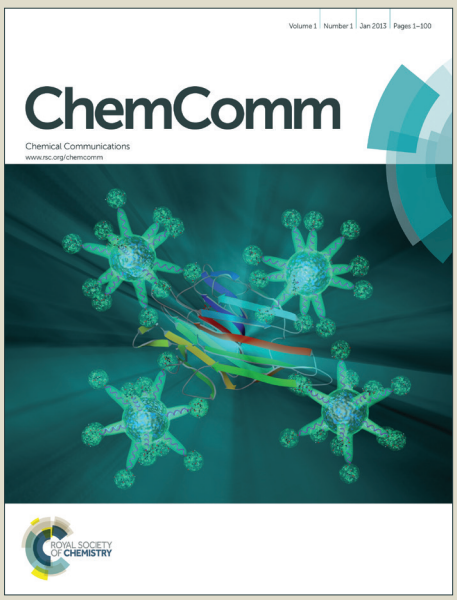
ChemComm

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

Journal Name

RSCPublishing

COMMUNICATION

Cellular activity of siRNA oligonucleotides containing synthetic isomorphic nucleoside surrogates

Cite this: DOI: 10.1039/x0xx00000x

Dongwon Shin,[a] Peter Lönn,[b] Steven F. Dowdy [b] and Yitzhak Tor[a*]

Received 00th January 2012,
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

Singly and multiply modified synthetic siRNA oligonucleotides, containing isomorphic surrogate nucleobases, show high interference potency in cell culture, suggesting the highly isomorphic RNA alphabet, based on a thieno[3,4-*d*]-pyrimidine core, is tolerated well by the cellular silencing machinery.

RNA interference (RNAi), a regulatory process induced by short interfering RNA (siRNA), is a powerful process capable of altering cellular phenotypes,[1-6] deciphering genetic pathways[7-8] and identifying new therapeutic targets.[9-10,11] Short double stranded RNA oligonucleotides are either generated endogenously by Dicer from longer precursors, or delivered exogenously. The passenger strand gets degraded, while the guide strand, loaded onto the RNA-induced silencing complex (RISC), guides the degradation of matched homologous mRNA sequences at a specific site (Figure 1). Related pathways mediated by micro- (miRNAs) or dicer-independent RNAs (piRNAs) have also been described.[12] The exquisite selectivity for all human mRNA targets has made RNAi the focus of intense research. Its promise to facilitate personalized medicine and to keep pace with evolving pandemics has made this technology vital for future therapeutic development. This potential has prompted the exploration of diverse chemical modifications with the hope of altering the activity, selectivity, stability and potency of therapeutically-relevant siRNAs.

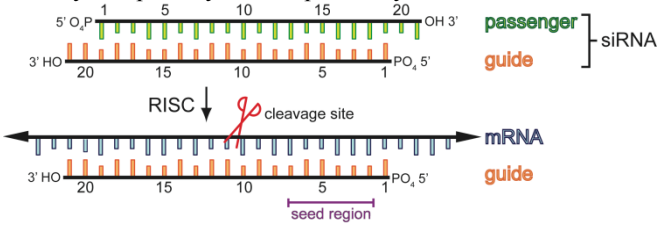


Fig. 1. Basic features of siRNA-mediated gene silencing.

Numerous backbone and ribose modifications have been examined in siRNA.[13-17] Much less has been done, however, with nucleobase surrogates capable of closely mimicking the native heterocycles.[18-20] Such modified isomorphic nucleosides have the potential to impact the cellular interference activity and to serve as built-in probes, potentially assisting in correlating biophysical properties with biological activity. Here we systematically explore the cellular silencing activity and biophysical properties of eighteen siRNA oligonucleotides, modified with a recently introduced isomorphic RNA alphabet (Figure 2).[21] We modify the guide strand at 13 out of the 21 possible nucleotide positions, including the seed region, investigating the effect of individual as well as multiple modifications on the stability of the guide/passenger duplex and on the *in vitro* interference potency. Our results illustrate that this alternative alphabet is well tolerated biophysically and cellularly, with oligonucleotides containing single substitutions displaying the same (or enhanced) activity compared to the native unmodified siRNAs. Multiple modifications were also well-accommodated, even with multiple base substitutions within the critical seed region. Importantly, a qualitative correlation is seen between the thermodynamic stability of the guide/passenger duplex and cellular interference activity.

