ChemComm

ROYAL SOCIETY OF CHEMISTRY

COMMUNICATION

View Article Online
View Journal | View Issue

Cite this: *Chem. Commun.*, 2021, **57**, 4263

Received 27th January 2021,
Accepted 19th March 2021

DOI: 10.1039/d1cc00096a

rsc.li/chemcomm

Bio-orthogonal chemistry enables solid phase synthesis and HPLC and gel-free purification of long RNA oligonucleotides†

Muhan He, Xunshen Wu, Song Mao, Phensinee Haruehanroengra, Irfan Khan, Jia Sheng and Maksim Royzen *

Solid phase synthesis of RNA oligonucleotides which are over 100-nt in length remains challenging due to the complexity of purification of the target strand from failure sequences. This work describes a non-chromatographic strategy that will enable routine solid phase synthesis of long RNA strands.

Synthetic RNA occupies a very special place in modern research. Custom solid phase synthesis of 20–30 nucleotide-long strands has become a powerful driving force in many fields of biochemical and pharmaceutical research focusing on these relatively short RNAs. Discoveries of the twenty-first century created a strong need for a robust solid phase synthesis of longer RNA strands, over 100 nucleotides (-nt) in length. For example, 101-nt long single-guide RNA (sgRNA) is required for CRISPR, one of the most effective approaches to gene editing.[1,2] Solid phase synthesis of sgRNA allows sequence specific incorporation of RNA modifications that can improve the CRISPR efficiency and nuclease stability and reduce off-target activity.[3,4]

Despite the strong need, solid phase synthesis of RNAs that are 100-nt in length remains challenging and is rarely attempted.[5,6] The limiting step of the otherwise highly optimized process is purification, illustrated in Fig. 1A.[7,8] The standard purification process entails cleavage of oligonucleotides from the solid support and concomitant deprotection of the nucleobases (step 1). Subsequently, desilylation of the 2-hydroxy groups is done using a fluoride reagent (step 2). After ethanol precipitation, the target RNA strands are purified using reverse phase HPLC or preparative gel electrophoresis (step 3). The latter is the most labor intensive, time consuming and challenging step.

A number of innovative approaches that allow HPLC-free purification of synthetic oligonucleotides have emerged. The 'DMT-on' approach has been extensively applied towards purification of oligonucleotides used in miRNA and RNAi research.[9] The hydrophobic DMT group serves as a handle for separation of DMT-protected (DMT-on) full-length oligonucleotides from the failure sequences. However, this approach is typically utilized for purification of RNA strands that are less than 50-nt long and the DMT group is usually not very stable before HPLC purification. More recently, Fang described an approach for capping the failure sequences with an acrylated phosphoramidite.[10] Subsequent polymerization of the failure sequences allows isolation of the target stand using extraction. The same research group also reported a similar method for catching the target strands by polymerization.[11] Bergstrom reported reversible 5′-end biotinylation of synthetic RNAs.[12] After cleavage and deprotection, the target strands were captured with NeutrAvidin coated microspheres. Beaucage synthesized DNA strands carrying 5′-siloxyl ether linkers that can be captured through an oximation reaction with aminopropylated silica gel.[13] The Minakawa group was the first to describe a 'catch and release' oligonucleotide purification strategy that combined strain promoted alkyne-azide cycloaddition and photocleavage.[14] To our knowledge, none of these approaches reported successful purification of 100-nt long RNA strands that present a higher degree of complexity.

This work describes a non-chromatographic method that facilitates construction of long synthetic RNA strands, schematically illustrated in Fig. 1B. Our strategy is based on bio-orthogonal inverse electron demand Diels–Alder (IEDDA) chemistry[15,16] between *trans*-cyclooctene (TCO) and tetrazine (Tz) that allows selective tagging and purification of structurally complex and increasing long RNA strands from the failure strands that accrue during solid phase synthesis. RNA synthesis is done on a controlled pore glass (CPG) solid support modified with a photolabile linker. After the final synthetic cycle, our approach takes advantage of the free 5′-OH group on the target strand, which provides an opportunity for selective bio-orthogonal tagging. Upon installing Tz on the target strand,

