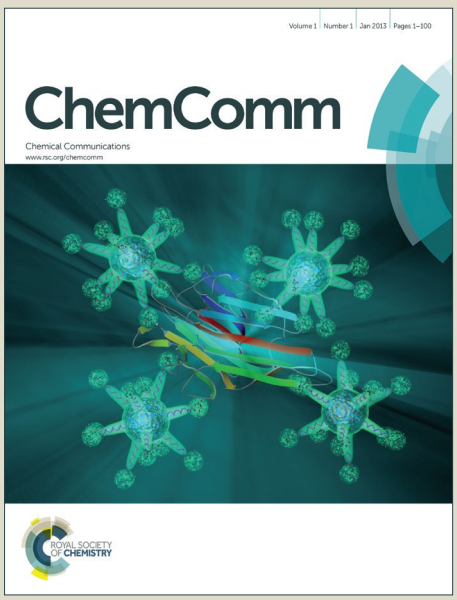
ChemComm

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

Journal Name

COMMUNICATION

Fast Click-Slow Release Strategy Towards HPLC-Free Synthesis of RNA.

E. Agustin,[‡] P. N. Asare Okai,[‡] Irfan,Khan, M. R. Miller, R. Wang, J. Sheng and M. Royzen[a]

Received 00th January 20xx,
Accepted 00th January 20xx

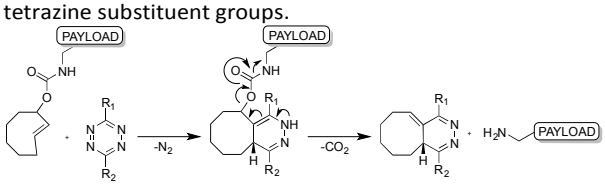
DOI: 10.1039/x0xx00000x

www.rsc.org/

A general strategy for purification of oligonucelotides synthesized by solid phase synthesis is described. It is based on a recently developed concept involving bio-orthogonal inverse electron demand Diels-Alder reaction between *trans*-cyclooctene and tetrazine, termed 'click-to-release'. The strategy has been applied towards synthesis and purification of a model hairpin RNA strand, as well as a 34-nt long aptamer.

The bio-orthogonal inverse electron demand Diels-Alder (IEDDA) reaction between *trans*-cyclooctene (TCO) and tetrazine has found a wide range of applications in imaging spanning the fields of chemical biology, molecular diagnostics and medicine.[1] The pre-targeting approach utilizing IEDDA has evolved into a powerful method for selective delivery of pharmaceutical agents to cancerous cells and tumors.[2] However, there are few examples to the use of this chemistry in construction of complex molecular structures.[3] This paucity of applications is in drastic contrast with the now classical 'click chemistry', which has been consistently used for crosslinking DNA, functionalization of block polymers and peptide ligation, to name a few. We envisioned that the IEDDA reaction may also be a useful synthetic tool as this process is orthogonal to a wide range of organic functional group reactions.

In this work, a general strategy was developed to facilitate solid phase synthesis of oligonucleotides using the IEDDA reaction. In particular, one of the latest developments in IEDDA chemistry, termed 'click-to-release',[4] was utilized (Scheme 1). The unique feature of this system is attachment of a releasable payload at the allylic position of the TCO. After the initial cycloaddition step, the payload is spontaneously released as shown. The kinetics of the click and release steps are very different and are optimized through appropriate tetrazine substituent groups.

