Electronic Supplementary Material (ESI) for ChemComm.

Cycloaddition enabled the detection of 5-vinyluridine in RNA

Mrityunjay Gupta,[1] Jingtian Wang,[2] Chely M. Garfio,[2] Abigail Vandewalle,[3] Robert C. Spitale[1,2,3,*]

1. Department of Chemistry, University of California, Irvine, Irvine, CA 92697

2. Department of Molecular Biology and Biochemistry, University of California, Irvine, Irvine CA 92697

3. Department of Pharmaceutical Sciences, University of California, Irvine, Irvine, CA 92697

TABLE OF CONTENTS

I. General Information

All reagents were purchased from commercial suppliers, were of analytical grade, and were used without further purification unless otherwise noted. NMR spectra were acquired with Bruker NEO-400 spectrometer (UC Irvine). All ^{1}H NMR spectra were acquired at 298 K and 400 MHz. Chemical shifts are reported in ppm relative to residual non-deuterated NMR solvent unless noted otherwise. All NMR spectra were analyzed using MestreNova software.

II. Experimental Methods

Maleimide substrate scope study

5-VUrd (1-2 mg) was dissolved in D_2O or DMSO-d_6, and 5 equivalents of maleimide substrate were added to the solution. The solution was mixed thoroughly and the mixture was transferred to a fresh NMR tube. The tube was then incubated at ambient room temperature for 3 hours after which ^{1}H NMR spectrum was acquired. Peaks corresponding to the internal proton of terminal alkene in 5-VUrd (6.3 – 6.5 ppm, dd) and 1' proton of the cycloadduct product (5.5 – 5.7 ppm, d) were identified. Both peaks were integrated, and the integration of the 5-VUrd peak was set to 1. Reaction conversion of 5-VUrd with each maleimide was calculated as follows –

$$Conversion\ (\%) = \frac{Integration\ of\ 1'cycloadduct\ proton}{Integration\ of\ 1'cycloadduct\ proton + Integration\ of\ 5{-}VUrd\ alkene\ proton}\ x\ 100$$

RNA Transcription

Double-stranded template DNA was obtained from Integrated DNA Technologies (IDT) and resuspended in nuclease-free water as 100 mM stock. *In-vitro* transcription was performed using T7 RNA polymerase (#M0251S, New England Biolabs) in 20 µL reactions as follows – 2 µL of 10X Reaction buffer, 1 µL of 10 mM CTP, ATP, and GTP each, 1 µL of 10mM UTP or 5-VUrd triphosphate (#CLK-069, Jena Biosciences), 1 µL 100 mM template DNA, 1 µL RNasin ribonuclease inhibitor (#N2611, Promega), 2 µL T7 RNA polymerase and nuclease-free water up to 20 µL. The mixture was incubated at 37°C for 8 hours. The reaction was adjusted to 50 µL volume with nuclease-free water, and RNA was purified by ethanol precipitation by adding 5 µL of 3M NaOAc, 150 µL of pure ethanol, and 1 µL of glycoblue (#AM9515, ThermoFisher Scientific). The solution was incubated at -20°C overnight and precipitated by centrifugation at 15,000 rpm for 30 minutes. The supernatant was discarded, and cold 70% ethanol was added to the pellet. The solution was centrifuged for 10 minutes at 15,000 rpm, after which the supernatant was discarded and the pellet was air-dried for 10 minutes. The pellet was then resuspended in nuclease-free water and the concentration was measured with Nanodrop. RNA was analyzed by running ~ 100 ng samples on a denaturing polyacrylamide gel [10% w/v acrylamide/bisacrylamide 19:1 solution, 7 M urea, and 0.5x TBE Buffer (50 mM tris(hydroxymethyl)aminomethane hydrochloride, 45 mM boric acid, and 0.5 mM EDTA, pH 8.4)] along with an ssRNA ladder (#N0364S, New England Biolabs). The gel was

stained with 1X v/v SYBR Gold (#S11494, ThermoFisher Scientific) for 30 minutes and imaged with ChemiDoc MP imaging system (Bio-Rad).

RNA biotinylation and dot-blot assay

1 µg RNA was reacted with 5 µL of 2 mM maleimide-biotin (#B1267, Sigma) or DMSO (negative control) in 50 µL reactions and incubated at 37°C for 3 hrs. The biotinylated RNA was then purified by ethanol precipitation as previously described and resuspended in 2 µL nuclease-free water. To detect biotinylated RNA with dot blot analysis, biotinylated RNA was fixed to a Hybond-N+ (GE Healthcare) membrane with 254 nm light using a UV stratalinker (Stratagene). The membrane was first equilibrated in 2x SSC buffer before spotting RNA and fixing with UV light. After fixing RNA, membranes were blocked in blocking buffer (0.12 M NaCl, 0.016 M Na_2HPO_4, 0.008 M NaH_2PO_4, 0.17 M SDS) for 30 minutes and followed by incubation with high sensitivity streptavidin-HRP (#PI21130, Fisher Scientific) at 1:5000 dilution in blocking buffer for 5 minutes. The membrane was washed twice in a Wash A buffer (1:10 dilution of blocking buffer in D.I. water) for 30 minutes and twice in Wash B buffer (0.1 M Tris-base, 0.1 M NaCl, 0.02M $MgCl_2$, pH 9.5) for 5-10 minutes. Membranes were then incubated for 1-5 minutes in ECL Chemiluminescent Substrate (Fisher Scientific, Cat#: PI32106) and imaged on a ChemiDoc MP imaging system (Bio-Rad) to detect chemiluminescence on the membrane. After imaging is completed, methylene blue (0.04% methylene blue and 0.3 M sodium acetate) was used to stain onto membranes overnight, rinsed in D.I water, followed by colorimetric imaging to detect consistent loading of RNA onto the membranes.