University at Albany, Department of Chemistry, Albany, NY,12222, USA.
E-mail: mroyzen@albany.edu

† Electronic supplementary information (ESI) available. See DOI: 10.1039/d1cc00096a

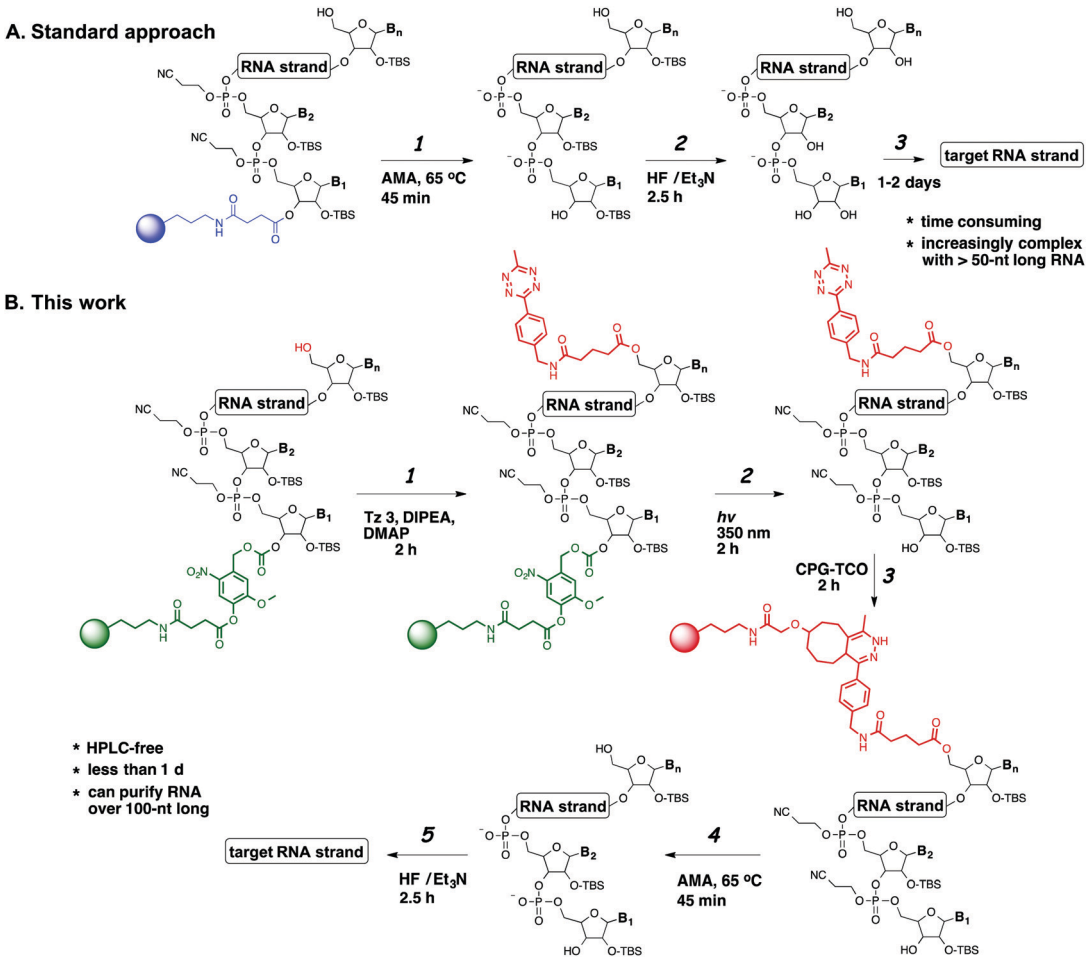
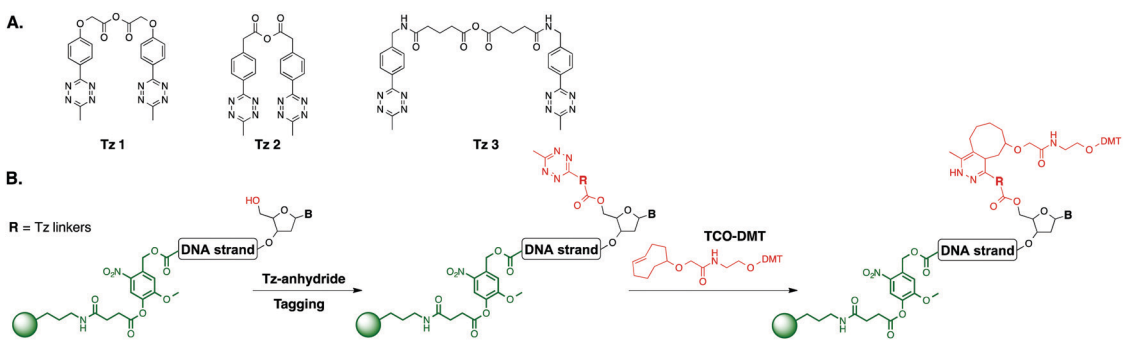