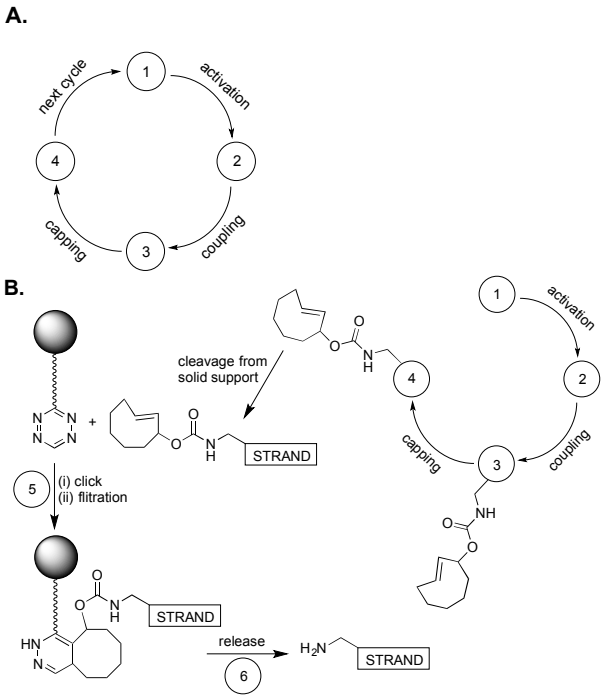


Scheme 1. The concept of click-to-release systems employing a bio-orthogonal IEDDA reaction.

A typical solid-phase oligonucleotide synthesis, follows the logic described in Scheme 2A.[5] It involves activation of the solid phase-immobilized first subunit, typically achieved by cleavage of a protecting group, followed by coupling of the second subunit. Failure strands, generated due to incomplete coupling, are capped to prevent further propagation. After oxidation of the phosphite triester, the dimer strand is carried over to the next cycle. One of the most challenging aspects of this approach is purification of the final product from the failure sequences that occur to some extent during every coupling step. In a typical synthesis of 20-nt long RNA, failure sequences may constitute as much as 50-60% of the total oligonucleotide content. Typically, purification is achieved through RP-HPLC, which is time and labor intensive and increasingly problematic with longer strands.

We envisioned modification of the final cycle of the solid-phase synthesis, as shown in Scheme 2B. The final coupling step will be carried out using a unique subunit functionalized with a releasable TCO group. Because all prior generated failure sequences are capped, only the desired strand will incorporate the releasable TCO. Upon completion of the final cycle, the synthetic oligo will be cleaved from the original solid support and treated with agarose-immobilized tetrazine (step 5). The full strand will 'click' to the tetrazine-modified agarose, while all the species, that didn't 'click' including the failure sequences, will be filtered out. The target strand will be isolated upon the spontaneous 'release' step (step 6).

[a] *Address here.*
[b] *Address here.*
[c] *Address here.*
† Footnotes relating to the title and/or authors should appear here.
Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

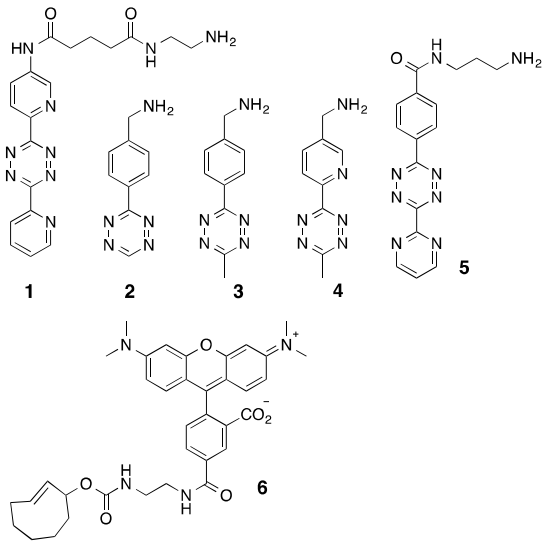


Scheme 2. (A) The logic of a typical solid-phase oligonucleotide synthesis. (B) Modification of the final cycle using 'click-release'.

As a proof-of-principle we have applied the idea described in Scheme 2B towards solid phase synthesis of a hairpin RNA strand, SL3, which is a known substrate for the HIV nucleocapsid protein, NCp7.[6] There are other known strategies that facilitate purification of synthetic RNA strands, including biotinylation, polymerization and perfluorination.[7] The method described herein is based on IEDDA chemistry which allows for robust purification of RNA under mild conditions due to orthogonality of TCO and tetrazine to the functional groups found in RNA.