^{32}P Primer labeling

500 nmol of FL forward primer DNA was phosphorylated with ATP[γ-^{32}P] using T4 PNK enzyme (#M0121S, New England Biolabs). The reaction was incubated for 30 min at 37°C, followed by a 20 min incubation at 60°C to deactivate the kinase. Gel-loading dye was added to the reaction and electrophoresed on a 15% denaturing PAGE gel as described above. The band of interest was visualized by using a phosphor screen and phosphorimager (GE healthcare, Typhoon). The band was excised and eluted overnight in 400 mM KCl. The solution was EtOH precipitated and resuspended in water to obtain a 10,000 c.p.m/µL stock.

Primer extension assay with various maleimide-conjugated RNA

All maleimide substrates were dissolved as stock solutions of 2 mM concentration in pure DMSO. ~500 RNA was reacted with 5 µL of 2 mM respective maleimide and nuclease-free water was added up to 50 µL and incubated at 37°C for 3 hrs. The conjugated RNA was then purified by ethanol precipitation as previously described and resuspended in 10 µL nuclease-free water. 2 µL of ^{32}P labeled FL forward primer stock was annealed to the reacted RNA samples by incubating at 65 °C for 5 min, 35 °C for 5 min, and then incubated on ice for 1 min. 4 µL 5X First-strand buffer, 1 µL DTT, and 1 µL 10 mM dNTPs were added to the mixture and incubated at 52 °C for 1 min. Sequencing samples for G sequencing and U sequencing were prepared similarly with unmodified RNA and the addition of 1 uL of 100mM dideoxycytidine triphosphate (ddCTP) or 100 mM dideoxyadenosine triphosphate (ddATP) respectively to the RT reaction. SuperScript III (#18080093, ThermoFisher Scientific) (10 unit/µL; final concentration) was added and extension was performed at 55°C for 15 min. To hydrolyze the RNA, 1 µL of 4 M sodium hydroxide was added to the reaction and incubated at 95°C for 5 min. The reaction was

then placed on ice for 5 min. The resulting cDNA was EtOH precipitated and resuspended in loading dye. Samples were run on 12% denaturing PAGE gel for 1 hr at 80V and imaged.

Full-length cDNA qPCR

~500 ng of T1 RNA was reacted with various maleimide substrates as described above and resuspended in 10 µL nuclease-free water. RT primer (FL forward) was annealed to RNA by adding 2 µL of 10 mM primer and 1 µL of 10 mM dNTPs and incubating at 65 °C for 5 min, 35 °C for 5 min, and then incubated on ice for 1 min. 4 µL 5X First-strand buffer, 1 µL DTT, 1 µL RNasin, and 1 µL of SuperScript III were added to the mixture and incubated at 55°C for 1 hr. RNA was then degraded by adding 5 U RNase H (#M0297, New England Biolabs) and incubating the mixture at 37°C for 30 minutes. 10 µL qPCR reactions were performed in a 96 well plate (#HSP9601, Biorad) with 1 µL of RT reaction, 0.5 µL of 10 mM FL forward and FL reverse primers each and 5 µL of 2X Luna Universal qPCR master mix (#M3003S, New England Biolabs) with single annealing step at 72°C for 40 cycles. The mean Cq value for each sample was calculated from the triplicates, and full-length cDNA yield was normalized to the DMSO negative control as follows –

$$normalized\ full-length\ cDNA\ yield = 2^{-(mean\ Cq\ of\ sample - mean\ Cq\ of\ DMSO\ control)}$$

Primer extension assay with different RT conditions

RNA conjugation was performed with N-phenyl-maleimide as described previously. RT reactions were performed with ^{32}P labeled FL forward primer as described previously with the following modifications – (i) for TGIRT-III and HIV RT, 1 µL of the respective enzyme was added instead of SuperScript III, (ii) for SuperScript II in Mn^{+2} buffer, 4 uL of 5X

modified first-strand buffer (250 mM tris(hydroxymethyl)aminomethane hydrochloride, 375 mM KCl, 50 mM DTT) and 4 µL 50 mM $MnCl_2$ were added with 1 µL of SuperScript II and (iii) for AMV RT, 2 µL AMV RT buffer and 1 µL AMV RT were added. All RT reactions were incubated for 1 hr at the following temperatures – TGIRT-III at 60°C, SuperScript II at 50°C, AMV, and HIV RT at 42°C. All RT reactions were inactivated by incubation at 85°C for 5 minutes. To hydrolyze the RNA, 1 µL of 4 M sodium hydroxide was added to the reaction and incubated at 95 °C for 5 min. The reaction was then placed on ice for 5 min. The resulting cDNA was EtOH precipitated and resuspended in loading dye. Samples were run on 12% denaturing PAGE gel for 1 hr at 80V and imaged.