Fig. 1 (A) The standard approach to isolate synthetic RNA. (B) The non-chromatographic approach described in this work: (1) tagging of the target strand with Tz anhydride; (2) photocleavage; (3) capture of the target strand with CPG-TCO; (4) deprotection; and (5) desilylation and ethanol precipitation.

oligonucleotides are cleaved from the solid support using UV irradiation. The target strand can be selectively immobilized using CPG-modified with TCO, while all failure strands dissolved in the supernatant can be removed. Subsequently, the target RNA strand is isolated using standard cleavage, deprotection, desilylation and ethanol precipitation steps. The 5-step process yields pure RNA strands that do not require any further purification.

Implementation of our strategy requires optimization of three key steps: (1) immobilization of RNA on the solid support; (2) tagging of the target RNA strands with Tz; and (3) capture of the target RNA strands. To address the first challenge, we decided to immobilize RNA on CPG using a previously reported photolabile linker. We could not utilize the standard succinate linker, as its cleavage requires ammonia-methylamine (AMA) treatment, which would inevitably also cleave Tz. Photocleavage



Scheme 1 (A) Tz anhydrides that were explored for tagging of 5′-OH of oligonucleotides; and (B) schematic representation of the approach used to evaluate the efficiency of tagging.

provides an orthogonal chemical procedure that preserves the tagging reagent. The photolabile linker was synthesized using the procedure described by Greenberg and co-workers[17] and attached to CPG1000, as described in Scheme S16 (ESI†). The photocleavage efficiency was determined to be 85.2% relative to the standard AMA cleavage (Fig. S5, ESI†).

To address the second challenge, we synthesized a series of Tz anhydrides, shown in Scheme 1A. Optimization of the tagging step was done using a model 20-mer DNA strand (5′-TCATTGCTGCTTAGATTGCT-3′). To quantify tagging, we also synthesized a **TCO-DMT** reagent (Scheme 1B). After the tagging step, the CPG beads were thoroughly washed to remove excess Tz and treated with **TCO-DMT** for 1 hour. The IEDDA reaction installs DMT groups on the immobilized oligonucleotides. After removal of the supernatant, the CPG beads were treated with a detritylation reagent. The absorbance at 504 nm was measured to determine the tagging yield (Fig. S2, ESI†). The attempted conditions are outlined in Table S1 (ESI†).

Optimal tagging of DNA was done in a 94% yield using **Tz 3** (Table S1, row 6, ESI†). These conditions were used for the rest of the studies described herein.

To optimize the capture step, we synthesized three model DNA strands using a photolabile linker-modified CPG solid support. The 20-mer DNA strand (5′-TCATTGCTGCTTAGATTGCT-3′) containing a free 5′-hydroxy group was a model target strand intended for tagging with **Tz 3**. We also synthesized two model failure strands: 17-mer DNA strand (5′-TTGCTGCTTAGATTGCT-3′) and 10-mer DNA strand (5′-TTAGATTGCT-3′), both of which were capped at the 5′-end with acetic anhydride. CPG beads containing the three model DNA strands were mixed and split into two equal portions. The first portion was processed by the standard AMA cleavage and deprotection, resulting in a crude DNA mixture (Fig. 2A, lane 2).