For the 'click-to-release' system to be successful in purification of synthetic RNA strands, the kinetics of the 'click' and 'release' steps have to be significantly different to enable filtration of the unbound species. The kinetics of these steps will be sensitive to variation of the tetrazine's substituent groups. The kinetics of the 'click' step will also be dependent on the concentration of the TCO-bound payload, while the intramolecular-triggered 'release' will not. To get an estimate of the timescale of the payload release, we screened a library of five known tetrazines shown in Scheme 3. Tetrazines **2** and **5** are commercially available, while **1**, **3**, and **4** were synthesized using reported procedures.[8] The tetrazines were coupled to NHS-derivatized agarose beads and the kinetics of the payload release was studied using rhodamine-derivatized TCO, **6**.

Scheme 4 illustrates the experimental design of the kinetic studies. Agarose beads modified with 3.6 μmol of **1**-**5** were placed into spin columns (33 mg per column). The columns were treated with three different concentrations of **6** (5, 1, 0.5 mM), dissolved in the minimal amount of water (300 μL) to form a homogeneous suspension of the solid support. After 2 min, the supernatant was collected after quick centrifugation and the agarose was resuspended in 300 μL of water. The supernatants were subsequently collected at regular time intervals and analyzed by fluorescence measurements (Ex. 560 nm; Em. 580 nm).



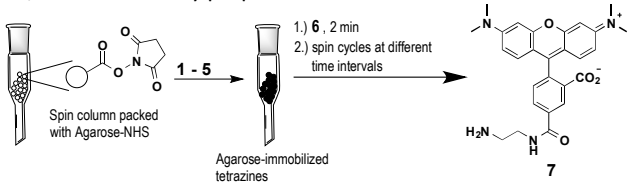
Scheme 3. 1,2,4,5-tetrazines and rhodamine-labelled TCO utilized to study the kinetics of 'click-to-release'.

Figure 1 shows the observed fluorescence of the supernatant fractions, as a function of time and concentration of the TCO-bound payload, **6**. The supernatant fractions of **1**a-**5**a, collected after 2 min, contained trace amounts of **6**, indicating that virtually all of the substrate 'clicked' to the solid support. The observed release product, **7**, in these fractions ranged from 17-32%. The efficiency of the RNA purification process described in Scheme 2b is dependent on the differential kinetics of the 'click' and release steps to allow filtration of the solid support unbound species. Agarose-immobilized with **1** released the lowest amount of **7** after 2 min, making it the best candidate out of the tested tetrazines. The loss of material due to premature release is significantly lower than the loss of RNA that typically occurs during HPLC purification. The agarose samples treated with lower concentrations of the substrate resulted lower conversions of the 'click' step. The supernatant fractions of **1**b-**5**b, as well as **1**c-**5**c, collected after 2 min, contained **6**, as determined by HPLC. This is consistent with the fact that the 'click' step follows the second order kinetics, resulting in a slower rate of IEDDA reaction at the lower concentrations of **6**.

The observed kinetic trends are in agreement with the previous report describing ability of electron withdrawing groups at the 3- and 6-positions of tetrazine to increase the rate of the 'click' step and decrease the rate of the payload release.[4] The bis-pyridyl tetrazine, **1**, is the most electron deficient diene in the examined series. The electron withdrawing groups lower the LUMO energy level of the diene, thus minimizing the HOMO-LUMO gap and the activation

energy of the IEDDA reaction.[9] The mechanism of the release step is not fully understood. However, the kinetic study by Robillard and co-workers also describes slow payload release by the electron deficient tetrazines.[4]

Subsequently, uridine analog containing releasable TCO was prepared and converted to a phosphoramidite suitable for solid phase synthesis (Scheme 5). The synthesis commenced with 5-iodouridine, which was coupled to TFA-protected propargylamine using Sonogashira conditions. After hydrolysis of the TFA group, the releasable TCO group was installed. The 5'-OH group was protected with a DMT group, followed by protection of the 2'-OH group with a TBDMS group. The phosphoramidite suitable for solid phase synthesis was installed in the final step. Analytically pure phosphoramidite, **12**, was obtained by preparative TLC.



Scheme 4. Schematic representation of kinetic study of 'click-to-release'.

