

COMMUNICATION

Reverse transcription through a bulky triazole linkage in RNA: implications for RNA sequencing

Cite this: DOI: 10.1039/x0xx00000x

Xiong Chen^a, Afaf H. El-Sagheer^{b,c} and Tom Brown^{c*}Received 00th January 2012,
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

A triazole linkage can be formed in RNA by strain-promoted chemical ligation of a 3'-azide modified oligonucleotide with a 5'-cyclooctyne labelled oligonucleotide under denaturing conditions. Reverse transcriptase reads through this artificial triazole linkage with omission of one nucleotide. This surprising result has implications for RNA sequencing and other biological applications.

Small RNAs account for about 1% of the total RNA population in eukaryotic cells and include microRNAs (miRNAs), small interfering RNAs (siRNAs) and piwi-interacting RNAs (piRNAs).¹ They are crucial in regulating many biological processes including transcription and translation. Analysis of small RNAs can be carried out by high-throughput next generation sequencing (HTS) techniques.^{2,3} In order to generate RNA sequencing libraries, the isolated native RNAs are normally reverse transcribed into DNA then amplified by PCR. For this procedure the RNAs need to be ligated to adaptors of known sequence at their 3'-end and 5'-ends. Ligation of the adaptors can lead to sequencing bias⁴⁻⁷ due to RNA secondary structures, which drastically reduce ligation efficiency. The resultant distorted representation of the RNA population greatly affects the reliability of studies to measure RNA expression levels.^{7,8} If T4-mediated ligation could be changed to chemical ligation under RNA-denaturing conditions the problem might be overcome. The choice of ligation chemistry is important; copper-catalysed alkyne-azide cycloaddition (CuAAC)⁹ and strain-promoted copper-free¹⁰ click ligation (SPAAC) should be considered, as these reactions are robust and can be performed under conditions in which RNA secondary structures do not exist, e.g. in salt-free water. In order to apply click ligation to RNA sequencing and related applications, a method is required to add the alkyne or the azide function to the 3'-end of the native RNA pool. Such a method has recently been developed to efficiently add a 2'-azido NTP to RNA.¹¹ In addition, the resultant triazole linkage has to be compatible with the reverse transcriptase (RT) enzyme used to produce complementary DNA (cDNA). This has not yet been achieved, but we have previously shown that a specific triazole mimic of the DNA

phosphodiester backbone of the kind depicted in Figure 1A can be read through by DNA and RNA polymerases. The biocompatibility of this artificial linkage has been established in *E. coli*⁹ and is the first example of a non-natural DNA linkage being functional in mammalian cells.¹² RNA strands can also be ligated by the CuAAC reaction to produce catalytically active ribozymes.¹³ With these developments in mind we decided to use this linkage to ligate RNA adaptors to the 3'-end of RNA for applications involving reverse transcription. We now report that RT can read through the above triazole backbone with the omission of one nucleotide. Surprisingly, a much bulkier linkage produced by copper-free click ligation between a strained cyclooctyne and an azide is also read through by reverse transcriptase, as described below.

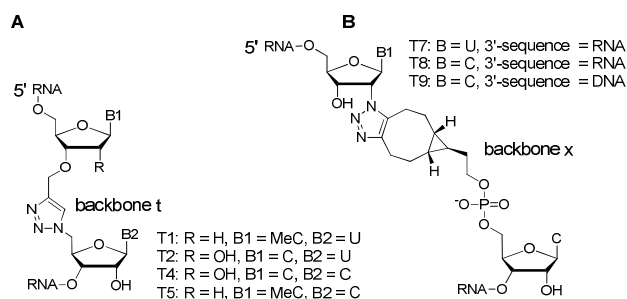


Figure 1. Structures of triazole linkages used in this study

We carried out a series of studies using a synthetic version of miR-155, a micro-RNA which plays a key role in the function of the immune system.¹⁴ An RNA reverse transcription template with the internal MedC-triazole-U linkage MedCtU (T1 in Figure 1A) was synthesized using the CuAAC reaction. Templated click ligation of 3'-propargyl-5-MedC oligonucleotide (K1) (Table S1) and the 5'-azido-U oligonucleotide (Z1) using a complementary splint (S1) afforded T1 in good yield. An additional template (T2) was synthesised in the same way using 3'-propargyl C instead of 3'-propargyl 5-MedC (Figure S1). To study the compatibility of these backbones with reverse transcriptase (RT) enzymes, M-MLV RT and M-MLV RT (RNase H-) were used. The latter is

engineered for reduced RNase H activity and is used to minimize RNA degradation.¹⁵

Five DNA primers of varying lengths (P1 to P5) were used with the two RNA triazole templates T1 and T2. These primers generated reverse transcription starting points at various loci; before the triazole linkage (-7, P1), (-3, P2), next to the triazole (+0, P3), one nucleotide after (+1, P4) and four nucleotides after (+4, P5). In the last two cases the triazole linkage is bridged by the bound primer. Reverse transcription of the triazole templates T1 and T2 stopped around the triazole site when primers P1, P2 and P3 were used, but gave the full length product with the bridged primers P4 and P5. Addition of Mg^{2+} to the Mg^{2+} -free buffer allowed the enzyme to read through the triazole¹⁶ even with the non-bridged primers. This gave a product that lacked a single deoxyguanosine, as shown by polyacrylamide gel-electrophoresis (PAGE) and mass spectrometry (MS) (Figures S2 and S3). This omission is due to the enzyme not recognising the presence of the template C base adjacent to the triazole linkage. The inability of RT to read through the triazole linkage efficiently in the presence of Mg^{2+} could be partly due to instability in the duplex formed between the triazole template and the primer.¹⁷ Therefore we changed the dinucleotide around the triazole linkage from CtU to give a more stable template/primer duplex. Two new templates T4 (CtC) and T5 (MedCtC) were synthesized by CuAAC click ligation and encouragingly M-MLV RT read through both triazole linkages in Mg^{2+} buffer with bridged and non-bridged primers (Figure 2). This indicates that duplex stabilisation allows the enzyme to copy the RNA template accurately with bridged primers, but with omission of one nucleotide after the triazole linkage with non-bridged primers (Figure S4).