Sequencing for misincorporation analysis

RNA was reacted with maleimide substrates as described previously. Various RT reactions were performed with FL forward primer as described previously. RNA was degraded by adding 5 U RNase H (#M0297, New England Biolabs). To increase diversity for short-read sequencing, each cDNA was amplified for 15 cycles using Q5 high-fidelity DNA polymerase with forward and reverse primers including 0-3 spacers.[1] The length of the amplicons was kept constant by the use of complimentary length primers (primer sequences in DNA sequences table), and all PCR reactions for each RT condition were pooled and purified using Zymo DNA Clean and Concentrator-5 kit. The resulting dsDNA was tagged with Illumina barcodes with 15-cycle PCR using Q5 DNA polymerase, purified with Zymo DNA Clean and Concentrator-5 kit, and analyzed by Bioanalyzer to check for correct amplicon length. Barcoded DNA samples from all RT reactions were pooled at equal concentrations and sequenced on the Illumina MiSeq platform.

Sequencing data analysis

Reads were quality checked using fastQC and trimmed using trimmomatic (LEADING quality threshold: 15, TRAILING quality threshold: 15, MINLEN: 30). Reads were then aligned to the reference sequence using the Burrows-Wheeler Aligner and mutational data was generated from the BAM files using bam-readcount. Mutational data was then cleaned and tabularized using Pandas and visualized with GraphPad.

III. Supplementary Figures

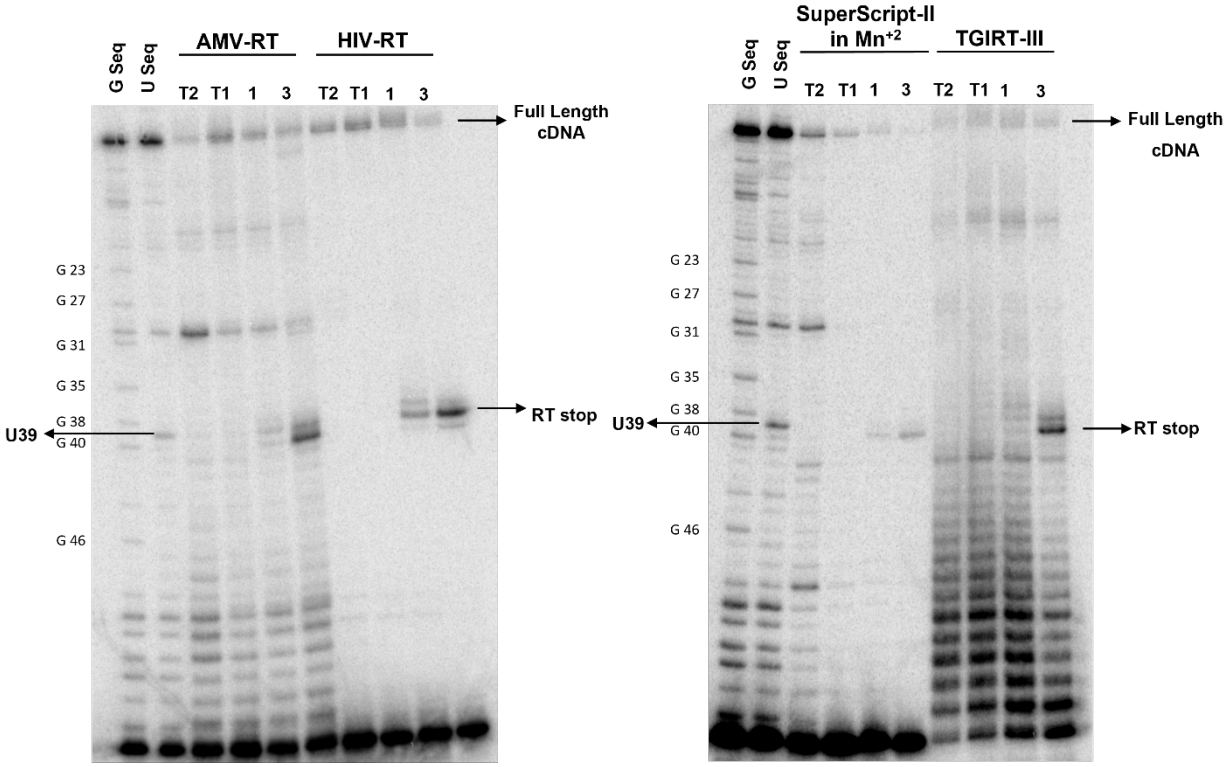

Figure S1. RT Primer extension assay with various RT conditions on T1 treated with 1 or 3.