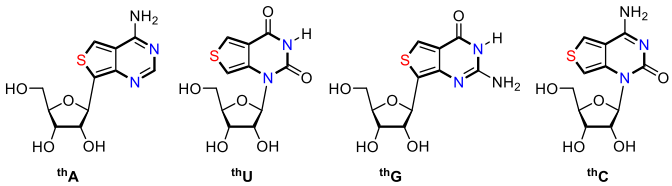
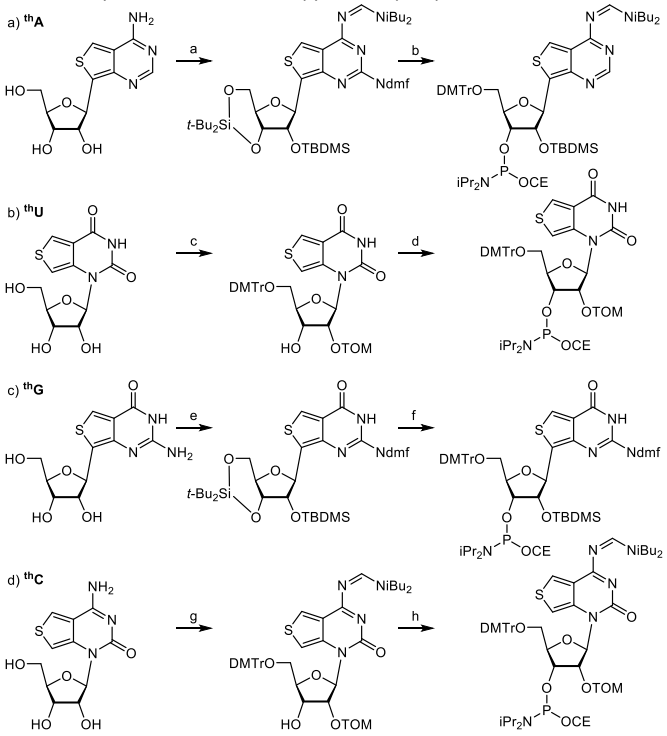


Fig. 2. Isomorphic RNA alphabet.

We have recently reported the preparation of an alternative RNA alphabet, where all nucleosides were derived from thieno[3,4-*d*]pyrimidine as the common heterocyclic nucleus (Figure 2).[21] Structural analysis of their crystal structures suggested high level of "isomophicity". While some perturbations to the D-ribose pucker were seen, all modified ribonucleosides displayed an *anti* orientation at their glycosidic linkage, similar to the preference seen with the native nucleosides.[22] A more critical test, assessing the potential utility of this alternative RNA alphabet, would be to evaluate their impact on the biophysical features and function of biologically relevant oligonucleotides. Although likely to be context, system and sequence dependent, we felt rigorously evaluating their *in vitro* RNA interference activity in cell culture would provide useful insight into their ability to properly replace the native nucleobases in a functionally-demanding context and perhaps inspire future utilizations.