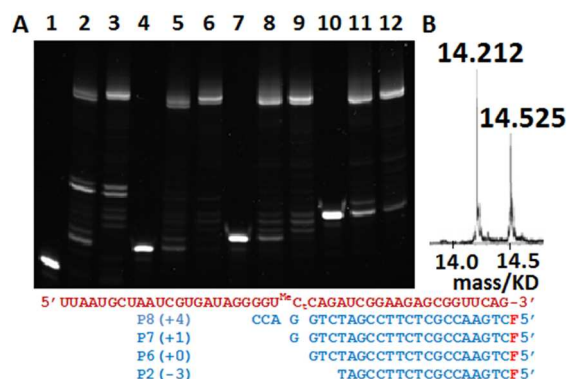


Figure 2. A) Bridged and non-bridged primers read through triazole linkage. 20% PAGE for reverse transcription of T5 (MedCtC, lanes 2, 5, 8, 11) and control T6 (lanes 3, 6, 9, 12) using M-MuLV Reverse Transcriptase in Mg^{2+} buffer at 37 °C for 2 h. Lane 1, 2 and 3; primer P2(-3) and reverse transcription using it, lane 4, 5 and 6; primer P6 (+0) and reverse transcription using it; lane 7, 8 and 9; primer P7 (+1) and reverse transcription using it; lane 10, 11 and 12; primer P8 (+4) and reverse transcription using it. B) Mass spectrum of reverse transcription product using template T5 and primer P7. Calc. Mass: 14212 (M); 14525 (M + dA).

Copper-free click ligation has potential advantages over the CuAAC reaction as it will not lead to degradation of RNA in biological media. In this context the ring-strain promoted alkyne-azide [3+2] cycloaddition (SPAAC) reaction has recently been applied to DNA strand ligation,^{10, 18} so we decided to investigate its use in templates for reverse transcription. We chose bicyclo[6.1.0]nonyne (BCN) as the strained alkyne¹⁹ because the resultant triazole linker is less bulky than the alternatives, and is therefore more likely to be accommodated by the RT enzyme. Two RNA templates T7 and T8 with internal UxC and CxC were synthesised by SPAAC

ligation of 5'-BCN oligonucleotide O1 with 2'-azide oligonucleotides Z2 or Z3 to produce 5'-2' linkages in the RNA backbone (Figure 1B). The key building blocks for the synthesis of the 3'-azide labelled RNA strand are 5'-DMT-2'-azido-2'-dU and 5'-DMT-2'-azido-2'-dC.²⁰ These were synthesised, coupled to succinylated aminoalkyl solid supports and used in the synthesis of 2'-azide oligonucleotides. To synthesise the cyclooctyne oligonucleotide O1, BCN phosphoramidite was added to the 5'-terminus during solid-phase synthesis. This was reacted with the 2'-azide oligonucleotides in water to generate the RNA templates T7 and T8 in excellent yield (Figure 3A). For ease of purification, and to mimic ligation of biologically-derived RNA, no template was used; instead concentrated reaction mixtures were employed to ensure good ligation yields.

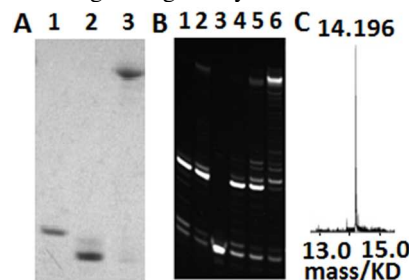


Figure 3. Copper-free RNA ligation and reverse transcription. A) 20% PAGE for non-templated copper free click ligation to synthesize (T8). Lane 1; 2'-azide oligo Z3, lane 2; alkyne oligo O1, lane 3; crude reaction mixture. B) 20% PAGE for reverse transcription of T8 and T9. Lane 1 and 2; reverse transcription of template T9 at 2h and 18h, lane 3; primer P1 (-7); lane 4, 5 and 6 reverse transcription of template T8 at 10 min, 2h and 18h. All reactions in Mg^{2+} buffer. C) Mass spectrum of RT products from template T8 and primer P7 after incubation at 37 °C overnight, Calc. mass: 14196 (M - dG + dA).

Because the BCN/azide triazole linker generated in the SPAAC reaction is much bulkier than the linker generated by CuAAC, and is a 5'-2' linkage rather than 5'-3', it was expected to hinder reverse transcription. Consequently, Click BCN backbones were reverse transcribed for an extended period (overnight) and M-MLV RT (RNase H-) was used to avoid RNA degradation. First the template with a CxC linkage (T8) was investigated as it would be expected to give a more stable duplex than the UxC linkage (T7). After 2h incubation at 37 °C, both PAGE and MS showed that the RT enzyme paused at the triazole site (CxC), but then copied through during overnight incubation (Figure 3B, lanes 5 and 6 and S5). The 2h incubation gave two products; "stopped before triazole" and "full-length minus deoxyguanosine" ("-dG"). After overnight incubation, the "stopped before triazole" product disappeared, the primer extended to the end of the template and produced the "full-length -dG + dA" product (Figure 3C). This extra dA was most likely added to the 3'-end of the RT product. Changing the divalent cation in the buffer from Mg^{2+} to Mn^{2+} accelerated the reverse transcription reaction, which was nearly completed in 2h instead of overnight (Figure 4A lanes 5 and 6) and the (full length - dG + dA) product was obtained. The shortest P1 (-7) primer worked more efficiently than the bridged primer P4 (+1), indicating that it is easier for the enzyme to pass the triazole-BCN linkage when it has a "running start" from a stretch of unmodified RNA than to start on top of the triazole-BCN link. This inefficiency of bridged primer extension was not observed when using the RNA template with the biocompatible CtC triazole linkage (T5), for which the bridged primer worked as efficiently as the -7 primer. This difference in behaviour can be explained by the length and bulkiness of the

BCN triazole backbone. Template T7 with the sequence “UxC” instead of “CxC” was also investigated. It was anticipated that it would be copied inefficiently, as it would be expected to give a less stable duplex than the corresponding CxC template. Indeed, reverse transcription of this backbone in Mg^{2+} buffer read through the triazole-BCN linkage with difficulty. A reaction buffer containing 3 mM Mg^{2+} and 3 mM Mn^{2+} improved the situation, but reverse transcription produced a mixture of products including one and two nucleotide deletions. Both DNA²¹ and RNA oligonucleotides have been used as 3'-adaptors for RNA ligation.²² However, M-MLV RT reads through RNA more efficiently than DNA. To evaluate if this is also the situation with BCN-triazole-modified templates, T9 with a DNA adaptor instead of RNA adaptor was synthesized. Compared to template T8 which contains an RNA adaptor, reverse transcription of T9 was much less efficient (Figure 3B lanes 1 and 2). Addition of Mn^{2+} to the buffer greatly improved the situation but reverse transcription of T8 was still faster and finished in 2h compared to 16h for T9.