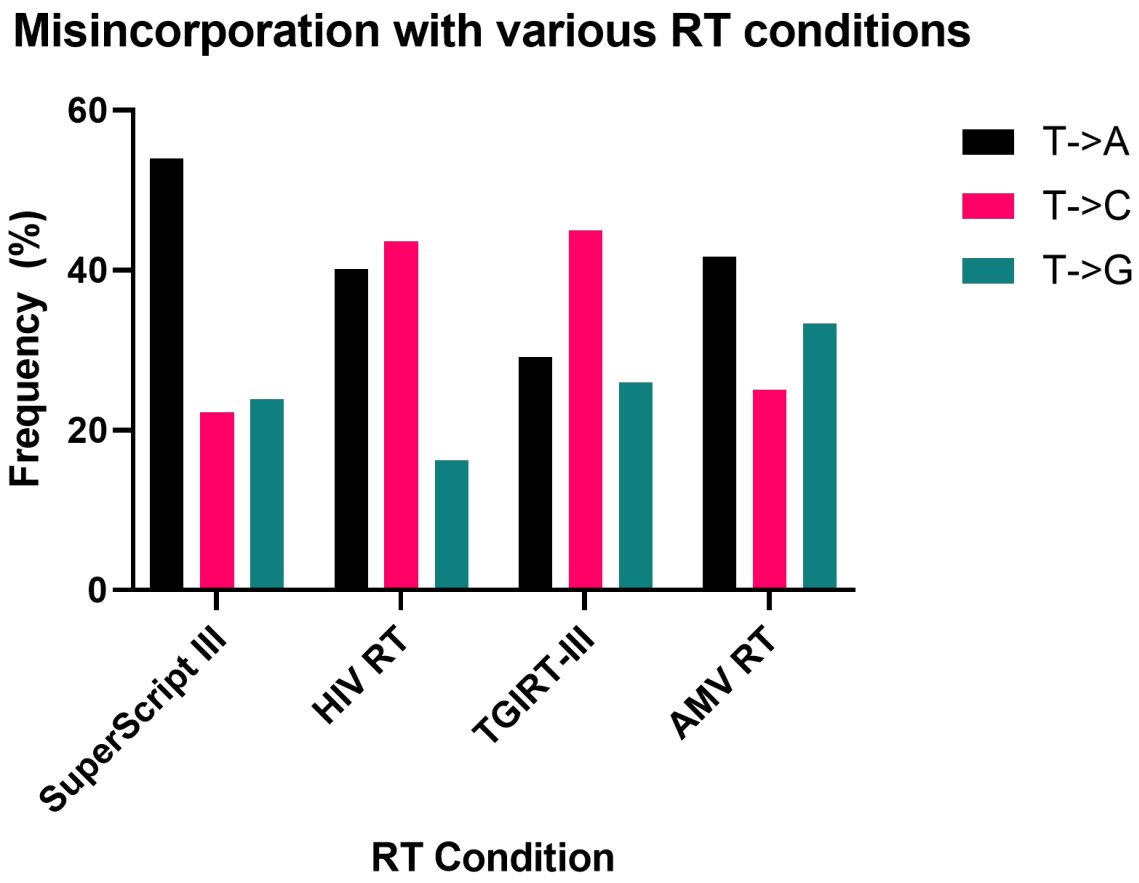

Figure S2. Nucleotide misincorporation pattern with various RT conditions on **T1** treated with **3**.

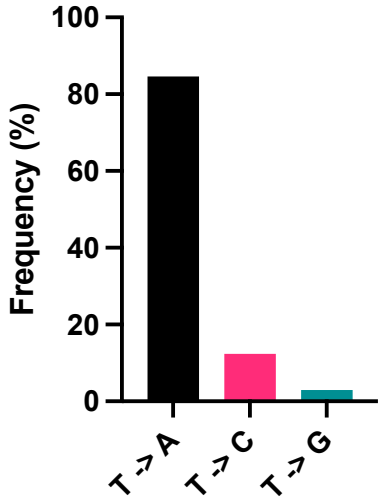

Figure S3. Mutation observed in cDNA from RT with SuperScript II in Mn+2 buffer on **T1** treated with **3.**