Scheme 1. Synthesis of thieno[3,4-*d*]pyrimidine phosphoramidites.[a]



[a] Reagents and conditions: (a) (i) di-*t*-BuSi(OTf)$_2$, DMF, 0°C; (ii) TBDMSCl, Im, 0 °C–RT, 68 % for 2 steps; (iii) *N*,*N*-diisobutylformamidine dimethyl acetal, DMF, 87 %; (b) (i) HF-Py, CH_2Cl_2, 0 °C, 86 %; (ii) DMTrCl, Py, 69 %; (iii) 2-cyanoethyl *N*,*N*-diisopropylchlorophophoramidite, *i*Pr$_2$NEt, DCM, 0 °C–RT, 95 %; (c) (i) DMTrCl, Py, 76 %; (ii) TOMCl, di-*tert*-butyltin dichloride, *i*Pr$_2$NEt, DCE, 80 °C, 19 %; (d) 2-cyanoethyl *N*,*N*-diisopropylchlorophophoramidite, *i*Pr$_2$NEt, DCM, 0 °C–RT, 94 %; (e) (i) DMF-DMA, DMF, 89 %; (ii) di-*t*-BuSi(OTf)$_2$, DMF, 0 °C–RT; (iii) TBDMSCl, Im, 66 % for 2 steps; (f) (i) HF-Py, CH_2Cl_2, 92 %; (ii) DMTrCl, Py, 0°C, 95 %; (iii) 2-cyanoethyl *N*,*N*-diisopropylchlorophophoramidite, *i*Pr$_2$NEt, DCM, 0 °C–RT, 57 %; (g) (i) *N*,*N*-diisobutylformamidine dimethyl acetal, DMF, 85 %; (ii) DMTrCl, Py, 74 %; (iii) TOMCl, di-*tert*-butyltin dichloride, *i*Pr$_2$NEt, DCE, 80 °C, 41 %; (h) 2-cyanoethyl *N*,*N*-diisopropylchlorophophoramidite, *i*Pr$_2$NEt, DCM, 0 °C–RT, 81 %;

To explore the compatibility of the isosteric RNA alphabet shown in Figure 2 with the cellular interference machinery, an established siRNA sequence, known to target a destabilized green fluorescent protein (dGFP) in human H1299 lung adenocarcinoma reporter cell line,[23] was selected. This assay has proven to be a sensitive and robust platform for evaluating cellular siRNA activity.[7-8] Following the selection of a reporter assay, which identifies a specific siRNA sequence, specific modification positions needed to be selected. Although all positions can, in principle, be modified, we have identified selected positions to challenge the alternative alphabet. Specifically, modifications have been introduced into the seed region as well as opposite the cleavage site.

Eighteen single stranded oligonucleotides, **5**–**22**, containing thieno[3,4-*d*]pyrimidine-based nucleosides replacing their natural counterparts at a single or multiple substitutions in the guide strand, were prepared (Table 1). The new nucleobases were strategically placed to explore the impact of the surrogate nucleobases on positions naturally occupied by both pyrimidines and purines. Additionally, substitutions were made to explore their impact in distinct domains within the 21-mer siRNA guide strand, including residues 2–8 at the 5'-end, known as the seed region, and substitutions neighboring or opposite the RISC induced cleavage site.

The modified 21-mer RNA oligonucleotides were all prepared using solid-phase synthesis (Table 1). The required protected phosphoramidites of th**G**, th**A** , th**U** and th**C** have been synthesized as shown in Scheme 1 relying, in the case of the purine analogs th**G** and th**A** (whose synthesis is more demanding), on the simultaneous protection of the 3' and 5' hydroxyls using di-*tert*-butylsiloxane to facilitate selective TBDMS protection of the 2'-hydroxyl.[24] Following completion of the solid phase synthesis, the oligonuelcotides were deprotected and purified by PAGE. They were further characterized by MALDI-TOF mass spectrometry (Table S1).