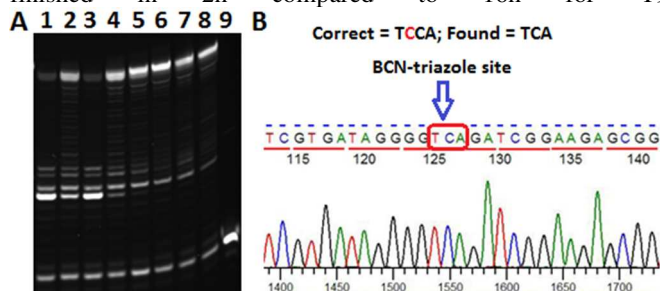


Figure 4. A) Reverse transcription of T8 in presence of Mg^{2+} or Mn^{2+} ions for 2h and 7h respectively. 20% PAGE. Lane 1 and 2; 3 mM Mg^{2+} , lane 3 and 4; 10 mM Mg^{2+} , lane 5 and 6; 3 mM Mn^{2+} , lane 7 and 8; 10 mM Mn^{2+} , lane 9; primer P1 (-7). B) Sanger sequencing analysis of the DNA. The cytidine (C) in red at the TCxCA site in the template is the one that is not copied by RT.

Sanger sequencing of reverse transcribed DNA showed that the RT products from the BCN-triazole templates formed by copper-free click chemistry failed to copy one of the two bases around the modified linkage (Figure 4B and S7). However, this would not be an issue in RNA sequencing applications; the identity of this base would be known, as it would be added to the 3'-end of the native RNA pool using an azide-modified NTP and Yeast poly-A polymerase (PAP).¹¹

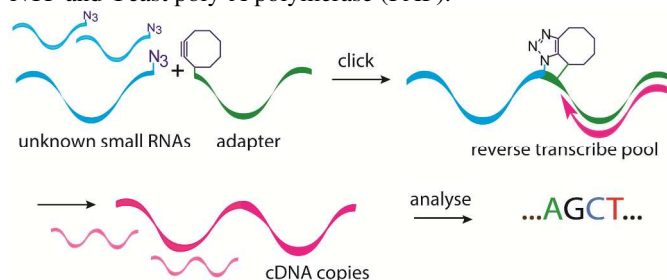


Figure 5. SPAAC ligation, reverse transcription and c-DNA analysis

In conclusion, using bridged primers the M-MLV reverse transcriptase enzyme accurately copies an RNA template containing biocompatible triazole linkage 1A. When templates containing this or the BCN-triazole linkage 1B are used with normal primers, one nucleotide is omitted. The use of the SPAAC reaction to ligate RNA strands is significant because it should enable small natural RNAs to be linked to synthetic adaptors in non-templated copper-free RNA-denaturing conditions without degradation (Figure 5). The ability of RT to

read through bulky linkages in RNA is likely to be a general property extending to other modified nucleic acid backbones. These findings could have applications in RNA isolation, sequence analysis and other biological applications.

Acknowledgments

This research was funded by the BBSRC sLOLA grant BB/J001694/1 “Extending the boundaries of nucleic acid chemistry.”

Notes and references

^a School of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ. ^b Chemistry Branch, Dept. of Science and Mathematics, Faculty of Petroleum and Mining Engineering, Suez University, Suez, 43721, Egypt. ^c Department of Chemistry, University of Oxford, Chemistry Research Laboratory, 12 Mansfield Road, Oxford, OX1 3TA, UK. *e.mail: tom.brown@chem.ox.ac.uk.

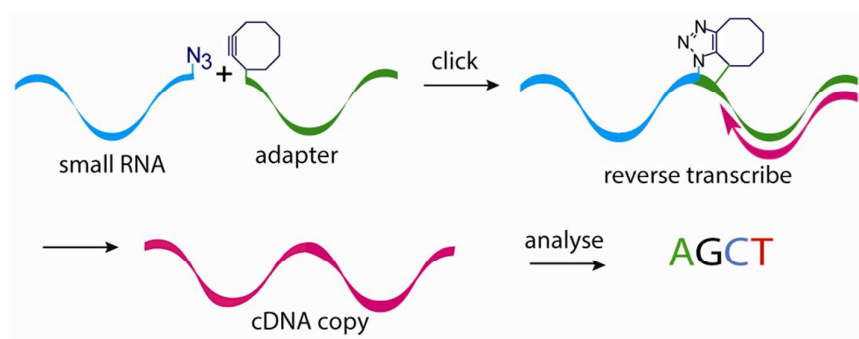
Electronic Supplementary Information (ESI) available: Chemical synthesis, table of oligonucleotide sequences, other reverse transcriptase reactions, DNA sequencing. See DOI: 10.1039/c000000x/