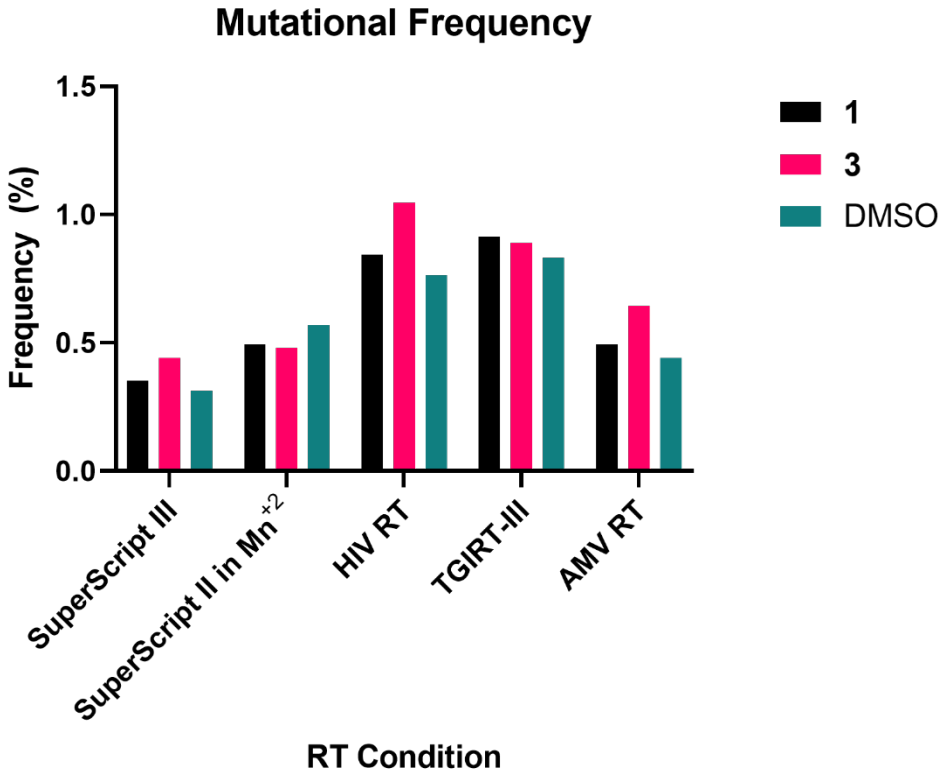

Figure S4. Mutational frequency in cDNA of **T2** treated with **1**, **3**, or DMSO in various RT conditions.

IV. DNA Sequences

DNA Template	5'TGGTCCGCCGTGCCCGGTCTTCGCTGTCGTTTGCACGGCT TTCTGTCGTTCTTTGGGTGCCGCTGTCGTTCGCTATAGTGAG TCGTATTA3'
RT primer	5' TGGTCCGCCGTGCCCGGT3'
FL forward primer	5' TGGTCCGCCGTGCCCGGT3'
FL reverse primer	5' CCACGGCGACAGCAAGCG3'
Forward Phasing Primer 1	5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTGGTCCG CCGTGCCCGGT3'
Forward Phasing Primer 2	5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGATGGTCC GCCGTGCCCGGT3'
Forward Phasing Primer 3	5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGATTGGTCC GCCGTGCCCGGT3'
Forward Phasing Primer 4	5'TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGATATGGTC CGCCGTGCCCGGT3'
Reverse Phasing Primer 1	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCGAACG ACAGCGGCACCC3'

Reverse Phasing Primer 2	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGACGAAC GACAGCGGCACCC3'
Reverse Phasing Primer 3	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGATCGAA CGACAGCGGCACCC3'
Reverse Phasing Primer 4	5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGATACGA ACGACAGCGGCACCC3'

V. NMR Spectra

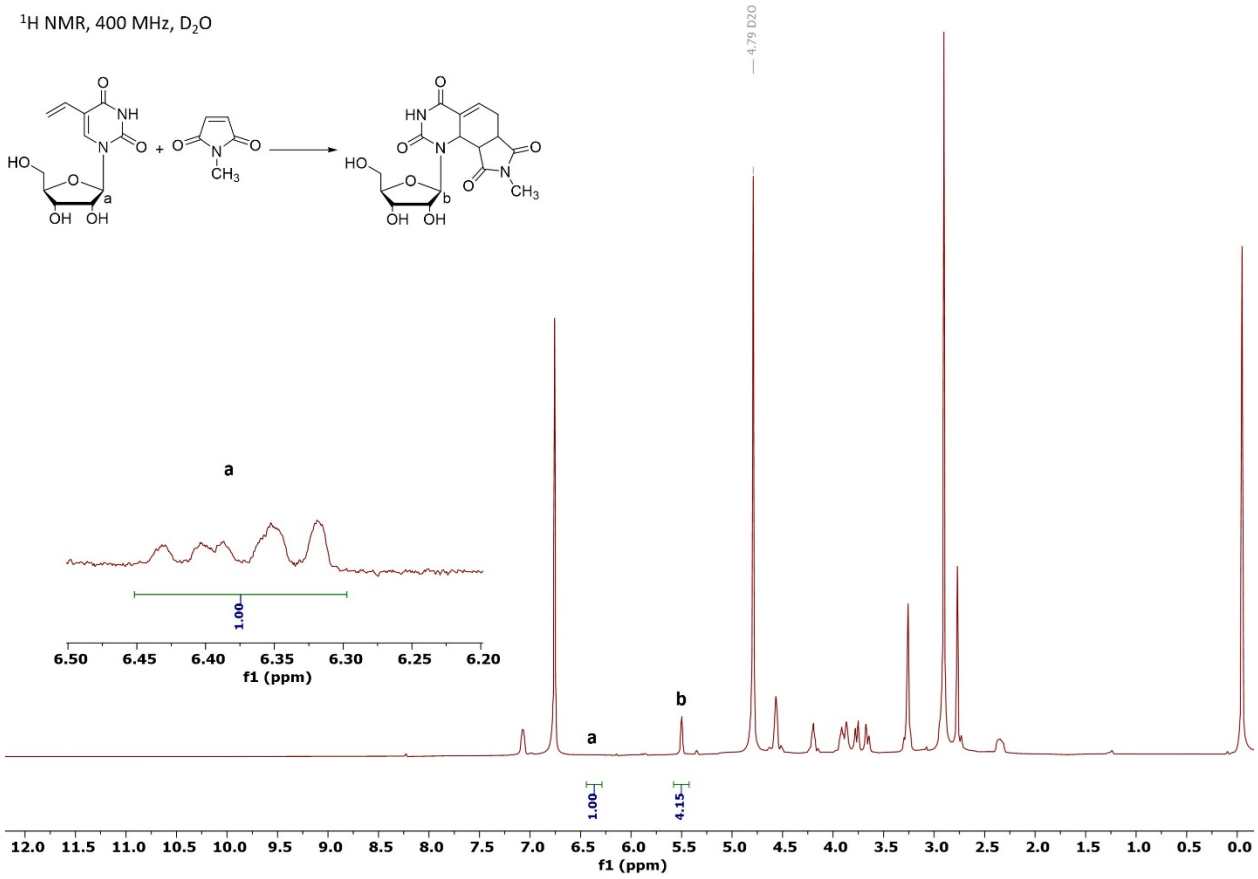

Figure 1. Reaction of 5-VUrd with **2**.

