

Chemical Science

Accepted Manuscript



This article can be cited before page numbers have been issued, to do this please use: S. Walsh, A. H. El-Sagheer and T. Brown, *Chem. Sci.*, 2018, DOI: 10.1039/C8SC02418A.



This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the [author guidelines](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the ethical guidelines, outlined in our [author and reviewer resource centre](#), still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.



Journal Name

ARTICLE

Fluorogenic thiazole orange TOTFO probes stabilise parallel DNA triplexes at pH 7 and above

Sarah Walsh,^{a,b} Afaf Helmy El-Sagheer^{a,c} and Tom Brown^aReceived 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

The instability of DNA triplexes particularly at neutral pH and above severely limits their applications. Here, we demonstrate that the introduction of a thiazole orange (TO) intercalator onto a thymine nucleobase in triplex forming oligonucleotides (TFOs) resolves this problem. The stabilising effects are additive; multiple TO units produce nanomolar duplex binding and triplex stability can surpass that of the underlying duplex. In one example, a TFO containing three TO units increased the triplex melting temperature at pH 7 by a remarkable 50 °C relative to the unmodified triplex. Notably, TO intercalation promotes TFO binding to target sequences other than pure polypurine tracts by the use of 5-(1-propynyl) cytosine (pC) against C:G inversions. By overcoming the instability of triplexes across a broad range of pH and sequence contexts, these very simple 'TOTFO' probes could expand triplex applications into many areas including diagnostics and cell imaging.

Introduction

Triplex-forming oligonucleotides (TFOs) bind in the major groove of double stranded DNA (dsDNA) through Hoogsteen or reverse Hoogsteen hydrogen bonding.^{1,2} While most methods of sequence-specific DNA recognition rely on denaturation of dsDNA, TFOs do not require this, so they have great potential in areas such as gene inhibition,^{3–5} DNA repair and recombination studies,⁶ inhibition of protein-DNA binding,^{7,8} photo-crosslinking of duplex DNA,⁹ DNA mutagenesis,¹⁰ and nanotechnology.¹¹ Unfortunately the most stabilising TFOs, those that form parallel triplexes, have much reduced duplex-binding affinity at neutral pH and above owing to the necessity for cytosine protonation. There have been many efforts to overcome the pH-related instability by introducing a variety of sugar, base and phosphate modifications into TFOs.^{12–16} Furthermore, the requirement for polypurine:polypyrimidine target duplexes is a significant limitation. In order to overcome this, a range of modified bases have been designed to stabilise pyrimidine:purine C:G and T:A inversions and thereby increase the range of accessible target duplexes.^{17,18}

Fluorescent TFOs have been used previously. One important application is visualisation of DNA for detection of genomic sequences and selective molecular recognition,^{19–21} but these applications require a mechanism upon triplex formation for fluorescence enhancement, usually through the use of both a fluorophore and quencher. In principle, this could instead be achieved via a fluorogenic intercalator which

would simultaneously increase triplex stability, potentially across a wide pH-range. Indeed intercalators, both chemically attached to the TFO or untethered, have been shown to stabilise triplexes (and other non-canonical DNA structures).^{22–26} A suitable design for stabilized fluorescent triplexes could be achieved by labelling TFOs with an intercalative fluorophore such as TO (**Figure 1A**). For many applications, the fluorogenic nature of TO is crucial.^{27,28} It acts as a "light switch", experiencing a large enhancement in fluorescence emission upon intercalation into DNA due to inhibition of rotation of the dye around its methine bridge. In this context, a TO dimer has been incorporated into modified TFOs that are known as 'ECHO probes' (exciton controlled hybridization sensitive fluorescent oligonucleotide).²⁹ However, no triplex formation was seen at pH 8 (essential for important *in vitro* enzymatic reactions such as PCR), and the stabilising effect was not additive; two TO dimers did not enhance triplex stability over that obtained from a single addition. Other studies have been carried out on parallel triplexes using terminal (3' and 5') or single internal TO units, but in neither case was stable parallel triplexes reported at pH 7.^{30,31} In these studies, only a limited range of target sequences was examined.

Here, we demonstrate that the chemical attachment of a very simple TO_{B6} modification (**1**) to the TFOs (TOTFO probes, **Figure 1A**) significantly stabilises parallel triplexes at neutral pH and above with a large increase in fluorescence on triplex formation. We also show that this TO modification can be combined with a readily available modified cytosine base, 5-(1-propynyl)-dC (pC) to expand triplex target sequences to those containing pyrimidines (C:G inversions in the duplex, **Figure 1C**), which are otherwise destabilising. Importantly, multiple additions of TO are shown to extend the duplex binding affinity of the TOTFO into the nanomolar region.