1. R. W. Carthew and E. J. Sontheimer, *Cell*, 2009, 136, 642-655.
2. D. C. Koboldt, K. M. Steinberg, D. E. Larson, R. K. Wilson and E. R. Mardis, *Cell*, 2013, 155, 27-38.
3. L. Mamanova and D. J. Turner, *Nature Protocols*, 2011, 6, 1736-1747.
4. S. E. V. Linsen, E. de Wit, G. Janssens, S. Heater, L. Chapman, R. K. Parkin, B. Fritz, S. K. Wyman, E. de Bruijn, E. E. Voest, S. Kuersten, M. Tewari and E. Cuppen, *Nature Methods*, 2009, 6, 474-476.
5. M. Hafner, N. Renwick, M. Brown, A. Mihailovic, D. Holoch, C. Lin, J. T. G. Pena, J. D. Nusbaum, P. Morozov, J. Ludwig, T. Ojo, S. J. Luo, G. Schroth and T. Tuschl, *RNA-a Publication of the RNA Society*, 2011, 17, 1697-1712.
6. A. D. Jayaprakash, O. Jabado, B. D. Brown and R. Sachidanandam, *Nucleic Acids Res.*, 2011, 39.
7. F. L. Zhuang, R. T. Fuchs, Z. Y. Sun, Y. Zheng and G. B. Robb, *Nucleic Acids Res.*, 2012, 40, 14.
8. K. Sorefan, H. Pais, A. E. Hall, A. Kozomara, S. Griffiths-Jones, V. Moulton and T. Dalmay, *Silence*, 2012, 3, 4.
9. A. H. El-Sagheer, A. P. Sanzone, R. Gao, A. Tavassoli and T. Brown, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, 108, 11338-11343.
10. M. Shelbourne, A. H. El-Sagheer and T. Brown, *Chem. Commun.*, 2012, 48, 11184-11186.
11. M. L. Winz, A. Samanta, D. Benzinger and A. Jaschke, *Nucleic Acids Res.*, 2012, 40 e78.
12. C. N. Birts, A. P. Sanzone, A. H. El-Sagheer, J. P. Blydes, T. Brown and A. Tavassoli, *Angew. Chem. Int. Edit.*, 2014, DOI: 10.1002/anie.201308691.
13. A. H. El-Sagheer and T. Brown, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, 107, 15329-15334.
14. A. Rodriguez, E. Vigorito, S. Clare, M. V. Warren, P. Couttet, D. R. Soond, S. van Dongen, R. J. Grocock, P. P. Das, E. A. Miska, D. Vetrie, K. Okkenhaug, A. J. Enright, G. Dougan, M. Turner and A. Bradley, *Science*, 2007, 316, 608-611.
15. M. L. Kotewicz, C. M. Sampson, J. M. Dalessio and G. F. Gerard, *Nucleic Acids Res.*, 1988, 16, 265-277.
16. E. Scolnick, E. Rands, S. A. Aaronson and G. J. Todaro, *Proc. Natl. Acad. Sci. U. S. A.*, 1970, 67, 1789-8.
17. A. Dallmann, A. H. El-Sagheer, L. Dehm, C. Mugge, C. Griesinger, N. P. Ernting and T. Brown, *Chem.-Eur. J.*, 2011, 17, 14714-14717.
18. M. Shelbourne, X. Chen, T. Brown and A. H. El-Sagheer, *Chem. Commun.*, 2011, 47, 6257-6259.
19. J. Dommerholt, S. Schmidt, R. Temming, L. J. A. Hendriks, F. Rutjes, J. C. M. van Hest, D. J. Lefebvre, P. Friedl and F. L. van Delft, *Angew. Chem. Int. Edit.*, 2010, 49, 9422-9425.
20. M. Aigner, M. Hartl, K. Fauster, J. Steger, K. Bister and R. Micura, *ChemBioChem*, 2011, 12, 47-51.
21. M. Hafner, P. Landgraf, J. Ludwig, A. Rice, T. Ojo, C. Lin, D. Holoch, C. Lim and T. Tuschl, *Methods*, 2008, 44, 3-12.
22. R. D. Morin, M. D. O'Connor, M. Griffith, F. Kuchenbauer, A. Delaney, A.-L. Prabhu, Y. Zhao, H. McDonald, T. Zeng, M. Hirst, C. J. Eaves and M. A. Marra, *Genome Res.*, 2008, 18, 610-621.

Reverse transcription through a bulky triazole linkage in RNA: implications for RNA sequencing

Xiong Chen, Afaf H. El-Sagheer and Tom Brown

Reverse-transcriptase reads through the bulky artificial linkage with omission of one nucleotide with implications for RNA isolation and analysis



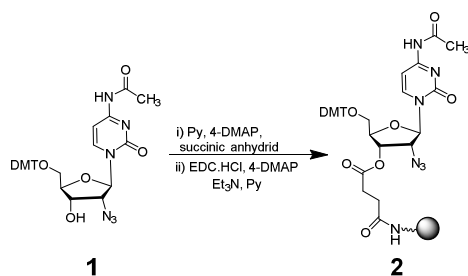
Supporting Information

Reverse transcription through a bulky triazole linkage in RNA: implications for RNA sequencing

Xiong Chen^a, Afaf H. El-Sagheer^{b,c} and Tom Brown.^{c*}

^aSchool of Chemistry, University of Southampton, Highfield, Southampton. SO17 1BJ, UK. ^bChemistry Branch, Dept. of Science and Mathematics, Faculty of Petroleum and Mining Engineering, Suez University, Suez, 43721, Egypt. ^cDepartment of Chemistry, University of Oxford, Chemistry Research Laboratory, 12 Mansfield Road, Oxford, OX1 3TA, UK. *e.mail: tom.brown@chem.ox.ac.uk.

Preparation of 5'-O-(4,4'-dimethoxytrityl)-2'-azido-2'-deoxycytidine on solid support (2)



The solid support, Amino-SynBase resin 1000/100 (Link Technologies, Glasgow, UK) (1000 Å pore size, loading 59 µmol/g, 0.5 g, 29.5 µmol) was treated with 3% trichloroacetic acid (TCA) in dichloromethane (DCM) for 4 h in a stoppered glass vessel fitted with a sinter and tap. The solvents were then removed by filtration and the support was washed with triethylamine:di-isopropylethylamine (9:1), dichloromethane (DCM) and diethyl ether. The support was dried under vacuum for 1 h then soaked in dry pyridine for 10 min. A solution of succinic anhydride (250 mg, 2.5 mmol) and 4-dimethylaminopyridine (DMAP) (25 mg, 0.21 mmol) in dry pyridine (5 mL) was added and the vessel was rotated for 20 h. The support was washed with pyridine, DCM and diethyl ether, dried and soaked in pyridine for 10 min. A solution of ethyldimethylaminopropylcarbodi-imide hydrochloride (EDC.HCl) (96 mg, 0.5 mmol), DMAP (4 mg, 0.033 mmol) and triethylamine (20 µL) in dry pyridine (5 mL) was added to the solid support in the vessel followed by compound **1** (90 mg, 0.147 mmol)¹ in dry pyridine (2 mL). The reaction vessel was left to rotate for 20 h at room temperature then pentachlorophenol (34 mg, 0.13 mmol) was added and the vessel was rotated for a further 3 h. The solvent was then removed by filtration and the support was washed with pyridine, DCM and diethyl ether. Piperidine (10% in DMF, 10 mL) was added and after rotating the

vessel for 1 min the solid support was washed with DCM and diethyl ether (note that during treatment of the support with piperidine to cap the unreacted succinic acid groups the loading of nucleoside decreased with time due to cleavage of the succinyl linkage. Therefore only a brief 1 min piperidine treatment was carried out).