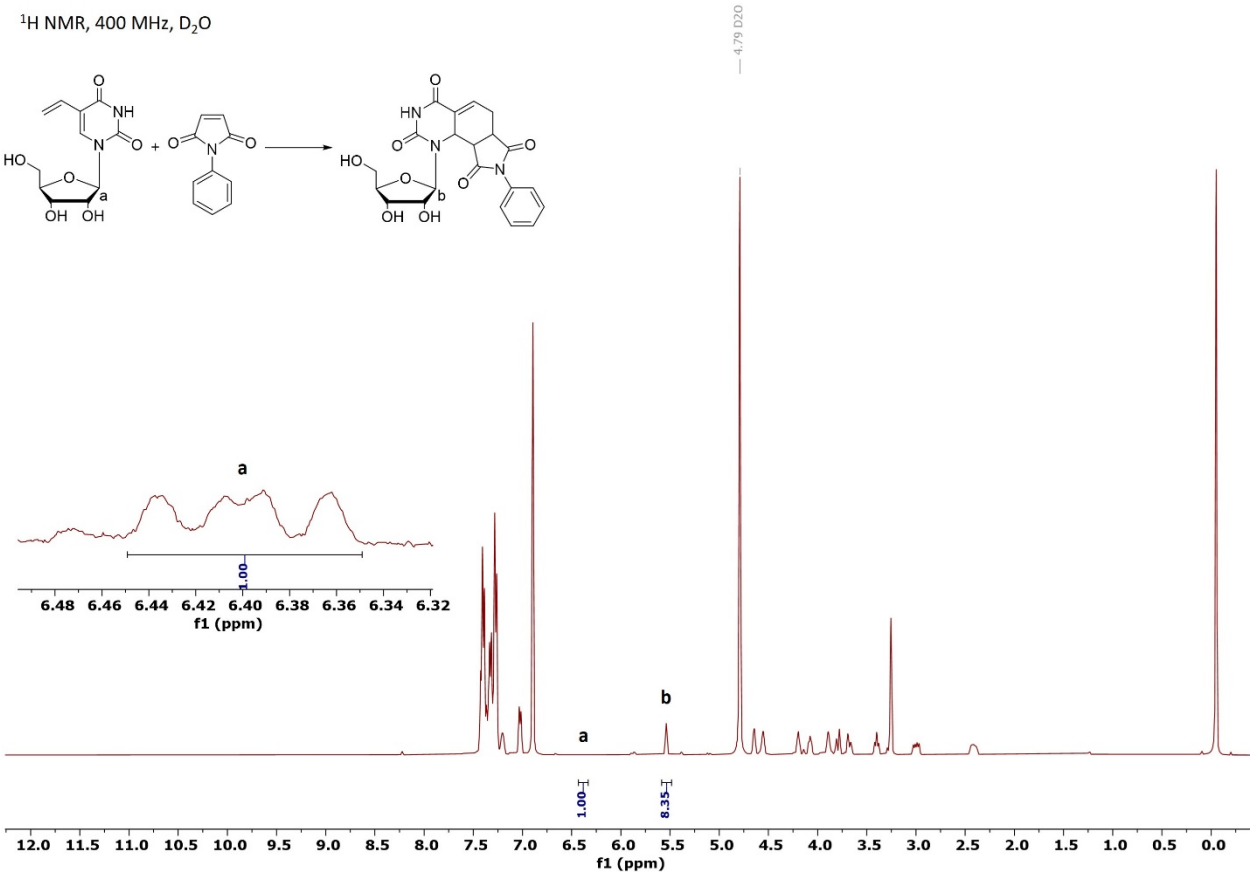

Figure 2. Reaction of 5-VUrd with **3.**

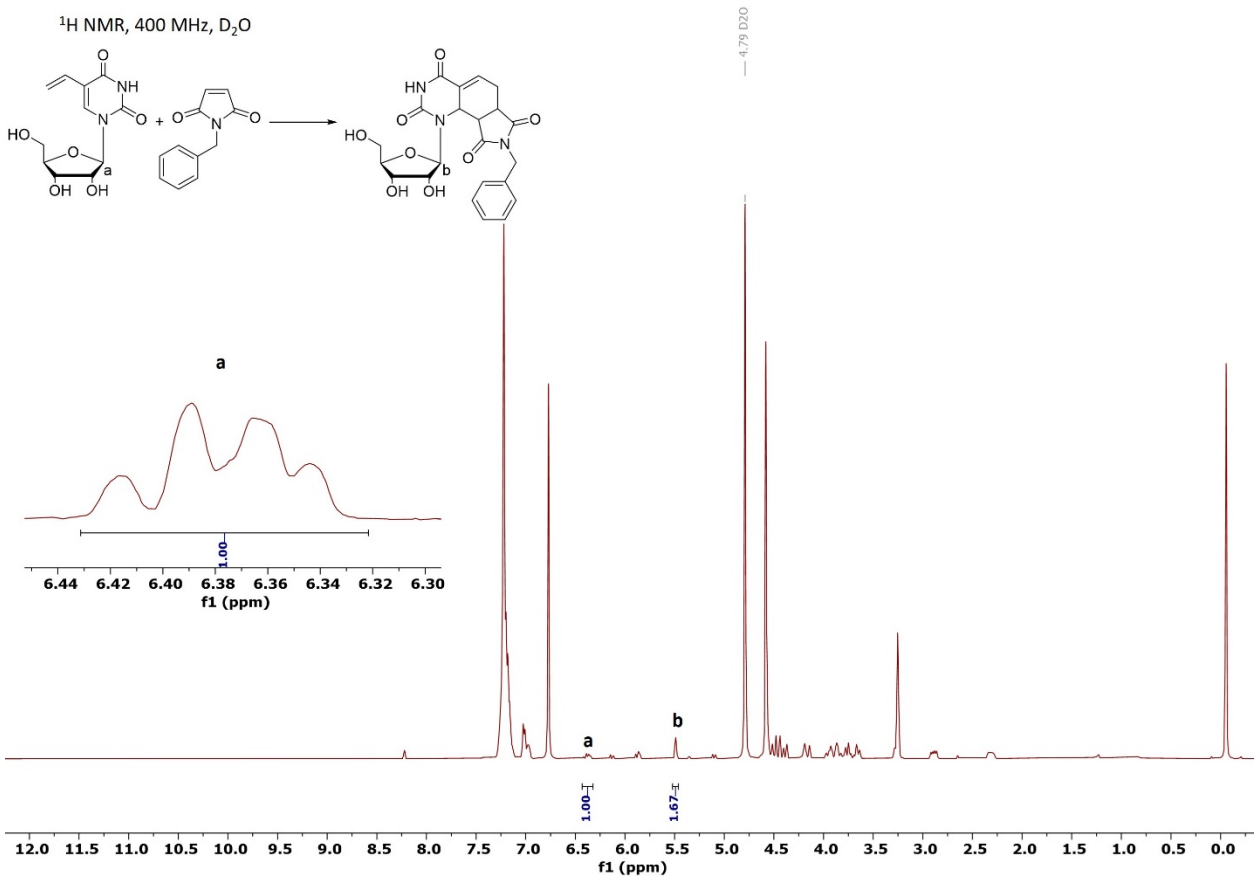

Figure 3. Reaction of 5-VUrd with **4.**

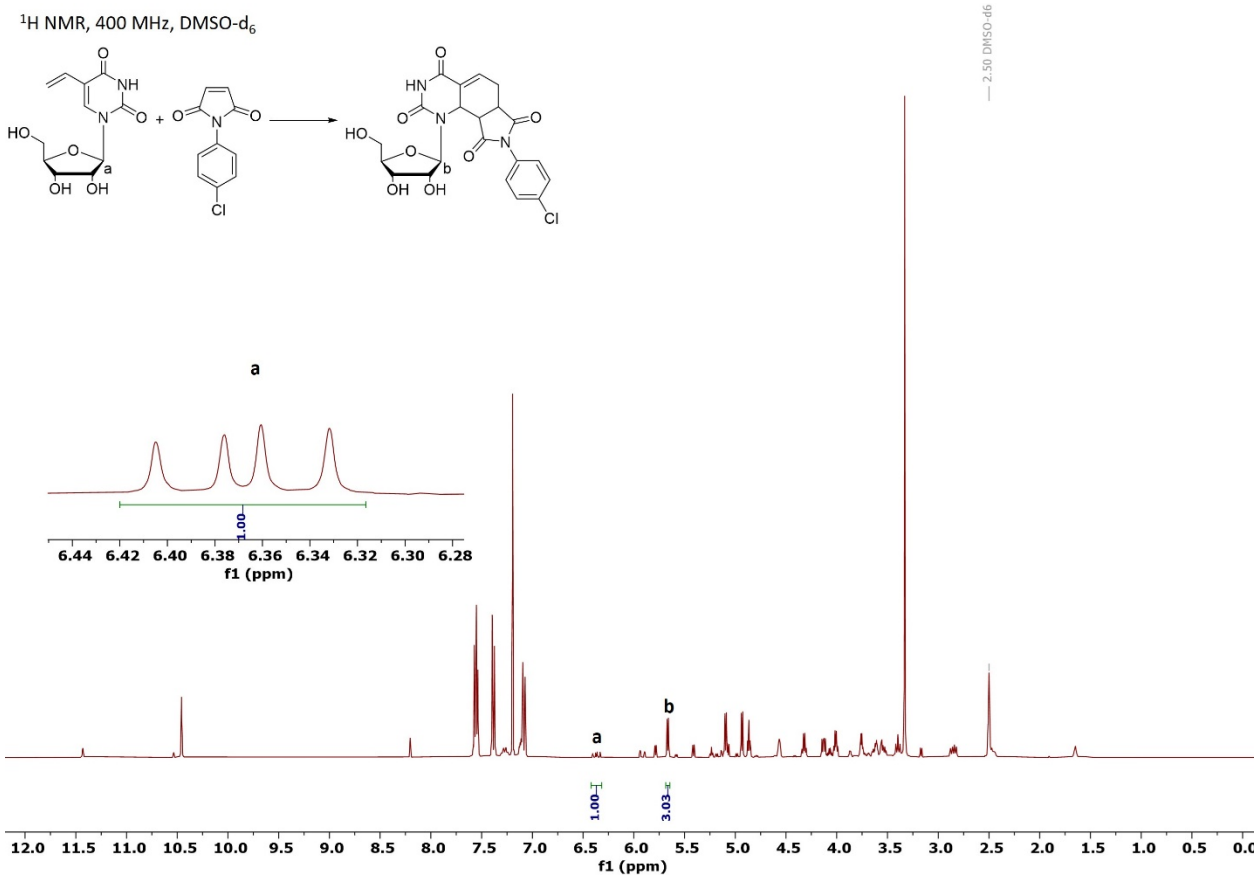