Prior to evaluating the cellular activity of the modified siRNAs, their duplex stability was assessed. The effect of base substitutions on the thermal stability of the resulting siRNA duplexes formed upon hybridization to the corresponding passenger strand was first examined and compared to the unmodified siRNA duplex **1•2** (Table 1). Singly modified siRNAs, incorporating the purine mimics th**A** and th**G**, appear to be more stable than their unmodified natural counterpart, where duplex stabilization of about + 0.5 °C per modification is observed at pH 7.0 (20 mM sodium cacodylate, 20 mM NaCl). The single exception was the modification in position G4, which was slightly destabilizing by –0.5 °C. Duplex siRNA oligonucleotides, containing the pyrimidine mimics th**U** and th**C**, were destabilized by –0.4 to –2.9 °C per base modification, when compared to the native duplex. Multiple incorporation of pyrimidine mimics yielded siRNA oligonucleotides, which were destabilized by –1.5 °C and –1.0 °C per modification for U2·6 (**1•16**) and C8·10 (**1•19**), respectively. In contrast, a duplex incorporating two th**A** at positions A11 and A15 (**1•22**) was found to be stabilized by +1.6 °C per modification, yielding a double stranded construct which appears more stable than the corresponding singly modified duplexes with th**A** at positions A11 or A15 (**1•20** and **1•21**, respectively).

siRNA-mediated interference studies were conducted in H1299 cells with a dGFP reporter gene as described.[23] All modified siRNA duplexes containing th**A**, th**C**, th**G** and th**U** substitutions were examined and compared to the activity displayed by the unmodified RNA (WT) and a Luciferase2-specific siRNA (as a negative control).[23]

Although the deviation in thermal stability of all modified siRNAs appears to be relatively subtle when compared to the unmodified siRNA duplex, native polyacrylamide gels

NO.	Duplex		Sequence[a]	T_m (°C)[b]	Δ T_m [c]
1	WT[d]	P	3′-TUG ACC CAC GAG UCC AUC ACC -5′	74.8	0.0
2		G	5′- C UGG GUG CUC AGG UAG UGG UT-3′		
3	Luc2[e]	P	3′-TUA AAC UUA GAA CAU UAG GAC -5′	ND[f]	ND
4		G	5′- U UUG AAU CUU GUA AUC CUG UT-3′		
5	G4	G	5′- C UGG GUG CUC AGG UAG UGG UT-3′	74.3	-0.5
6	G7	G	5′- C UGG GUG CUC AGG UAG UGG UT-3′	75.4	+0.6
7	G13	G	5′- C UGG GUG CUC AGG UAG UGG UT-3′	75.3	+0.5
8	G16	G	5′- C UGG GUG CUC AGG UAG UGG UT-3′	75.3	+0.5
9	G19	G	5′- C UGG GUG CUC AGG UAG UGG UT-3′	75.8	+1.0
10	G4·5	G	5′- C UGG GUG CUC AGG UAG UGG UT-3′	75.8	+1.0
11	5Gs	G	5′- C UGG GUG CUC AGG UAG UGG UT-3′	71.8	-3.0
12	U2	G	5′- C UGG GUG CUC AGG UAG UGG UT-3′	74.4	-0.4
13	U6	G	5′- C UGG GUG CUC AGG UAG UGG UT-3′	73.9	-1.1
14	U9	G	5′- C UGG GUG CUC AGG UAG UGG UT-3′	71.9	-2.9
15	U14	G	5′- C UGG GUG CUC AGG UAG UGG UT-3′	73.0	-1.8
16	U2·6	G	5′- C UGG GUG CUC AGG UAG UGG UT-3′	71.8	-3.0
17	C8	G	5′- C UGG GUG CUC AGG UAG UGG UT-3′	72.4	-2.4
18	C10	G	5′- C UGG GUG CUC AGG UAG UGG UT-3′	73.9	-1.1
19	C8·10	G	5′- C UGG GUG CUC AGG UAG UGG UT-3′	72.9	-1.9
20	A11	G	5′- C UGG GUG CUC AGG UAG UGG UT-3′	75.3	+0.5
21	A15	G	5′- C UGG GUG CUC AGG UAG UGG UT-3′	75.3	+0.5
22	A11·15	G	5′- C UGG GUG CUC AGG UAG UGG UT-3′	77.9	+3.1