Capping reagent (oligonucleotide synthesis grade, acetic anhydride/pyridine/tetrahydrofuran: *N*-methyl imidazole in tetrahydrofuran, 1:1, 10 mL, Applied Biosystems) was added and the vessel was rotated for 1 h after which the support was washed with pyridine, DCM and diethyl ether, then left to dry under vacuum overnight. The loading of (**1**) on the support was 25 $\mu\text{mol/g}$, as determined colorimetrically from the cleaved DMT group.

5'-*O*-(4,4'-dimethoxytrityl)-2'-azido-2'-deoxyuridine was attached to the solid support in the same manner as explained above.

Oligonucleotide Synthesis and purification

General method

Standard DNA phosphoramidites, solid supports, and additional reagents were purchased from Link Technologies, Berry&Associates, Glen Research, Sigma-Aldrich and Applied Biosystems Ltd. All oligonucleotides were synthesized on an Applied Biosystems 394 automated DNA/ RNA synthesizer using a standard 0.2 or 1.0 μmol phosphoramidite cycle of acid-catalyzed detritylation, coupling, capping, and iodine oxidation. Stepwise coupling efficiencies and overall yields were determined by the automated trityl cation conductivity monitoring facility and in all cases were >98.0%. All β -cyanoethyl phosphoramidite monomers were dissolved in anhydrous acetonitrile to a concentration of 0.1 M immediately prior to use. The coupling time for normal A, G, C, and T monomers was 40 s and the coupling time for the 5'-BCN phosphoramidite monomer was extended to 480 s. 2'-azido oligonucleotides were synthesized on the 1.0 μmol scale using the 5'-*O*-(4,4'-dimethoxytrityl)-2'-azido-2'-deoxy(cytidine) or (uridine) solid support (25 $\mu\text{mol/g}$ loading) which were synthesised as explained above. The resin was packed into a twist column (Glen Research) then used to assemble the required oligonucleotide sequence in the 3'- to 5'-direction by standard phosphoramidite oligonucleotide methods. Cleavage of oligonucleotides from the solid support and deprotection was achieved by exposure to concentrated aqueous ammonia solution for 60 min at room temperature followed by heating in a sealed tube for 5 h at 55 °C.

RNA oligonucleotides were prepared using 2'-TBS protected RNA phosphoramidite monomers with *t*-butylphenoxyacetyl protection of the A, G and C nucleobases and unprotected U (Sigma-Aldrich). A solution of 0.3 M benzylthiotetrazole in acetonitrile (Link Technologies) was used as the coupling agent, *t*-butylphenoxyacetic anhydride as the capping agent and 0.1 M iodine as the oxidizing agent (Sigma-Aldrich). All phosphoramidite monomers were dissolved in anhydrous acetonitrile to a concentration of 0.1 M immediately prior to use, and the coupling time for all monomers was 6 min. Stepwise coupling efficiencies were determined by automated trityl cation conductivity monitoring and in all cases were >97%. Cleavage of oligonucleotides from the solid support and deprotection were achieved by exposure to concentrated aqueous ammonia/ethanol (3/1 v/v) for 2 h at room temperature followed by heating in a sealed tube for 45 min at 55 °C.

2'-TBS deprotection of RNA oligoribonucleotides

After cleavage from the solid support and deprotection of the nucleobases and phosphodiester, oligonucleotides were concentrated to a small volume *in vacuo* (until turbidity starts to appear), transferred to 15 mL plastic tubes and freeze dried. The residue was dissolved in DMSO (300 µL) and triethylamine trihydrofluoride (300 µL) was added after which the reaction mixtures were kept at 65 °C for 2.5 h. Sodium acetate (3 M, 50 µL) and butanol (3 mL) were added with vortexing and the samples were kept at -80 °C for 30 min then centrifuged at 4 °C at 13,000 rpm for 10 min. The supernatant was decanted and the precipitate was washed twice with ethanol (0.75 mL) then dried under vacuum.

Purification of oligonucleotides (DNA or RNA)

The fully deprotected oligonucleotides were purified by reversed-phase HPLC on a Gilson system using a Luna 10µ C8 100Å pore Phenomenex 10x250 mm column with a gradient of acetonitrile in triethylammonium acetate or ammonium acetate (0% to 50% buffer B over 20 min, flow rate 4 mL/min), (buffer A: 0.1 M triethylammonium acetate, pH 7.0, buffer B: 0.1 M triethylammonium acetate, pH 7.0, with 50% acetonitrile). Elution was monitored by UV absorption at 295 nm. After HPLC purification, oligonucleotides were desalted using NAP-25 then NAP-10 columns (GE Healthcare). For RNA oligonucleotides, HPLC using triethylammonium bicarbonate buffer (buffer A: 0.1 M triethylammonium bicarbonate, pH 7.5, buffer B: 0.1 M triethylammonium bicarbonate, pH 7.5, with 50% acetonitrile) was used

to avoid desalting and degradation of the RNA. The fractions from HPLC were evaporated without need for additional desalting.

All oligonucleotides were characterised by negative-mode electrospray HPLC-mass spectrometry in water, using a Bruker Daltronics microTOF mass spectrometer, using an Acquity UPLC BEH C18 1.7 μm HPLC column (Waters), with a gradient of TEAA/ CH_3CN in TEAA/HFIP buffer, increasing from 5-40% buffer B over 14 minutes, with a flow rate of 0.1 mL min^{-1} (buffer A: 10 mM TEAA, 100 mM HFIP (H_2O); buffer B: 20 mM TEAA (CH_3CN)). Raw data was processed/deconvoluted using the DataAnalysis function of the Bruker Daltronics CompassTM 1.3 software package.