^a Department of Chemistry, University of Oxford, Oxford, OX1 3TA, UK.

tom.brown@chem.ox.ac.uk

^b ATDBio Ltd., Oxford Science Park, Oxford, UK.^c Chemistry Branch, Department of Science and Mathematics, Faculty of Petroleum and Mining Engineering, Suez University, Suez 43721, Egypt.

Electronic Supplementary Information (ESI) available: See DOI: 10.1039/x0xx00000x.



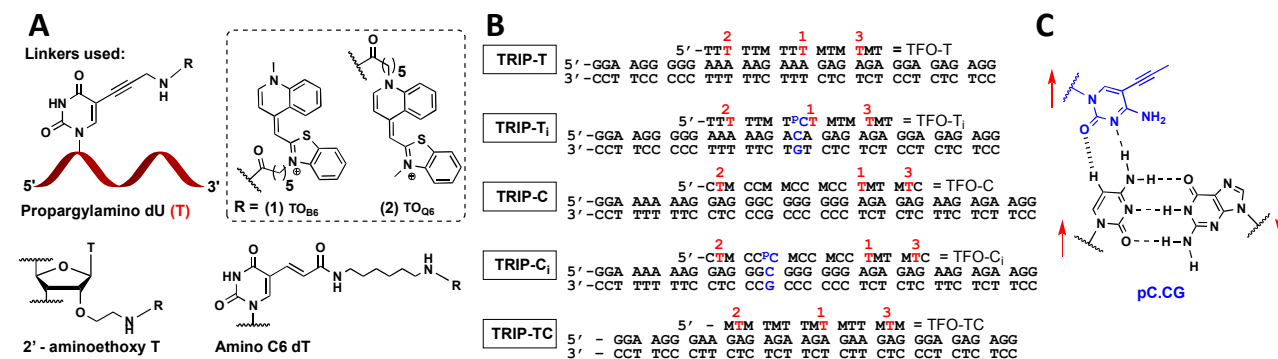


Figure 1: A) Structure of propargylamino deoxyuridine (pdU), 2' aminoethoxy thymidine and amino-C6 dT scaffolds investigated for the attachment of TO_{B6} (1) and TO_{O6} (2) to produce 1xTO, 2xTO and 3xTO TFOs. B) Triplex sequences analysed in this study. Blue text = position of C:G inversion, M = 5-methyl cytosine (^{Me}C), pC = 5-(1-propynyl) cytosine (pC), T = pdU linkage for attachment of TO_{B6} (1) and TO_{O6} (2). In (B), the single-labelled TFOs have TO at position 1, the double-labelled TFOs have TO in positions 1 and 2 and the triple-labelled TFOs have TO in positions 1, 2 and 3. C) Pyrimidine:purine inversion triplet pC.CG. The pC base forms two H-bonds with cytosine of the C:G base pair; one is N-H...N and the other is C-H...O which is a weak H-bond. Red arrows indicate relative directionality of DNA strands. Other base triplets used in this study are shown in Figure S6.

Results and discussion

Triplex design, synthesis and initial evaluation

With the aim of increasing the stability of parallel triplexes, a series of modified 15-18mer TFOs were designed to contain 5-(1-propargylamino)-dU (pdU) at various positions for internal attachment of TO. Since TO binds by intercalation, we chose well separated linker positions in the TFO to avoid hindrance of intercalation that might arise from clustered TO units. The triple bond in pdU is known to have a stabilising effect in triplexes due to increased base stacking.^{32,33} Other attachments such as 2'-aminoethoxy thymidine and a longer linker at the 5-position of thymine (amino C6-dT) (Figure 1A) were also investigated, but pdU was ultimately the most effective (Figure S1). Importantly, the triplexes studied were chosen to cover a range of duplex sequence contexts and stacking environments (Figure 1B). A 'mixed' sequence with 47% G:C duplex content was designed (TRIP-TC) as well as two T-rich sequences (TRIP-T and TRIP-T₁), and two C-rich sequences, (TRIP-C and TRIP-C₁). The subscript 'i' indicates the presence of a C:G inversion in the purine tract of duplex. 5-methyl-cytosine (^{Me}C) (M in Figure 1B) was used at some loci in the TFOs because triplexes containing ^{Me}C are slightly more stable compared to those containing cytosine.³⁴

