124.11, 113.78, 109.37, 102.39, 42.91, 25.23.

HRMS (CI) $m/z$: calcd. for C$_{31}$H$_{21}$N$_5$O$_6$ [M+1] 560.1492; found 560.1553.
Fig. S1 Denaturing gel analysis of the *in vitro* transcription experiments used to determine incorporation efficiency of 1. Lane 1 shows the 33 nt. long RNA sequence prepared by solid phase synthesis; lanes 2, 3 and 4 show *in vitro* transcripts obtained using variable amounts of CTP and 1. Lane 2: 100 % CTP; lane 3: 50 % CTP, 50 % 1; lane 3 100 % 1. Based on the relative pixel intensity, the transcription efficiency of the CTP analogue was calculated to be 60 ± 5 % (lane 2 was used as a reference).
Fig. S2  Tandem MS analysis of RNase T1 fragment ACAAUG.
Fig. S3  Tandem MS analysis of RNase T1 fragment ACAG
<table>
<thead>
<tr>
<th></th>
<th>ACAAUG&gt;p</th>
<th>ACAG&gt;p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fragment</td>
<td>Calculated mass (Da)</td>
<td>Observed mass (Da)</td>
</tr>
<tr>
<td>c₁</td>
<td>331.068</td>
<td>-</td>
</tr>
<tr>
<td>c₂</td>
<td>859.267</td>
<td>859.253</td>
</tr>
<tr>
<td>c₃</td>
<td>1188.32</td>
<td>1188.31</td>
</tr>
<tr>
<td>c₄</td>
<td>1517.37</td>
<td>1517.36</td>
</tr>
<tr>
<td>c₅</td>
<td>1823.38</td>
<td>1823.39</td>
</tr>
<tr>
<td>y₁</td>
<td>345.047</td>
<td>345.041</td>
</tr>
<tr>
<td>y₂</td>
<td>651.073</td>
<td>651.067</td>
</tr>
<tr>
<td>y₃</td>
<td>980.125</td>
<td>980.121</td>
</tr>
<tr>
<td>y₄</td>
<td>1309.18</td>
<td>1309.18</td>
</tr>
<tr>
<td>y₅</td>
<td>1837.38</td>
<td>1837.37</td>
</tr>
</tbody>
</table>

**Table S1** Summary of calculated and observed CID fragments for ACAAUG>p and ACAG>p.

**Fig. S4** Fluorescence spectra showing: (a) 1 µM 1, (b) 0.1 µM 2 and (c) 0.1 µM 2 treated with 0.1µM 1. Experiments were performed in 1X PBS buffer pH 7.4 (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄ and 1.8 mM KH₂PO₄)
**Fig. S5** A 20% denaturing gel analysis of *in vitro* transcription products from the 435 nt template. Panel A, shows the fluorescence image of the unstained gel and panel B is the image of the same gel after SYBR Green staining.

**Fig. S6** The sequence of the sense strand DNA template that was inserted into a pUC57-Kan vector and used for the run off transcription is shown.

5'-TAATACGACTCACTATAGGTCTCTTGTTAGACCATCTAGCTGAGCTGGGCTCTCGTACAC
T7 promoter
TAGGAAAGCCACTGTTAAAGCGCTCAAACGCTTGGCTAGTCTCAAGTGATGTTGTGCGCG
TCTTGTGTGACTCCTGAACTAGAGATCCCTCCAGCCCTTTATTAGTCTGTTGGAAAAATCTCTA
GCAGTGGCGCCCAAGAGGACTTGAAGAGCTAAGGCTAAGGAGAGGAGCGGGCGCAGACTCTGAGGTGAGTACGC
AGGACTCCGCTTGGCTGAGAGACCTGCAAGAGACAGAGAAGAGGAGGGCGGCGGCGCAGACTCCGAGATGACGC
CAAAAATTTGACTAGCGGAGAGGCTAGAGGAGATGGGTGAGAGGCTGAGGCTCGGTATTAAG
TCGTCGTCGTCCTAGAGCGACGACCTGAGCTAGGCCGAAAGGCCGAACCTTAATACCG-3'
Ribozyme

S20