Non-templated and template copper catalysed click reactions

A solution of degassed Cu^{I} click catalyst was prepared from *tris*-hydroxypropyltriazole ligand (THPTA) (2.1 mg in 25 μL water), sodium ascorbate (14 μL of 0.5 M solution in water) and $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ (7 μL of 0.1 M solution in water). In case of the non-templated reaction, this solution was added to a solution of the azide and alkyne oligonucleotides (30 nmol of each in 30 μL water). In case of the templated reaction, a splint oligonucleotide (30 nmol) was added to the mixture of the azide and alkyne oligonucleotides followed by annealing before adding the copper catalyst solution. The reaction mixture was kept under argon at r.t. for 2 h. Reagents were removed using NAP-25 gel-filtration columns and the ligated RNA was analysed and purified by denaturing 20% polyacrylamide gel electrophoresis (PAGE). The product bands were cut then soaked in Tris-HCl buffer (50 mM Tris-base, 25 mM NaCl, pH 7.5) at 37 $^\circ\text{C}$ overnight and desalted by NAP-25 followed by NAP-10 gel-filtration then lyophilized.

Non-templated copper free click reaction

A mixture of the alkyne (5'-BCN) (DNA or RNA) and azide (2'-azide) (RNA) oligonucleotides (30 nmol each) in 60 μL water was kept at room temperature for 2 h then lyophilized before loading to 20% polyacrylamide gel electrophoresis.

The product bands were cut then soaked in Tris-HCl buffer (50 mM Tris-base, 25 mM NaCl, pH 7.5) in case of RNA, or in water in case of DNA at 37 $^\circ\text{C}$ overnight, then desalted by NAP-25 followed by NAP-10 gel-filtration and lyophilized.

Reverse transcription of BCN templates

In a total reaction volume of 20 μL , 3 μM primer, 3 μM template, 1x M-MuLV Reverse Transcriptase (RNase H⁻) reaction buffer (NEB, 50 mM Tris-HCl, 75 mM KCl, 3 mM MgCl₂, pH 8.3 at r.t.), 10 mM DTT (NEB), 0.5 mM dNTP (Promega), 200 U M-MuLV Reverse Transcriptase (1 μL , NEB[®]) or Reverse Transcriptase (RNase H⁻, 1 μL , NEB[®]) were mixed and incubated on a BIO-RAD T100TM Thermal Cycler at 37 °C for different time courses. In some experiments, a specific concentration of Mn²⁺ was added to Mg²⁺ free buffer. In other experiments, Mn²⁺ was added to the buffer in the presence of Mg²⁺. The enzymatic reaction was stopped by freezing the sample in liquid nitrogen then mixing with an equal volume of formamide and directly loaded onto the gel for PAGE analysis (20 % acrylamide, 600v for 5 h) or diluted to 1 mL and desalted by NAP-10 then freeze-dried overnight for HPLC-MS analysis. To give clear spectra, the RNA template in some HPLC-MS samples was digested by RNase H after reverse-transcription (NEB, 1x MuLV RT (RNase H⁻) buffer, 10 mM additional DTT, 5 U RNase H (25 μL) was added to 50 μL reverse-transcription solution, incubated at 37 °C for 6 h, gel-filtered (NAP-10), lyophilized then dissolved in 10 μL water for MS.

To minimize the co-migration of the RNA template and the DNA products during PAGE that results from duplex formation, a 10-fold excess of the full length unlabelled DNA strand complementary to the RNA template was added together with the reverse-transcription product. The mixture was heated at 90 °C for 10 min then cooled down to RT. The samples were then mixed with an equal volume of formamide and loaded for PAGE.

Table 1: oligonucleotide used in this study

1a alkyne and azide ODNs and splints

Code	Sequence (5' - 3')	5'	3'	Calc.	Found
K1	r ^p UUA AUGCUAAUCGUGAUAGGGGU ^{Me} dC ^K	phosphate	propargyl MedC	7809	7810
Z1	r ^z UAGAUCGGAAGAGCGGUUCAG	5'-azido U		6839	6840
K2	r ^p UUA AUGCUAAUCGUGAUAGGGGUC ^K	phosphate	propargyl rC	7811	7812
K3	rUUA AUGCUAAUCGUGAUAGGGGUC ^K		Propargyl rC		
O1	r ^o CAGAUCGGAAGAGCGGUUCAG	BCN		7040	7040
Z2	r ^p UUA AUGCUAAUCGUGAUAGGGGUU ^z	phosphate	2'-azido-2'-dU	7799	7800
Z3	rUUA AUGCUAAUCGUGAUAGGGGUC ^z		2'-azido-2'-dC	7718	7719
O2	d ^o CAGATCGGAAGAGCGGTTTCAG	BCN		6746	6746
Z4	r ^z CAGAUCGGAAGAGCGGUUCAG	5'-azido C		6839	6839
S1	dCCGATCTAGACCCCT			4473	4473
S2	dCCGATCTGGACCCCT			4489	4488

1b. Reverse Transcriptase Templates, sequence in green is DNA

Code	Sequence (5' - 3') RNA	Calc.	Found
T1 (K1+Z1)	^p UUA AUGCUAAUCGUGAUAGGGGU ^{dC^{Me}} UAGAUCGGAAGAGCGGUUCAG	14648	14650
T2 (K3+Z1)	UUA AUGCUAAUCGUGAUAGGGGU ^C UAGAUCGGAAGAGCGGUUCAG	14570	14571
T3 (Control)	UUA AUGCUAAUCGUGAUAGGGGU ^{CU} AGAUCGGAAGAGCGGUUCAG	14570	14572
T4 (K2+Z4)	^p UUA AUGCUAAUCGUGAUAGGGGU ^C CAGAUCGGAAGAGCGGUUCAG	14649	14651
T5 (K1+Z4)	^p UUA AUGCUAAUCGUGAUAGGGGU ^{dC^{Me}} CAGAUCGGAAGAGCGGUUCAG	14647	14648
T6 (Control)	UUA AUGCUAAUCGUGAUAGGGGU ^{CC} AGAUCGGAAGAGCGGUUCAG	14569	14570
T7 (Z2+O1)	^p UUA AUGCUAAUCGUGAUAGGGGU ^U CAGAUCGGAAGAGCGGUUCAG	14839	14839
T8 (Z3+O1)	UUA AUGCUAAUCGUGAUAGGGGU ^C CAGAUCGGAAGAGCGGUUCAG	14758	14759
T9 (Z3+O2)	UUA AUGCUAAUCGUGAUAGGGGU ^C CAGATCGGAAGAGCGGTTTCAG	14464	14465