TO_{B6} and TO_{O6} carboxylic acid active esters were synthesised according to protocol in Section SA.1.1 and were incorporated into the TFOs by labelling the aminopropargyl groups in the DNA strands using PyBOP coupling reagent (Section SA.2). Efficient labelling was confirmed by mass spectrometry (Table S1 and S2). The melting temperatures (T_m values) of the triplexes were then analysed (Table 1 and Table S3). The unlabelled T-rich TFO-T produced a triplex (TRIP-T) with a T_m of 43.2 °C in 10 mM sodium phosphate/10 mM MgCl₂ buffer at pH 5.8 ('high Mg²⁺ buffer') (Entry 1, Table 1). Initially, TO_{O6} was attached to the TFO (i.e. through the quinoline moiety) and this provided a large increase (+14.2 °C) in triplex stability in high Mg²⁺ buffer compared to the unlabelled triplex. However, as found previously for other TO

modifications,²⁹ a second TO_{O6} unit did not confer any additional triplex stabilisation (Figure S2). When a single TO_{B6} unit was attached to the TFO, there was a huge increase in T_m compared to the unlabelled triplex (+23.3 °C) and fortunately, the stabilising effect of TO_{B6} increased with the addition of a second and third fluorophore, reaching 78.0 °C ($\Delta T_m = +34.8$ °C). Therefore, TO_{B6} was used for all remaining studies. The melting studies were also carried out in a buffer closer to cellular salt conditions but still at acidic pH and using Na⁺ ions instead of K⁺ ions (low Mg²⁺ buffer, 10 mM sodium phosphate/150 mM NaCl/2 mM MgCl₂, pH 5.8). Strong stabilisation was observed with the addition of three TO_{B6} units giving a remarkable increase in T_m of 44.8 °C (Table 1, Entries 1-4). Stabilisation was also seen in the TRIP-TC triplex with the addition of three TO units in the TFO, but the overall enhancement in stability (up to 9.7 °C) was not as great, presumably because the unlabelled triplex is already very stable at this pH ($T_m > 70$ °C), surpassing the stability of the underlying duplex. However, at pH 7 when the unmodified triplex becomes unstable, the increase in T_m from TO incorporation is impressive (> 20 °C) (Table 1, Entries 10-11). The effect of sequence on stability of triplexes can be observed in TRIP-C, which is highly unstable but exhibited major stabilisation with the introduction of 3x TO (+26.2 °C in low Mg²⁺ buffer) (Table 1, Entries 12-15). Melting curves and additional melting data are in Figure S3 and Table S3.

To further investigate the sequence specificity of TOTFOs, a scrambled duplex sequence was mixed with 1xTO TFO-T. No triplex formation was observed by UV melting, confirming that the stabilisation gained from intercalation of TO is coupled with retained sequence specificity (Figure S4). To confirm that the TO intercalator in the TFO does not reduce base pair selectivity we investigated triplexes in which TFO-T (3xTO) was mixed with a duplex containing an incorrect base pair to give a T.CG, T.TA or T.GC base triplet instead of T.AT. As expected, there is a decrease in triplex stability (-10 to -14 °C) when a mismatched triplet is introduced, with triplet stability in the following order: T.AT > T.CG > T.TA > T.GC. This confirms that TO intercalation does not destroy the selectivity of the TFO for its fully complementary duplex (Figure S5).



Entry	Triplex	TFO Sequence (5'-3')	pH = 5.8		pH = 7.0	
			10 mM MgCl ₂	150 mM NaCl 2 mM MgCl ₂	10 mM MgCl ₂	150 mM NaCl 2 mM MgCl ₂
T_m (°C)						
1	TRIP-T	TTT TTM TTT MTM TMT	43.2	31.9	39.3	28.4
2	1xTO TRIP-T	TTT TTM TTT MTM TMT	66.5	58.0	60.0	50.6
3	2xTO TRIP-T	TTT TTM TTT MTM TMT	76.4	65.1	65.5	58.7
4	3xTO TRIP-T	TTT TTM TTT MTM TMT	78.0	76.7	74.0	72.0
5	TRIP-T _i Control	TTT TTM TTT MTM TMT	49.0	29.6	-	-
6	TRIP-T _i	TTT TTM T ^F CT MTM TMT	50.3	31.2	23.0	-
7	1xTO TRIP-T _i	TTT TTM T ^F CT MTM TMT	66.3	48.7	45.0	35.9
8	2xTO TRIP-T _i	TTT TTM T ^F CT MTM TMT	74.6	66.9	57.5	51.3
9	3xTO TRIP-T _i	TTT TTM T ^F CT MTM TMT	76.1	73.2	73.4	65.9
10	TRIP-TC	MTM TMT TMT MTT MTM	77.0	72.7	52.6	49.0
11	3xTO TRIP-TC	MTM TMT TMT MTT MTM	86.7	76.3	73.2	73.4
12	TRIP-C	CTM CCM MCC MCC TMT MTC	44.0	35.6	<i>nt</i>	<i>nt</i>
13	1xTO TRIP-C	CTM CCM MCC MCC TMT MTC	50.6	46.2	25.7	19.9
14	2xTO TRIP-C	CTM CCM MCC MCC TMT MTC	54.3	55.1	30.4	26.0
15	3xTO TRIP-C	CTM CCM MCC MCC TMT MTC	62.5	61.8	34.6	34.8
16	TRIP-C _i Control	CTM CCT MCC MCC TMT MTC	31.5	21.0	-	-
17	TRIP-C _i	CTM CC ^F C MCC MCC TMT MTC	39.5	26.6	<i>nt</i>	<i>nt</i>
18	1xTO TRIP-C _i	CTM CC ^F C MCC MCC TMT MTC	42.9	29.7	<i>nt</i>	<i>nt</i>
19	2xTO TRIP-C _i	CTM CC ^F C MCC MCC TMT MTC	45.3	36.7	<i>nt</i>	<i>nt</i>
20	3xTO TRIP-C _i	CTM CC ^F C MCC MCC TMT MTC	56.4	44.7	25.4	24.7