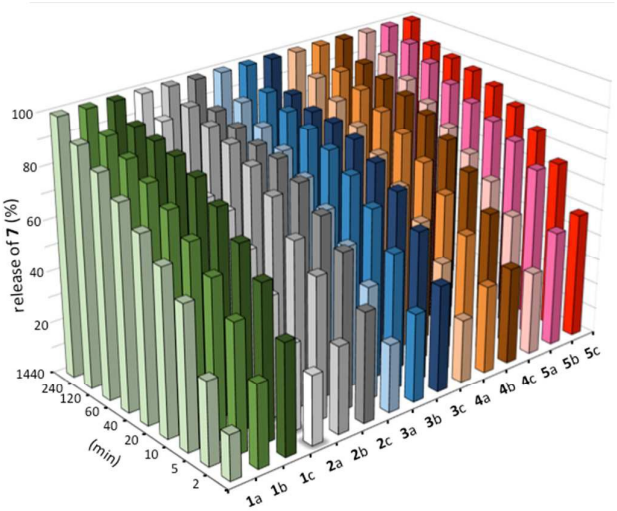
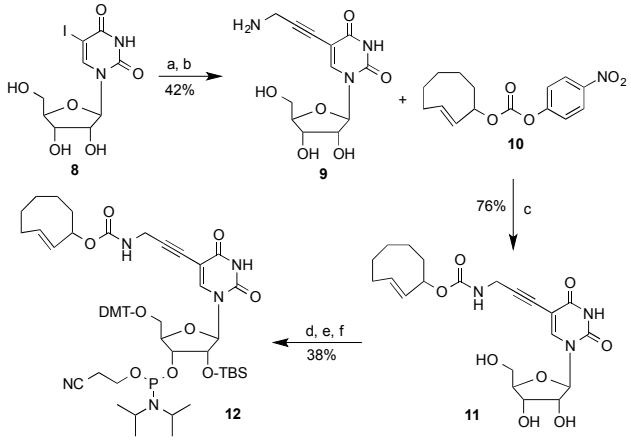


Fig. 1. Fluorescence measurements of the supernatants fractions obtained upon treatment of agarose beads modified with the tetrazines **1**-**5**, with three different concentrations of **6**: 5 mM (a), 1 mM (b), 0.5 mM (c). The numbering of the X-axis corresponds to the immobilized tetrazines, while the letters correspond to the concentrations of **6**.

The SL3 strand, having the sequence: 5'-U*GGACUAGCGGAGGCUAGUCC-3', was synthesized using standard solid phase synthesis techniques, described in detail in the supporting information (U* represents the TCO-modified uridine). Following the strategy described in Scheme 2B, compound **12** was the last phosphoramidite coupled to the RNA strand during the last cycle of the solid phase synthesis. Subsequently, the RNA strand was detached from the solid support, desilylated, and treated with agarose modified with **1** in 300 μL of water for 2 min to allow the 'click' step to occur. The agarose beads were washed three times with water to remove all unbound species (a process that took 1 min), and then suspended in water for 5 h at rt, to allow the 'release' step to occur. The results of 'click-to-release' were analyzed by PAGE, shown in Figure 2A. Lane 1, termed 'unretained' material, represents the supernatant collected after the first 2 min. It contains the synthetic failure sequences as well as the SL3 strand that failed to 'click' to the solid support. Lanes 2-4 correspond to the three washes that contain traces of the released RNA strand. Lane 5 is the released product that was collected after 5 hours of incubation. The supernatant solutions collected at every stage of this process were analyzed by ESI-MS to establish their composition (Figure S1).



Scheme 5. Synthesis of the TCO-modified uridine phosphoramidite: (a) TFA-propargylamine, $Pd(PPh_3)_4$, CuI, Et_3N, DMF; (b) NH_4OH, H_2O; (c) DMF; (d) DMT-Cl, pyridine, CH_2Cl_2; (e) TBDMS-Cl, pyridine, $AgNO_3$, THF; (f) CEP-Cl, Et_3N, CH_2Cl_2.