1c. Reverse Transcriptase primers

Code	Sequence (5' - 3')	position	Calc.	Found
P1	Fam-dCTGAACCGCTCTTC	-7	4711	4711
P2	Fam-dCTGAACCGCTCTTCCGAT	-3	5947	5947
P3	Fam-dCTGAACCGCTCTTCCGATCTA	+0	6853	6854
P4	Fam-dCTGAACCGCTCTTCCGATCTAG	+1	7183	7183
P5	Fam-dCTGAACCGCTCTTCCGATCTAGACC	+4	8074	8075
P6	Fam-dCTGAACCGCTCTTCCGATCTG	+0	6869	6870
P7	Fam-dCTGAACCGCTCTTCCGATCTGG	+1	7199	7199
P8	Fam-dCTGAACCGCTCTTCCGATCTGGACC	+4	8090	8090

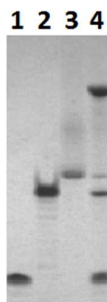
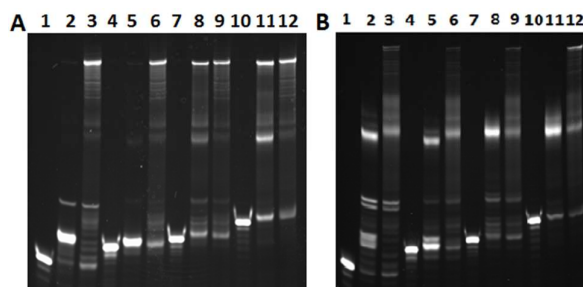


Figure S1: Templated CuAAC click ligation to synthesize T2. Lane 1; S1, lane 2; 5'-azide oligo Z1, lane 3; alkyne oligo K3, lane 4; crude reaction mixture. 20% Polyacrylamide gel.



5'-HO- UUAA UGCU AAUC GUGA UAGG **GGUC**,UAGA UCGG AAGA GCGG UUCA G -3'
CCAG ATCT AGCC TTCT CGCC AAGT C-F-5' P₅
G ATCT AGCC TTCT CGCC AAGT C-F-5' P₄
ATCT AGCC TTCT CGCC AAGT C-F-5' P₃
T AGCC TTCT CGCC AAGT C-F-5' P₂

Figure S2: Mn²⁺ ions enhance the reverse transcriptase read through of the RNA-RNA triazole linkage. Reverse-transcription of triazole template T2 (C₁U) and control T3 by M-MuLV reverse-transcriptase (RNase H⁻) in 3mM Mg²⁺ buffer (A) and 3mM Mn²⁺ buffer (B). Lane 1; primer P2 (-3), lane 2 and 3; reverse transcription product of T2 and T3 using primer P2, lane 4; primer P3 (+0), lane 5 and 6; reverse transcription product of T2 and T3 using primer P3, lane 7; primer P4 (+1), lane 8 and 9; reverse transcription product of T2 and T3 using primer P4, lane 10; primer P5 (+4), lane 11 and 12; reverse transcription product of T2 and T3 using primer P5. The transcription products run as two bands (single stranded and double stranded with the RNA template). This problem was overcome in the later gels by adding a 10-fold excess of the full length unlabelled DNA strand complementary to the RNA template to the reverse-transcription product. P1 (-7) was also used and gave similar results. In some experiments, Mn²⁺ was added to Mg²⁺ containing buffer and results were similar to those which formed from using Mn²⁺ in Mg²⁺ free buffer. This indicates that Mn²⁺ enhances the reverse transcription reaction in the presence or absence of the Mg²⁺ ions.

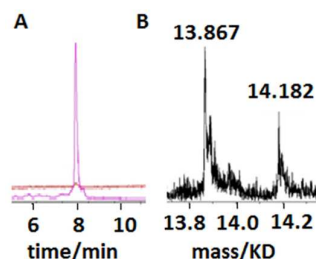


Figure S3: A) HPLC and B) mass spectrum of the M-MuLV reverse transcriptase (RNase H⁻) products of the triazole template T2: (M minus "G", cal. 13867), (M minus "G" plus "A", cal. 14180). The "plus A" fragment likely resulted from the terminal transferase activity of the reverse-transcriptase. (6 mM Mn²⁺ and primer P1 (-7) were used)

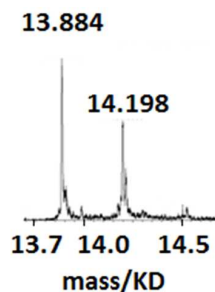


Figure S4: Mass spectrum of reverse transcriptase product of T5: M-MuLV reverse transcriptase with 3 mM Mg^{2+} and primer P1 (-7), incubated at 37 °C for 2 h. Calc. Mass: 13883 (M - G); 14196 (M-G+A).

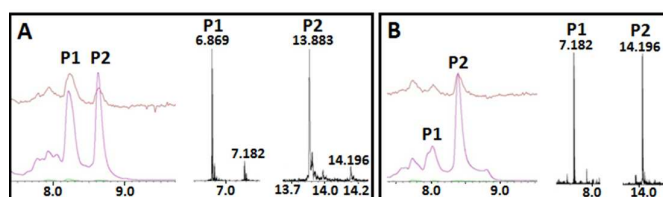


Figure S5: Reverse transcriptase reads through the triazole linkage after overnight incubation. HPLC traces and mass spectra of the reverse transcription product of T8, A) after incubation at 37 °C for 2 h, P1: shows the product formed from RT termination before triazole, Calc. 6869 and RT termination before triazole “+A”, Calc. 7183; P2: “M-G” Calc. 13883 and “M-G+A” Calc. 14196. B) After incubation at 37 °C overnight, P1: shows the product formed from RT termination before triazole “+A” Calc. 7183, P2: “M-G+A” Calc. 14196. M-MuLV reverse transcriptase (RNase H⁻), 3 mM Mg^{2+} buffer and primer P1 (-7) were used. After the incubation time, digestion of the RNA template by RNase H and desalting by gel-filtration (NAP-10) was carried out.