Table 1: T_m values were obtained from the first derivatives of the melting curves (black = UV melting analysis, blue = fluorescence melting analysis). Fluorescence melting was used when the T_m could not be determined by UV melting. Buffer at pH 5.8 = 10 mM sodium phosphate and varying salt conditions according to the table header, buffer at pH 7 = 10 mM MOPS and varying salt conditions. T_m values are an average of two independent experiments, both of which are an average of 5 measurements. pC = pC, M = MeC, T = TO attached to pU, *nt* = no triplex detected. Full triplex sequences are in Figure 1B. T_m of the underlying duplexes are as follows (pH 5.8/pH 7): TRIP-T = 73.1/73.4 °C, TRIP-T_i = 73.9/74.3 °C, TRIP-TC = 73.2/74.1 °C, TRIP-C = 74.6/75.1 °C, TRIP-C_i = 77.3/77.5 °C.

Stabilisation of triplex inversions

Having established strong stabilisation of parallel triplexes at uninterrupted polypurine tracts, attention was focussed on triplexes in which the duplex contains a triplex-destabilising base pair inversion. To study this, a C:G base pair was inserted in the poly-A:T or poly-G:C tract of the duplex DNA target (TRIP-T_i and TRIP-C_i). Previous studies on modified bases that can recognise and stabilise C:G inversions require multi-step synthesis of special phosphoramidite monomers.^{17,35} Here, 5-(1-propynyl) cytosine (pC), a commercially available monomer, was used to form a Hoogsteen interaction with the inverted cytosine base of the duplex (a pC.CG triplet). The pC nucleobase has been investigated in TFOs previously for binding to G:C base pairs, where it was destabilising due to its resistance to protonation at N3 (low pK_aH).³⁶ However, for C:G recognition, this is advantageous as protonation of pC is not required (Figure 1C, Figure S6). A comparison to thymine, the most effective natural base to stabilise a C:G inversion (Figure S5),³⁵ indicated that pC is slightly superior, giving a ΔT_m = +1.2 °C and +1.6 °C in high and low Mg²⁺ buffers respectively (Table 1, Entry 7). To investigate the selectivity of pC for the C:G base pair, TFO-T_i was mixed with three different duplexes to form triplexes containing a mismatch triplet (pC.GC, pC.AT and

pC.TA instead of pC.CG) (Figure S7). The pC.AT triplet is interesting because the A:T base pair would cause selectivity problems when using T to recognise C:G in a genomic context due to the fact that the T.AT triplet is more stable than T.CG. Fortunately, pC.AT produces a less stable triplex compared to pC.CG at pH 7 (ΔT_m = -12.3 °C) and is much less stable than T.AT (ΔT_m = -24.7 °C). pC.GC produces a triplex of similar stability to pC.CG in low pH but we see this stability is highly pH-dependent (ΔT_m = -24.6 °C) as the pH is increased whilst pC.CG remains far less pH-dependent.

Incorporation of the C:G inversion into the TRIP-T_i and TRIP-C_i sequence environments showed that the TRIP-C_i triplex is far less stable than TRIP-T_i (ΔT_m = -10.8 °C, pH 5.8, high Mg²⁺ buffer, Table 1, Entries 6 and 17). This is expected as contiguous runs of cytosine bases destabilise triplexes because the pK_aH of each individual C-nucleobase is reduced due to their close proximity.^{37,38} The inclusion of a single TO unit in TFO-T_i led to a large increase in the T_m of the TRIP-T_i triplex (ΔT_m > 16.0 °C) and a remarkable +43.6 °C increase in low Mg²⁺ buffer for the triplex with three units of TO (Table 1, Entries 7-9). A smaller increase in stability was observed in the TRIP-C_i triplex with one TO-addition (ΔT_m ~ 3 °C, Table 1, Entry 18). However, the T_m increased much more (by 16 °C) when three TO units were introduced into the TFO-C_i (Table 1, Entry