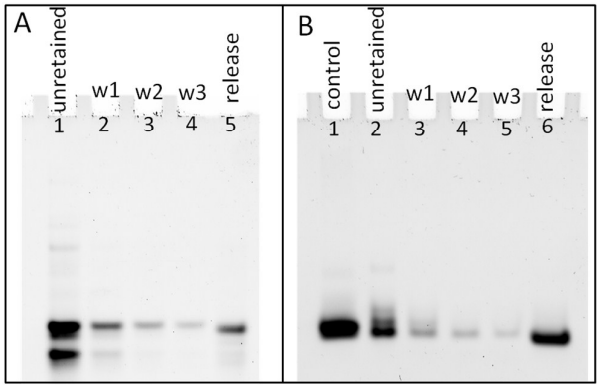


Fig. 2. PAGE analysis of the 'click-to-release' process. Panel **A** shows the supernatant fractions collected at different timepoints after the crude product of the solid phase synthesis of SL3 was treated with agarose modified with **1**. Lane 1 is the supernatant collected after 2 min; lanes 2-4 are the three wash steps and lane 5 is the supernatant collected after the 5 h incubation. Panel **B** represents the supernatant fractions collected at different timepoints after the HPLC purified SL3 was treated with agarose modified with **1**. Lane 1 is HPLC purified SL3. Lane 2 is the supernatant collected after 2 min; lanes 3-5 are the three wash steps and lane 6 is the supernatant collected after the 5 h incubation. Based on the pixel intensity, the efficiency of the click-to-release step was calculated to be 65 ± 5 %.

To quantify the yield of the 'click-to-release' step, we purified the synthetic SL3 strand using RP-HPLC and subjected it to the same conditions as above. Figure 2B illustrates the PAGE of the supernatant fractions obtained at the key steps of the protocol. Lane 1 represents the HPLC purified SL3 RNA strand (control). Lane 2 is the supernatant containing unretained materials after the first 2 min, while lanes 3-5 are the solutions obtained after the three wash steps. No release product was detected by ESI-MS in the supernatant loaded in lane 2. Lane 6 is the release product obtained after the 5 hour incubation step. By comparing the pixel intensity of the control lane (Figure 2B, lane 1) to lanes 2-6 the efficiency of release was calculated to be 70 % with 5 % of the sample lost in the wash steps, giving an overall recovery of 65% ± 5 %.

To compare our method with chromatographic purification, we divided the crude SL3 RNA obtained from the solid phase synthesis (1 μmol scale), containing trityl group at the 5'-end, into four equal fractions. One of the fractions was purified by RP-HPLC and desalted using Amicon Ultra Centrifugal Filter[10] (0.5 mL), yielding 61 nmol of SL3, as determined by nano-drop. Another fraction was purified using the 'click-to-release' procedure and similarly desalted, yielding 98 nmol of SL3.

To illustrate the generality of the 'click-to-release' approach, we synthesized a theophylline-binding RNA aptamer having the sequence: 5'-U*GGCGAUACCAGCCGAAAGGCCCUUGGCAGCGUC-3', (U* represents the TCO-modified uridine).[11] We chose to focus on this particular example, because it proved to be synthetically challenging in the hands of our collaborators. The solid phase synthesis of the 34-nt long aptamer RNA resulted in higher amounts of failure strands than typically observed during solid phase syntheses of RNAs of this length. Lower coupling yields at particular stages of the solid phase synthesis were attributed to structural elements of the aptamer's sequence. We thus chose this challenging example to illustrate the ability of our bio-orthogonal chemistry-based approach to isolate the 34-nt long aptamer RNA from a solution containing a series of failure strands.

Upon completion of the solid phase synthesis (1 μmol scale), the crude aptamer RNA was split into four equal fractions. One of the fractions was purified by the 'click-to-release' procedure. Every step of the process was analysed by PAGE, shown in Figure S2 and ESI-MS, shown in Figure S3. The supernatant fraction collected after 2 min (Figure S3, panel **A**) contains a series of failure strands, as well as a small amount of the target RNA that did not click to the agarose resin. Numerous other mass peaks were observed on the mass spectra due to fragmentation. Subsequent wash steps further removed failure strands (Figure S3, panels **A**-**D**). After 5 h of incubation, we were able to isolate 32 nmol of the target aptamer RNA (Figure S3, panel **E**) upon its release from the agarose resin. The efficiency of 'click-to-release' procedure was once again compared to RP-HPLC purification using the second fraction of the crude RNA. The conventional purification method yielded 23 nmol of the target RNA.