The second portion was utilized for optimization of the capture process. The mixture of CPG beads was treated with **Tz 3** to selectively tag the 20-mer DNA using optimized tagging conditions. Subsequently, the beads were thoroughly washed and all three DNA strands were photocleaved using UV light. The 20-mer DNA was selectively captured by treatment with **CPG-TCO** beads. The optimal capture conditions were determined to be 2 hours at 37 °C. The 17-mer and 10-mer DNA strands remained in the supernatant solution. To analyze the capture process, the **CPG-TCO** beads and the supernatant solution were separated and processed by AMA deprotection. Fig. 2A describes PAGE analysis of the experimental DNA purification process. For gel loading, each DNA sample was resuspended in water and the concentrations were adjusted based on nanodrop measurements. Lane 2 contains a mixture of the three DNA strands (the 10-mer DNA does not stain as well as the larger strands). The successful capture of the 20-mer DNA strand was confirmed by the single dominant band observed in lane 3. Removal of the artificial failure strands in the supernatant fraction was demonstrated in lane 4. We loaded a 2-times higher concentration of DNA in lane 4 to better illustrate the removal of the failure strands (especially the difficult to visualize 10-mer DNA). The band corresponding to the 20-mer DNA strand was also observed in lane 4, indicating a partial loss of the target DNA strand during purification. We believe that this was caused by partial hydrolysis of the Tz group during photocleavage. Evidence of that has been obtained by LC-MS analysis of the photocleavage products, illustrated in Fig. S3 (ESI†). Purification of the 20-mer DNA was further confirmed by the ESI-MS analysis shown in Fig. 2B and C. The ESI-MS analysis also confirmed that pyrimidine–pyrimidine photodimers did not form under the experimental photocleavage conditions. Based on the nanodrop measurements, the isolated yield of the target DNA strand was 168 nmol, which translates to a 33.6% isolated yield.

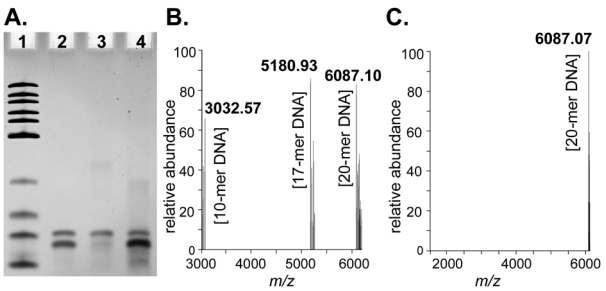


Fig. 2 (A) Urea polyacrylamide gel electrophoresis analysis of capture of target 20-mer DNA. Lane 1: ultra low range DNA ladder, containing DNA strands that are 300, 200, 150, 100, 75, 50, 35, 25, 20, and 10-nt long. Lane 2: mixture of three model DNA strands. Lane 3: captured 20-mer DNA strands. Lane 4: supernatant fraction. (B) Deconvoluted ESI-MS spectrum of the mixture of 20-mer, 17-mer and 10-mer DNA. (C) Deconvoluted ESI-MS spectrum of the purified 20-mer DNA.

The methodology was then applied towards purification of structurally more complex 76-nt long Lys transfer RNA ($tRNA^{Lys}$) containing canonical nucleobases. The synthesis of $tRNA^{Lys}$ was carried out on the photocleavable linker-modified CPG beads. The CPG beads were divided into two equal portions. The crude $tRNA^{Lys}$ from the first portion was purified using preparative PAGE. The isolated $tRNA^{Lys}$ was analyzed by urea PAGE, shown in Fig. 3A (lane 2). The isolated yield was 6.8 nmol or 1.35%. The second portion of the synthesized $tRNA^{Lys}$ was purified using our experimental procedure and analyzed by urea PAGE (Fig. 3A). Purification of $tRNA^{Lys}$ was confirmed in lane 4. Elimination of failure strands was proven by multiple fragment bands observed in lane 5. The target tRNA is also present in lane 4 due to the aforementioned inefficiency of the photocleavage. The sample loss was estimated to be 24.2%. HPLC analysis of the purified $tRNA^{Lys}$, shown in Fig. 3C and D, is also indicative of the removal of failure sequences. Based on the nanodrop measurements, the isolated yield of $tRNA^{Lys}$ was 77 nmol, which translates to an overall isolated yield of 15.4%.