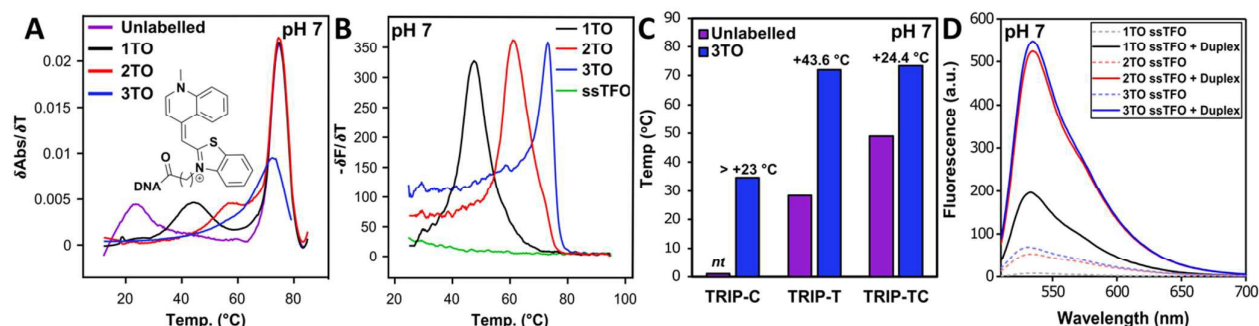


Figure 2: A) First derivatives of UV absorption curves of TRIP-T_i triplex; unlabelled, 1x, 2x and 3x TO at pH 7. The first transition is triplex denaturation and the second (higher temperature) is duplex denaturation. T_m of 3x TO could not be determined by UV melting as the triplex T_m overlapped with the duplex transition. Duplex concentration is 1.0 μM and TFO concentration is 2.5 μM, λ = 260 nm. B) First derivative of fluorescence melting curves of the 1x, 2x and 3x TO TRIP-T_i triplex with the ssTFO as a control at pH 7, 0.25 μM of duplex and 0.625 μM TFO, see Section SA.6 for excitation and emission channels. C) Triplex stability vs. sequence environment. C⁺.GC-rich sequence TRIP-C, T.AT-rich sequence TRIP-T and mixed sequence TRIP-TC. nt = no triplex detected, therefore, accurate ΔT_m could not be determined for the labelled triplex. Buffer contains 10 mM MOPS, 150 mM NaCl, 2 mM MgCl₂, pH 7. All T_m values are an average of two independent experiments, both of which are an average of 5 measurements. D) Fluorescence intensity of 1x, 2x, 3x TO (TRIP-T), ssTFO in the absence (dashed line) and presence (solid line) of the duplex at pH 7. Concentrations of 1.8 μM duplex and 1 μM TFO used.

20). Hence, in general TO intercalation and pC as a C:G recognition monomer more than compensates for the destabilising effect caused by the triplet inversion.

Fluorescence melting

As triplexes become more stable, UV monitoring of thermal denaturation is a less reliable determinant of T_m because melting of the triplex is obscured by the duplex melting transition (Figure S8) and this can also distort the triplex melting transition. In such cases, fluorescence melting can be used instead, because in this technique only the triplex to duplex transition is detectable.¹⁹ Fluorescence melting was therefore employed to analyse triplexes, and in cases where both UV and fluorescence melting could be measured the melting temperatures were in good agreement with slight differences due to the DNA concentrations used (1 μM for UV melting, 0.625 μM for fluorescence melting). Also, small differences in T_m result from the fact that the two techniques

monitor different DNA melting-related processes. UV melting is sensitive to the dissociation of strands *via* the unstacking of bases whilst fluorescence melting reports on intercalation of the TO into the triplex.

TOTFO triplex stabilisation at higher pH

As discussed previously, one of the limitations of triplexes and their applications is the necessity to conduct experiments at low pH. With stabilisation from TO, it is possible to increase the pH and still form stable triplexes. 3xTO TRIP-T_i is one of the most stable triplex at pH 7, with a T_m > 70 °C (Figure 2A, 2B) and the additive effective of TO is clear. Substantial stabilisation is also observed for TRIP-T and TRIP-TC compared to their unlabelled triplexes in low Mg²⁺ buffer (Figure 2C). However, 3xTO TRIP-C is much less stable (T_m ~ 35 °C). This is because it is rich in cytosine bases (78%) which require protonation to form stable C⁺.GC triplets and this will not occur at neutral pH. It is noteworthy that we could not observe any