Table 1. siRNA Duplexes studied. All siRNA, excluding Luc2 (see below), have a native sugar phosphate backbone. [a] Passenger strands are on top and guide strands below; [b] T_m values were measured at duplex concentrations of 1 μM in 20 mM sodium cacodylate, pH 7.0, 20 mM NaCl, and 0.5 mM EDTA. Errors of at least two independent measurements were less than 0.8 °C.; [c] Difference in T_m compared to **1•2**; [d] wild type; [e] Luciferase2 siRNA served as a negative control and had a modified backbone (2'-hydroxyl groups were fully modified as OMe for purines and replaced with F for pyrimidines); [f] Not determined.

were run before all transfection experiments (see Figure S1). The cells were transfected using lipofectamine with 1–50 nM of dsRNA in the cell media. Interference was evaluated by measuring the level of GFP expression at 24, 48 and 72 hr using flow cytometry. Figure 3 shows the data corresponding to 48 hr and additional time points (for 24 and 72 hours) are provided in the Supplementary Information (see Figure S2). Figure 4 depicts the normalized interference activity with respect to the native siRNA shown along the thermal stability of each duplex. The following sections summarize the interference observed per modification.

Interference results for siRNA duplexes containing guide strand modifications with th**G** replacing the native G in positions 4, 7, 13, 16 and 19 are shown in Figure 3. These modified duplexes were nearly as effective as the WT sequence in all concentrations, with the exception of the G4 substitution (**1•5**), which showed slightly attenuated activity and G13 (**1•7**), which consistently showed enhanced activity when compared to the positive control (Figure 4).

This particular siRNA contains only two A residues at positions 11 and 15 in the guide strand. Their substitution for th**A** yielded little effect on cellular mRNA degradation compared to the unmodified WT for both singly and doubly modified siRNAs (Figure 3). This is of significance, since A11 is near the cleavage site of the target mRNA, demonstrating unperturbed function by a strategically modified siRNA.

Substituting U for th**U** at positions 2, 6, 9 and 14 in the guide strands **12**–**15**, respectively, was somewhat disruptive for RNAi activity (Figure 3). In particular, incorporating th**U** at positions 2 and 9 (**1•12** and **1•14**, respectively), flanking the seed region, were most detrimental (Figure 4) with up to 25% attenuation in interference activity with 5 nM duplex concentration at 48 hours. Interestingly, and in stark contrast to the behavior seen with residues 2 and 9, substituting U6 within the seed region with th**U** (**1•13**) retained gene silencing activity at levels close to the RNAi activity displayed by the WT oligonucleotide. Incorporating th**U** at position 14, toward the 3'-end of the guide strand (siRNA **1•15**) did not disrupt RNAi activity, with potency similar to that displayed by the siRNAs containing the G mimic.

Since the siRNA used in this study has only three C residues in the guide strand, with one of them at the 5'-end, we examined the replacement of the internal residues only. Remarkably, the siRNA oligonucleotide (**17**) where C at position 8 within the seed region was substituted by th**C**, exhibited better gene knockdown than that displayed by the WT. Incorporating th**C** at position 10 had minimal impact on silencing activity, with potency similar to that of the unmodified siRNA.

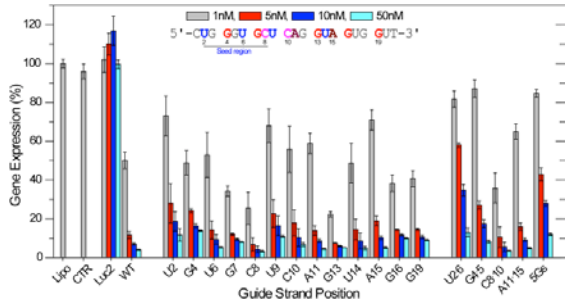


Fig. 3. Gene expression of dGFP at 48 hr for all siRNAs; Lipo = lipofectamin only; CTR = buffer; Luc2 = luciferase2; WT = wild type; grey = 1 nM, red = 5 nM, blue = 10 nM, cyan = 50 nM.