To illustrate the practicality of our method, we synthesized a 101-nt long sgRNA frequently used for CRISPR experiments. Once again, the CPG beads were divided into two equal portions. The first one was purified by preparative PAGE (Fig. 3B, lane 2). The isolated yield was 2.32 nmol or 0.46%. The second portion of synthetic sgRNA was purified using the HPLC-free process described above. The purification was

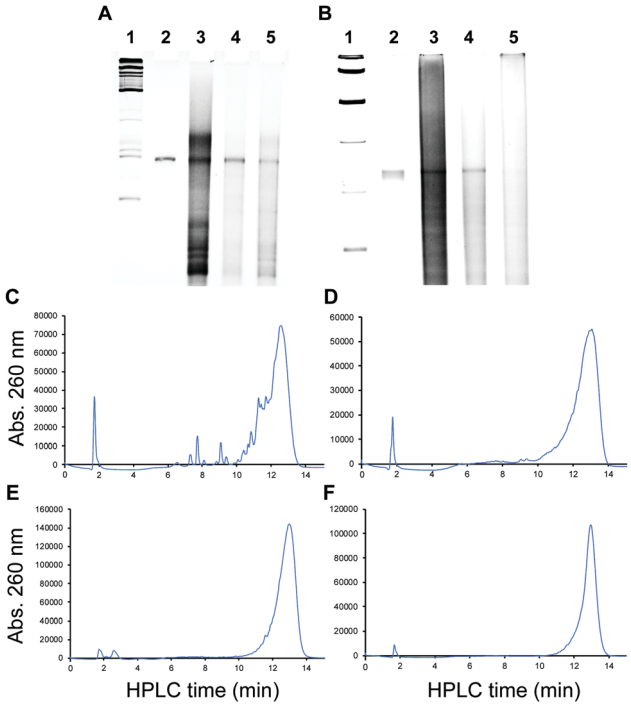


Fig. 3 (A) Urea polyacrylamide gel electrophoresis analysis of purification of $tRNA^{Lys}$. Lane 1: low range ssRNA ladder. Lane 2: $tRNA^{Lys}$ purified by preparative PAGE. Lane 3: crude synthetic $tRNA^{Lys}$. Lane 4: $tRNA^{Lys}$ purified using our experimental procedure. Lane 5: failure strands. (B) Urea polyacrylamide gel electrophoresis analysis of purification of sgRNAs. Lane 1: Low range ssRNA ladder. Lane 2: sgRNA purified by preparative PAGE. Lane 3: Crude synthetic sgRNA. Lane 4: sgRNA purified using our experimental procedure. Lane 4: Failure strands. (C) HPLC analysis of crude synthetic $tRNA^{Lys}$. (D) HPLC analysis of $tRNA^{Lys}$ purified using our experimental procedure. (E) HPLC analysis of crude synthetic sgRNA. (F) HPLC analysis of sgRNA purified using our experimental procedure.

characterized by 15% urea PAGE, shown in Fig. 3B, lane 4. The HPLC analysis of the purified sgRNA is shown in Fig. 3E and F. Based on the nanodrop measurements, the isolated yield of sgRNA was 50 nmol, which translates to an overall isolated yield of 10%. The amount of isolated sgRNA will be sufficient for over 800 CRISPR experiments.

To confirm the functional fidelity of the purified sgRNA, we carried out CRISPR-Cas9 experiments targeting the GFP gene in HEK293T cells, expressing GFP and Cas9.[18] The cells were transfected with the purified sgRNA, which would guide Cas9 to produce double-stranded DNA breaks at the GFP site. As a positive control, the cells were transfected with purchased HPLC-purified crRNA and tracrRNA, which form a two-component structure that has been reported to function analogous to sgRNA.[18] After three-day transfection, the cells were fixed and the GFP expression was evaluated using flow cytometry, shown in Fig. S4 (ESI†). Prior to the transfection, 65% of HEK293T cells were determined to express GFP (purple curve). The percentage of cells expressing GFP decreased to 42% after transfection with the commercial crRNA and tracrRNA mixture (red curve). Transfection of HEK293T cells with sgRNA purified by our experimental procedure resulted in a comparable attenuation of GFP expression, as illustrated in Fig. S4. In the latter case 49% of HEK293T cells were determined to express GPF (green curve). These results indicate that our experimental RNA purification procedure achieved sgRNA with sufficient functional integrity.