Triplex	pH = 5.8					pH = 7.0				
	λ _{em} ^{ss}	λ _{em} ^T	F ^{ss} @ λ _{em} ^T	F ^T @ λ _{em} ^T	F ^T /F ^{ss}	λ _{em} ^{ss}	λ _{em} ^T	F ^{ss} @ λ _{em} ^T	F ^T @ λ _{em} ^T	F ^T /F ^{ss}
	nm		a.u.			nm		a.u.		
1xTO TRIP-T	534	533	10.72	249.16	23.2	536	533	7.49	196.90	26.3
2xTO TRIP-T	532	534.5	72.23	605.85	8.4	533	534.5	52.41	523.88	10.0
3xTO TRIP-T	531.5	535	119.61	578.91	4.8	531	534.5	67.45	547.83	8.1
1xTO TRIP-T _i	531.5	535	87.77	704.85	8.0	530	531.5	51.50	642.48	12.5
2xTO TRIP-T _i	534.5	535	52.12	348.26	6.7	532	534.5	29.49	321.42	10.9
1xTO TRIP-C	538	535.5	62.88	485.31	7.7	531.5	532	100.26	668.40	6.7
2xTO TRIP-C	564	538	32.89	640.17	19.5	533	535	38.44	875.43	22.8
3xTO TRIP-C	585	539	18.68	468.42	25.1	533.5	535	19.93	711.12	35.7
1xTO TRIP-C _i	533	534.5	98.23	462.54	4.7	531.5	532.5	69.62	485.58	7.0
2xTO TRIP-C _i	535.5	536	80.78	400.23	5.0	565.5	536	42.88	440.69	10.3
3xTO TRIP-C _i	533	535.5	100.15	288.61	2.9	582.5	537.5	52.19	315.20	6.0

Table 2: Data from fluorescence intensity measurements. Buffer at pH 5.8 = 10 mM sodium phosphate, 10 mM MgCl₂. Buffer at pH 7 = 10 mM MOPS, 10 mM MgCl₂. Fluorescence values are averages of two independent measurements. λ_{em}^{ss} = max. emission wavelength of single strand TFO, λ_{em}^T = max. emission wavelength of formed triplex, F^{ss} @ λ_{em}^T = Fluorescence intensity of ssTFO at the max. emission wavelength of the triplex, F^T @ λ_{em}^T = Fluorescence intensity of triplex at the max. emission wavelength of the triplex, F^T/F^{ss} = ratio of fluorescence intensity calculated by comparison of the ssTFO fluorescence intensity and the formed triplex fluorescence intensity at λ_{em}^T.



triplex formation for unmodified TRIP-C at pH 7 but could overcome this instability with TO attachment. In some sequence contexts, very stable triplexes can be produced at pH 7 with only one TO modification e.g. 1xTO TRIP-T, $T_m = 60.0$ °C and 50.6 °C in high and low Mg^{2+} buffer respectively (Table 1, Entry 2). Stable triplexes can even be formed above neutral pH; TRIP-TC with 3xTO has a T_m of 54 °C at pH 8 (Table S4). Hence, these studies demonstrate that triplex formation in biologically compatible conditions is feasible using the simple TO_{B6} modification.

The influence of the T.AT/C⁺.GC triplet ratio on unlabelled and TO-labelled triplex stability at pH 7 is shown in Figure 2C. C⁺.GC-rich triplexes are unstable, T.AT-rich triplexes are much more stable and triplexes with an even mixture of T.AT and C⁺.GC triplets are the most stable. All benefit from the presence of TO immensely (some with $\Delta T_m > +40$ °C).

Evaluation of fluorescence enhancement on triplex formation

To quantify the enhancement of fluorescence intensity upon triplex formation, the fluorescence spectra of the TFOs both in the absence and presence of the complementary duplex were measured (Figure 2D). TFOs excited at 484 nm showed minimal fluorescence in single-strand form (ssTFO). Upon addition of the duplex, fluorescence intensity was greatly enhanced, 23-fold for the TRIP-T triplex at pH 5.8. At neutral pH, slightly greater enhancement (26-fold) was seen upon triplex formation (Table 2). The fluorescent intensity ratios between the free TFOs and the triplexes are given in Table 2. A decrease in the ratio is observed as additional TO units are introduced into the TFO, even though the triplexes with 2 and 3 TO units are much brighter than that with a single TO. This is due to the higher fluorescence of the single stranded TFOs that contain multiple TO units. Fluorescence increases slightly at higher pH, and the trends are similar. Further fluorescent analysis is provided in Figure S9. Quantum yield measurements were carried out on the triplexes (Table S5) and lie mostly between 0.13 and 0.23 at pH 7, varying as a function of sequence context. This is in line with published data on thiazole orange in DNA duplexes and higher ordered structures.^{39–41}

For many applications, the inherent fluorescence of thiazole orange will be adequate for detection purposes, but in specific cases a brighter and more stable fluorophore will be required e.g. for super-resolution imaging of duplex sequences in genomic DNA in live or fixed cells.⁴² Our TOTFO concept was designed with this in mind; the termini of the TFO are unmodified and therefore available for orthogonal fluorescent labelling. This was demonstrated by the synthesis of an ATTO 647N click-labelled TOTFO (TO-ATTO TRIP-T_i, Figure S10) which forms a stable triplex with its target duplex.