In summary, a general strategy for purification of oligonucleotide strands synthesized by solid phase synthesis using bio-orthogonal IEDDA reaction was described. This strategy entailed modification of the subunit incorporated in the final cycle of the solid phase synthesis with a releasable TCO group. The full strand can be isolated by 'clicking' to tetrazine-functionalized agarose and subsequent spontaneous release. This general strategy has been implemented towards synthesis of two model RNA strands. The isolated RNA strands contain a small modification at the 5-position of the nucleobase of the final uridine residue. Although it might be useful in some cases (one could couple a fluorescent group or a biotin tag), we are currently improving the design of the TCO-modified phosphoramidite to achieve 'traceless' release of RNA containing only canonical nucleobases.

Notes and references

[a] *University at Albany, SUNY, Department of Chemistry 1400 Washington Avenue, Albany, NY 1222, United States Email: mroyzen@albany.edu Tel: 917-974-4391.*

‡ E. Agustin and P.N. Asare-Okai contributed equally to this work.

‡‡ Electronic Supplementary Information (ESI) available: Experimental procedures, synthesis of small molecules and RNA, ESI-MS analyses.

1 (a) N. K. Devaraj, S. Hilderbrand, R. Upadhyay, R. Mazitschek, R. Weissleder, *Angew. Chem. Int. Ed.* 2010, **49**, 2869-2872. (b) C. M. Cole, J. Yang, J. Seckute, N. K. Devaraj, *Chembiochem*, 2013, **14**, 205-208.

2 R. Rossin, P. R. Verkerk, S. M. van den Bosch, R. C. Vulders, I. Verel, J. Lub, M. S. Robillard, *Angew. Chem. Int. Ed.* 2010, **49**, 3375-3378.

3 (a) A. M. Cok, H. Zhou, J. A. Johnson, *Macromol. Symp.* 2013, **329**, 108-112. (b) Z. Li, H. Cai, M. Hassink, M. L. Blackman, R. C. Brown, P. S. Conti, J. M. Fox, J. M. *Chem. Commun.* 2010, **46**, 8043-8045.

4 R. M. Versteegen, R. Rossin, W. ten Hoeve, H. M. Janssen, M. S. Robillard, *Angew. Chem. Int. Ed.* 2013, **52**, 14112-14116.

5 S. L. Beaucage, C. B. Reese, *Curr. Protoc. Nucleic Acid Chem.* 2009, 2.16.1-2.16.31.

6 R. N. de Guzman, Z. R. Wu, C. C. Stalling, L. Pappalardo, P. N. Borer, M. F. Summers, *Science*, 1998, **279**, 384-388.

7 (a) S. Fang, D. E. Bergstrom, *Tetrahedron Lett.* 2004, **45**, 7987-7990. (b) C. Beller, W. Bannwarth, *Helv. Chim. Acta*, 2005, **88**, 171-179. (c) S. Fang, S. Fueangfung, X. Lin, X. Zhang, W. Mai, L. Bi, S. A. Green, *Chem. Commun.* 2011, **47**, 1345-1347. (d) S. Fang, S. Fueangfung, *Org. Lett.* 2010, **12**, 3720-3723.

8 Compound **2** was purchased from Click Chemistry Tools (cat.# 1130-100); compound **5** was purchased from SiChem (cat.# SC-1192); compounds **3** and **4** were prepared following: Yang, M. R. Karver, W. Li, S. Sahu, N. K. Devaraj, *Angew. Chem. Int. Ed.* 2012, **51**, 5222-5225. Compound **1** was synthesized following: R. Selvaraj, S. Liu, M. Hassink, C. Huang, L. Yap, R. Park, J. M. Fox, Z. Li, P. S. Conti, *Bioorg. Med. Chem. Lett*. 2011, **21**, 5011-5014.

9 D. L. Boger, M. S. Subas, *J. Org. Chem.* 1988, **53**, 1415-1423.

10 Amicon ® Ultra Centrifugal Filters were purchased from Merck Millipore Ltd.

11 G. R. Zimmerman, C. L. Wick, T. P. Shields, R. D. Jenison, A. Pardi. *RNA*, 2000, **6**, 659-667.