In conclusion, this report describes a methodology that allows construction of RNA strands that are over 100-nt in length. The procedure consists of three key steps: photocleavage, tagging and capture, which were optimized using model DNA strands. The optimized protocol was implemented towards purification of 76-nt long tRNA and 101-nt long sgRNA. The isolated sgRNA was sufficient for over 800 CRISPR experiments. In parallel, the target RNAs were processed using the standard procedure (Fig. 1A) and purified by preparative gel electrophoresis. Our experimental method resulted in considerably higher overall yields of isolated RNAs than the standard procedure. The purity of the isolated RNAs was characterized by PAGE and HPLC.

Conflicts of interest

There are no conflicts of interest to declare.

Notes and references

1 M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna and E. Charpentier, *Science*, 2012, **337**, 816–821.
2 F. Zhang, *Hum. Gene Ther.*, 2015, **26**, 409–410.
3 H. Yin, C.-Q. Song, S. Suresh, S.-Y. Kwan, Q. Wu, S. Walsh, J. Ding, R. L. Bogorad, L. J. Zhu, S. A. Wolfe, V. Koteliansky, W. Xue, R. Langer and D. G. Anderson, *Nat. Chem. Biol.*, 2018, **14**, 311–316.
4 D. O'Reilly, Z. J. Kartje, E. A. Ageely, E. Malek-Adamian, M. Habibian, A. Schofield, C. L. Barkau, K. J. Rohilla, L. B. DeRossett, A. T. Weigle, M. J. Damha and K. T. Gagnon, *Nucleic Acids Res.*, 2018, **47**, 546–558.
5 L. Taemaitree, A. Shivalingam, A. H. El-Sagheer and T. Brown, *Nat. Commun.*, 2019, **10**, 1610.
6 A. H. El-Sagheer and T. Brown, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 15329–15334.
7 L. Baronti, H. Karlsson, M. Marusic and K. Petzold, *Anal. Bioanal. Chem.*, 2018, **410**, 3239–3252.
8 A. Grajkowski, J. Cieslak and S. L. Beaucage, *J. Org. Chem.*, 2016, **81**, 6165–6175.
9 R. R. Deshmukh, D. L. Cole and Y. S. Sanghvi, *Methods Enzymol.*, 2000, **313**, 203–226.
10 S. Fang, S. Fueangfung, X. Lin, X. Zhang, W. Mai, L. Bi and S. A. Green, *Chem. Commun.*, 2011, **47**, 1345–1347.
11 S. Fang, S. Fueangfung and Y. Yuan, *Curr. Protoc. Nucleic Acid Chem.*, 2012, **49**, 10.14.11–10.14.21.
12 S. Fang and D. E. Bergstrom, *Tetrahedron Lett.*, 2004, **45**, 7987–7990.
13 A. Grajkowski, J. Cieslak and S. L. Beaucage, *J. Org. Chem.*, 2016, **81**, 6165–6175.
14 Y. Igata, N. Saito-Tarashima, D. Matsumoto, K. Sagara and N. Minakawa, *Bioorg. Med. Chem.*, 2017, **25**, 5962–5967.
15 M. L. Blackman, M. Royzen and J. M. Fox, *J. Am. Chem. Soc.*, 2008, **130**, 13518–13519.
16 N. K. Devaraj, R. Weissleder and S. A. Hilderbrand, *Bioconjugate Chem.*, 2008, **19**, 2297–2299.
17 H. Venkatesan and M. M. Greenberg, *J. Org. Chem.*, 1996, **61**, 525–529.
18 H. Yin, C. Q. Song, S. Suresh, S. Y. Kwan, Q. Wu, S. Walsh, J. Ding, R. L. Bogorad, L. J. Zhu, S. A. Wolfe, V. Koteliansky, W. Xue, R. Langer and D. G. Anderson, *Nat. Chem. Biol.*, 2018, **14**, 311–316.