Figure 4. Reaction of 5-VUrd with **5.**

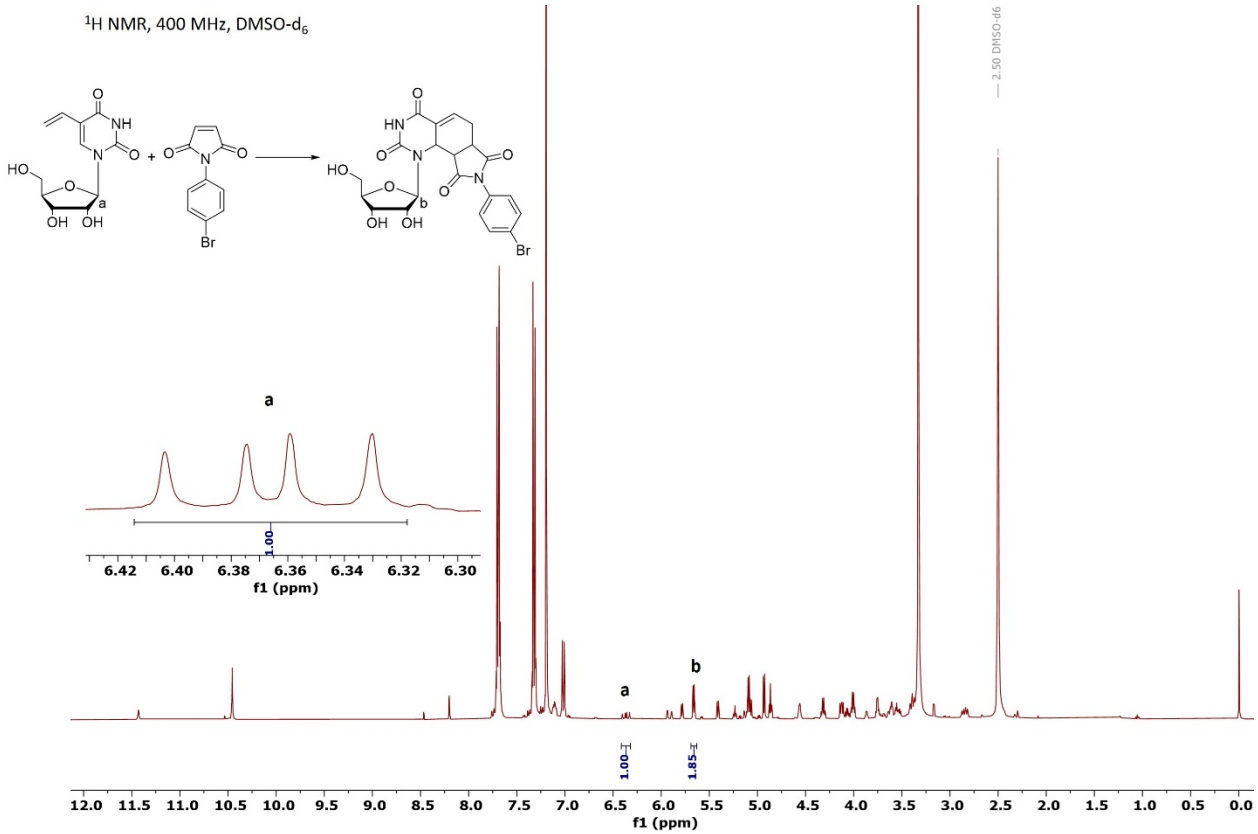

Figure 5. Reaction of 5-VUrd with **6.**

VI. References

(1) Wu, L.; Wen, C.; Qin, Y.; Yin, H.; Tu, Q.; Van Nostrand, J. D.; Yuan, T.; Yuan, M.; Deng, Y.; Zhou, J. Phasing Amplicon Sequencing on Illumina Miseq for Robust Environmental Microbial Community Analysis. *BMC Microbiol.* **2015**, *15* (1), 125. https://doi.org/10.1186/s12866-015-0450-4.