TOTFO Triplex studies by other biophysical techniques

Triplex formation and stability were also monitored by other techniques. The TRIP-T_i triplex was investigated by native polyacrylamide gel-electrophoresis (PAGE) at pH 7 (Figure 3).

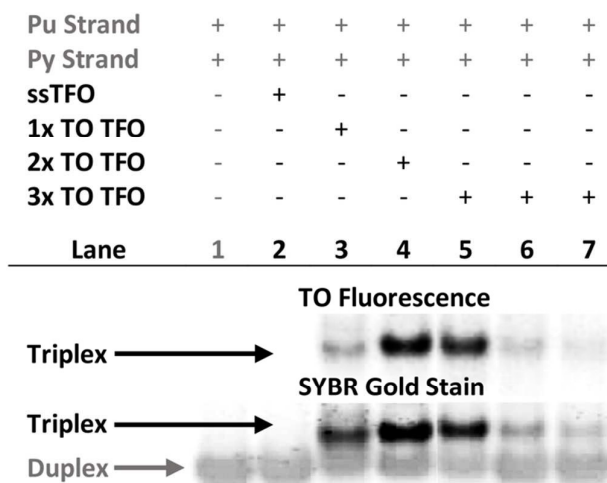


Figure 3: Fluorescence and SYBR Gold Staining, 15% non-denaturing polyacrylamide gels showing TRIP-T_i triplex formation. Buffer is 40 mM MOPS containing 10 mM NaOAc and 10 mM $MgCl_2$ (pH 7). All duplex strand concentrations are 1.0 μ M. Lane 2 - [10:1] Unlabelled TFO: duplex, Lane 3/4 - [10:1] TFO: duplex, Lanes 5,6,7 - [10:1], [5:1], [2.5:1] TFO: duplex respectively. Lane 2 shows no triplex formation due to the low T_m of the unlabelled triplex at neutral pH ($T_m = 23$ °C).

Triplexes were separated from the duplex by gel electrophoresis based on charge and size, then visualised by inherent TO fluorescence and SYBR Gold staining. The duplex can be identified from lane 1 and the triplex can be observed in lanes 3–7 where the TOTFOs were added to the duplexes. Under SYBR Gold staining it can be seen that the triplex band migrates slightly slower than the duplex band. As expected, decreasing the TFO:duplex ratio results in less intense triplex bands (compare lanes 5, 6, 7). The triplex band was extracted from the gel and analysed by mass spectrometry, which revealed three peaks of the expected masses, confirming the presence of each strand of the triplex (Figure S11).

SPR analysis was carried out to corroborate the gel-electrophoresis and thermal denaturation studies. The unlabelled triplexes TRIP-T_i produced a K_D of 1.5 μ M in accordance with previous reports in which the dissociation constants of triplexes are generally within the micromolar range.⁴³ A progressive increase in binding was observed with two then three additions of TO, three additions giving a binding constant of 3.9 nM (Figure S12). This data supports the trends observed in the triplex melting studies discussed above.

Conclusions

Triplex stability is greatly enhanced by incorporation of thiazole orange between the 5-position of thymine and the benzothiazole moiety of TO (TO_{B6}) in triplex forming oligonucleotides (TOTFOs). Very stable triplexes are obtained at (and above) neutral pH using this straightforward methodology. The stabilising effects of TO are cumulative; the most stable triplexes at pH 7 was formed with three additions of TO in TFOs that target a A:T-rich and mixed-sequence duplexes, and T_m increases up to 45 °C were achieved. This



greatly increased stability has enabled us to expand the range of target duplexes at neutral pH to polypurine sequences that contain a pyrimidine inversion using pC as a C:G recognition monomer. In one such case an increase in T_m of 50 °C was obtained at pH 7. TOTFOs undergo large enhancement in fluorescence upon triplex formation and this could be significant for many applications. TOTFOs are also substrates for dual dye labelling enabling the attachment of very bright fluorophores for stringent biological studies (ATTO-TOTFOs). This simple approach to triplex stabilisation could unlock many practical applications and might also be effective in the stabilisation of highly unstable antiparallel triplexes.

Finally, the method used in this work for oligonucleotide labelling involved reaction of TO active esters with amino-modified oligonucleotides. Now that we have established which TO-nucleotide is the most effective (TO_{B6}-pdU) it will be more efficient in future to introduce it directly into TFOs during solid-phase synthesis via its phosphoramidite monomer and equip future generations of the probes with substituents that interrogate the target sequence in both the major and minor groove.⁴⁴

Acknowledgements and Conflicts of Interests

This project received funding from the European Union Horizon 2020 research and innovation programme under Marie Skłodowska Curie agreement 642023 (H2020-MSCA-ITN-2014-642023) and UK BBSRC grant BB/J001694/2 (extending the boundaries of nucleic acid chemistry).