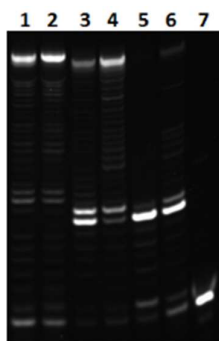


Figure S6: Mn^{2+} ions enhance the reverse transcriptase read through of the RNA-DNA triazole linkage. Lane 1 and 2; reverse transcription of control T6 after 2h and 16h using 3 mM Mn^{2+} , lane 3 and 4; reverse transcription of triazole template T9 after 2h and 16h using 3 mM Mn^{2+} , lane 5 and 6; reverse transcription of triazole template T9 after 2h and 16h using 3 mM Mg^{2+} , lane 7; primer P1 (-7). 20% polyacrylamide gel

DNA sequencing analysis

The reverse transcription products were obtained using unlabelled RT primer (5'-CTGAACCGCTCTTC, same sequence as P1 with no 5'-FAM), templates (T8, C_xC), (T6, control CC) or (T7, U_xC). Mg²⁺ buffer was used in the case of templates T8 and T6 while Mg²⁺ + Mn²⁺ buffer was used in case of template T7.

3 μM primer (300 pmol), 3 μM template (300 pmol), 1x supplied M-MuLV Reverse Transcriptase (RNase H⁻) Reaction Buffer (50 mM Tris-HCl, 75 mM KCl, 3 mM MgCl₂, or 3 mM MnCl₂, pH 8.3 at room temperature), 10 mM DTT, 0.5 mM dNTP (each triphosphate), 1000 U (5 μL) M-MuLV Reverse Transcriptase (RNase H⁻)(NEB[®]) in a total volume of 100 μL. Reverse transcription was carried out at 37 °C for 18h.

cDNAs formed from the reverse transcription of T7 (U_xC), T8 (C_xC) and T6 (CC control) were purified on a 20% polyacrylamide gel (PAGE). Three PCR reactions (50 μL) were carried out using GoTaq DNA polymerase, the PAGE purified DNA and two tailed primers as shown below. The PCR products were purified using a 2% agarose gel followed by extraction using QIAquick Gel Extraction kit (50) Cat. No.28704.

5' -UUAAGCUAAUCGUGAUAGGGGUCx CAGAUCGGAAGAGCGGUUCAG-3' C_xC template
3' -AATTACGATTAGCACTATCCCCAG-GTCTAGCCTTCTCGCCAAGTC-5' RT product

Primers
5' GCATTGAGCAACGTAAGTTAATGCTAATCGTGA GCCTTCTCGCCAAGTCGACGGCTGTGTATTGG-5'

PCR reaction mixture: 25 pmol each primer (final 0.5 μM), 0.002 pmol template, 1x Green GoTaq[®] Reaction Buffer, 2.5 μL 10 mM dNTP (Promega[®], final 0.5 mM), 6.25 u GoTaq[®] DNA polymerase (5u/μL) in total 50 μL.

PCR condition: 95 °C for 2 min, then 30 cycles of 95 °C for 30 sec, 45 °C for 30 sec, 72 °C for 30 sec.

Cloning and automated Sanger sequencing of these PCR products were performed and the results are shown below. M-MLV RT reads through the unnatural BCN-triazole linkages in RNA with omission of one nucleotide (the first base after the triazole). There are few mutations in other regions of the templates. However similar mutations also appeared in the control sequences suggesting that the mutations far from the triazole site may have occurred during the sequencing and cloning process.

T6 CC control, red GG in cDNA strand

Sequence 1 GGTATGTGTGTCGGCAGCTGAACCGCTCTCCGATCTGGACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 2 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTGGACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 3 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTGGACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 4 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTGGACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 5 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTGGACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 6 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTGGACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 7 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCG-CTGGACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 8 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTGGACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 9 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTGGACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 10 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTGGACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 11 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTGGACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 12 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTGGACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 13 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTGGACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 14 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTGGACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 15 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTGGACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 16 GGTATGTGTGTCGGCAGCTGAACCGC-CTTCCGATCTGGACCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 17 GGTATGTGTGTCGGCAGCTGAACCGCTCTTTCGATCTGGACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 18 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTGGACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 19 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTGGACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Required GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTGGACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC

T8 C_xC in RT template should be copied as GG in the cDNA strand

Sequence 1 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 2 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 3 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-GCCCTATCAGATTAGC-TTAACTTACGTTGCTCGAATGC
Sequence 4 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 5 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 6 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 7 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 8 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 9 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-A-CCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 10 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-GCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 11 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 12 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 13 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 14 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 15 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 16 GGTATGTGTGTC- GACGTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 17 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Required GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTGGACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC

T7 U_xC in RT template should be copied as GA in the cDNA strand

Sequence 1 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 2 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCC-TATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 3 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCC-TATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 4 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 5 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCC-TATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 6 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCC-TATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 7 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCC-TATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 8 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCC-TATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 9 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 10 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 11 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 12 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 13 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCC-TATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 14 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCC-TATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 15 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-AC-TATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 16 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 17 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 18 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 19 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Sequence 20 GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTG-ACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC
Required GGTATGTGTGTCGGCAGCTGAACCGCTCTCTCCGATCTGACCCCTATCAGATTAGCATTAACTTACGTTGCTCGAATGC

Figure S7: Sequencing results of clones of the PCR products of the reverse transcription products of templates T6 (CC control), T8 (C_xC) and (T7 U_xC): From the template sequence (T8) GGGGU_xCAGA, the sequence TCTGGACCCC should be produced but instead of GGA, GA is incorrectly produced. From the template sequence (T7) GGGGU_xCAGA, the sequence TCTGAACCCC should be produced but instead of GAA, GA is incorrectly produced. Therefore we can conclude that the purine base in red is the one that is omitted because the pyrimidine base at the 5'-side of the triazole linkage in the RT template is the one that is not copied (i.e. the one in blue).

Reference

1- K. Fauster, M. Hartl, T. Santner, M. Aigner, C. Kreutz, K. Bister, E. Ennifar and R. Micura, ACS Chem. Biol. 2012, 7, 581-589
2- M. Aigner, M. Hartl, K. Fauster, J. Steger, K. Bister and R. Micura, ChemBioChem, 2011, 12, 47-51.