In addition to the singly substituted siRNAs discussed above, we also investigated the impact of multiple modifications in the guide strand on the ability of the resulting siRNAs to inhibit protein expression. Two substitutions with each one of the nucleoside analogs th**C**, th**G**, th**U** and th**A** were explored and one challenging case with five incorporations of th**G**, was examined. Substitutions at positions 2+6 with th**U**, as well as at positions 4+5 and 4+7+13+16+19 with th**G** (oligonucleotides **16** and **10**–**11**, respectively) were somewhat detrimental to RNAi activity (Figure 4).

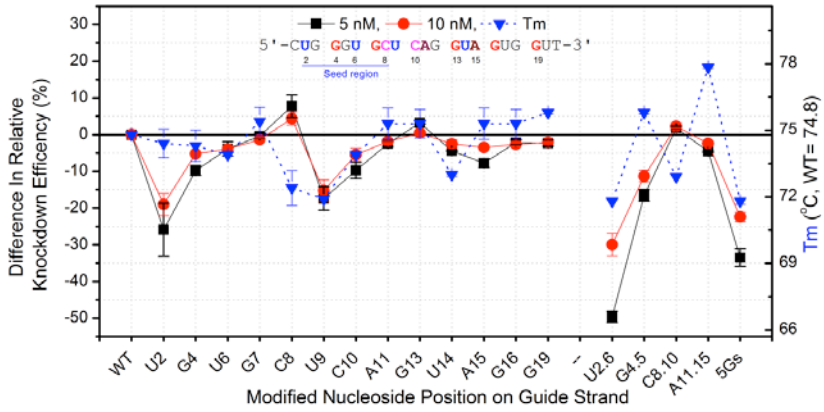


Fig. 4. Gene knockdown and duplex stability compared to WT; black square ■ = 5 nM, red dot ● = 10 nM, blue inverted triangle ▼ = T_m.

siRNAs with double incorporation of th**C** and th**A** at positions 8+10 and 11+15 (**1•19**, **1•22**, respectively), were nearly as effective as the WT oligonucleotide. Interestingly, the oligonucleotide with th**C** at positions 8 and 10 (**19**) exhibited slightly better activity that the native siRNA in all concentrations.

In summary, all modified siRNA duplexes containing substitutions of the native nucleosides with th**A**, th**C**, th**G** and th**U**, synthetic alphabet letters, were found to display potent cellular interference activity. In general, replacing native pyrimidines with their synthetic surrogates, th**C** and th**U**, appears to be slightly more disruptive when compared to the purine replacements. This is not entirely unexpected, due to the sterically more demanding fusion of the thiophene at the pyrimidine's 5,6-positions. Multiple incorporations of modified letters, while overall more detrimental, resulted in highly modified siRNA duplexes with respectable interference activity. Even with the poorest performance, seen for duplexes containing two modifications within the seed region, interference activity of above 50% (compared to the wild type activity) was seen at all concentrations. While not without exceptions, the interference activity followed the relative thermal stability of the siRNA duplexes. Although duplex stability alone cannot necessarily serve as a reliable "high resolution" predictor for cellular potency, this suggests that disrupting the conformational and solvation integrity of the siRNA duplex at certain key domains is likely to negatively impact its interaction with the interference machinery.

We recognize the broad spectrum of RNA interference activities and their dependency on numerous factors, including transfection efficiency, duplex stability, and recognition by the required proteins, etc. Hence, some of the observations made above might be somewhat specific to this particular dGFP-specific siRNA. Nevertheless, the overall excellent performance of highly modified letters, where a thiophene heterocyle replaces the native imidiazole, coupled to additional recent observations[25-27] suggest great utility for such isomorphic RNA building blocks.

We thank the NIH (grant No. GM 069773 to YT) and the W.M. Keck Foundation (to SFD and YT) for support.

Address

[a] Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA 92093-0358, USA.
[b] Cellular & Molecular Medicine, University of California, San Diego, La Jolla, CA 92093-0686, USA.
* Correspondence to: Email address ytor@ucsd.edu

Notes and references

1. G. J. Hannon, *Nature*, 2002, **418**, 244.
2. C. C. Mello and D. Conte, *Nature*, 2004, **431**, 338.
3. G. Meister and T. Tuschl, *Nature*, 2004, **431**, 343.
4. C. C. Mello, *Angew Chem Int Edit*, 2007, **46**, 6985.
5. A. Z. Fire, *Angew Chem Int Edit*, 2007, **46**, 6967.
6. J. Kurreck, *Angew Chem Int Edit*, 2009, **48**, 1378.
7. A. E. Carpenter and D. M. Sabatini, *Nat Rev Genet*, 2004, **5**, 11.
8. J. Moffat and D. M. Sabatini, *Nat Rev Mol Cell Bio*, 2006, **7**, 177.
9. Y. Dorsett and T. Tuschl, *Nat Rev Drug Discov*, 2004, **3**, 318.
10. C. J. Echeverri and N. Perrimon, *Nat Rev Genet*, 2006, **7**, 373.
11. P. H. E. Johnson, in *"RNA Interference, Application to Drug Discovery and Challenges to Pharmaceutical Development"*, John Wiley & Sons, NJ, 2011, pp. 307.
12. C. Ender and G. Meister, *J Cell Sci*, 2010, **123**, 1819.
13. M. A. Behlke, *Oligonucleotides*, 2008, **18**, 305.
14. J. K. Watts, G. F. Deleavey and M. J. Damha, *Drug Discov Today*, 2008, **13**, 842.
15. M. Terrazas and E. T. Kool, *Nucleic Acids Res*, 2009, **37**, 346.
16. E. L. Chernolovskaya and M. A. Zenkova, *Curr Opin Mol Ther*, 2010, **12**, 158.
17. J. B. Bramsen, et. al., *Nucleic Acids Res*, 2009, **37**, 2867.
18. H. Peacock, R. V. Fucini, P. Jayalath, J. M. Ibarra-Soza, H. J. Haringsma, W. M. Flanagan, A. Willingham and P. A. Beal, *J Am Chem Soc*, 2011, **133**, 9200.
19. A. S. Wahba, F. Azizi, G. F. Deleavey, C. Brown, F. Robert, M. Carrier, A. Kalota, A. M. Gewirtz, J. Pelletier, R. H. E. Hudson and M. J. Damha, *Acs Chem Biol*, 2011, **6**, 912.
20. H. Peacock, A. Kannan, P. A. Beal and C. J. Burrows, *J Org Chem*, 2011, **76**, 7295.
21. D. Shin, R. W. Sinkeldam and Y. Tor, *J Am Chem Soc*, 2011, **133**, 14912.
22. Theoretical support for their high similarity to the native nucleobases has been reported; Samanta, P. K.; Manna, A. K.; Pati, S. K. *J. Phys. Chem. B*. 2012, **116**, 7618.
23. A. Eguchi, B. R. Meade, Y. C. Chang, C. T. Fredrickson, K. Willert, N. Puri and S. F. Dowdy, *Nat Biotechnol*, 2009, **27**, 567.
24. For the synthesis of thU, see : Srivatsan, S. G.; Weizman, H.; Tor, Y. *Org Biomol Chem* 2008, **6**, 1334
25. W. Liu, D. Shin, Y. Tor and B. S. Cooperman, *ACS Chemical Biology*, 2013, **8**, 2017.
26. R. W. Sinkeldam, L. S. McCoy, D. Shin and Y. Tor, *Angewandte Chemie International Edition*, 2013, **52**, 14026.
27. L. S. McCoy, D. Shin and Y. Tor, *J Am Chem Soc*, 2014, **136**, 15176.