There are no conflicts to declare.

References

- H. E. Moser and P. B. Dervan, *Science*, 1987, **238**, 645–650.
- K. Hoogsteen, *Acta. Cryst.*, 1959, **12**, 822–823.
- A. R. Morgan and R. D. Wells, *J. Mol. Biol.*, 1968, **37**, 63–80.
- J. C. Hanvey, M. Shimizu and R. D. Wells, *Nucleic Acids Res.*, 1990, **18**, 157–161.
- Y. Taniguchi and S. Sasaki, *Org. Biomol. Chem.*, 2012, **10**, 8336.
- A. Mukherjee and K. M. Vasquez, *Biochimie*, 2011, **93**, 1197–1208.
- L. Maher, B. Wold and P. Dervan, *Science*, 1989, **245**, 725–730.
- S. B. Noonberg, G. K. Scott, C. A. Hunt, M. E. Hogan and C. Benz, *Gene*, 1994, **149**, 123–126.
- K. M. Vasquez, T. G. Wensel, M. E. Hogan and J. H. Wilson, *Biochemistry*, 1996, **35**, 10712–10719.
- R. Zain, C. Marchand, J. S. Sun, C. H. Nguyen, E. Bisagni, T. Garestier and C. Helene, *Chem. Biol.*, 1999, **6**, 771–777.
- Y. Hu, A. Ceconello, A. Idili, F. Ricci and I. Willner, *Angew. Chem., Int. Ed.*, 2017, **56**, 15210–15233.
- N. Ben Gaied, Z. Zhao, S. R. Gerrard, K. R. Fox and T. Brown, *ChemBioChem*, 2009, **10**, 1839–1851.
- V. Kumar, V. Kesavan and K. V. Gothelf, *Org. Biomol. Chem.*, 2015, **13**, 2366–2374.
- T. Sato, Y. Sato and S. Nishizawa, *J. Am. Chem. Soc.*, 2016, **138**, 9397–9400.
- L. Lacroix, J. Lacoste, J. F. Reddoch, J. L. Mergny, D. D. Levy, M. M. Seidman, M. D. Matteucci and P. M. Glazer, *Biochemistry*, 1999, **38**, 1893–1901.
- H. Torigoe, Y. Hari, M. Sekiguchi, S. Obika and T. Imanishi, *J. Biol. Chem.*, 2001, **276**, 2354–2360.
- V. Malnuit, M. Duca and R. Benhida, *Org. Biomol. Chem.*, 2011, **9**, 326–336.
- Y. Hari, S. Obika and T. Imanishi, *Eur. J. Org. Chem.*, 2012, **2012**, 2875–2887.
- R. J. Darby, M. Sollogoub, C. McKeen, L. Brown, A. Risitano, N. Brown, C. Barton, T. Brown and K. R. Fox, *Nucleic Acids Res.*, 2002, **30**, e39.
- Y. Wang, Q. Sun, L. Zhu, J. Zhang, F. Wang, L. Lu, H. Yu, Z. Xu and W. Zhang, *Chem. Commun.*, 2015, **51**, 7958–7961.
- T. N. Grossmann, L. Röglin and O. Seitz, *Angew. Chem., Int. Ed.*, 2007, **46**, 5223–5225.
- W. D. Wilson, F. Taniou, S. Mizan, S. Yao, A. Kiselyov, G. Zon and L. Strekowski, *Biochemistry*, 1993, **32**, 10614–10621.
- K. R. Fox and T. Brown, in *DNA Conjugates and Sensors*, 2012, pp. 75–102.
- J. S. Sun, J. C. François, T. Montenay-Garestier, T. Saison-Behmoaras, V. Roig, N. T. Thuong and C. Hélène, *Proc. Natl. Acad. Sci. U. S. A.*, 1989, **86**, 9198–202.
- M. Grigoriev, D. Praseuth, P. Robin, A. Hemar, T. Saison-Behmoaras, A. Dautry-Varsat, N. T. Thuong, C. Helene and A. Harel-Bellan, *J. Biol. Chem.*, 1992, **267**, 3389–3395.
- S. Bevers, S. Schutte and L. W. McLaughlin, *J. Am. Chem. Soc.*, 2000, **122**, 5905–5915.
- T. Sato, Y. Sato and S. Nishizawa, *Chem. - A Eur. J.*, 2017, **23**, 4079–4088.
- I. Lubitz, D. Zikich and A. Kotlyar, *Biochemistry*, 2010, **49**, 3567–3574.
- S. Ikeda, H. Yanagisawa, M. Yuki and A. Okamoto, *Artif. DNA PNA XNA*, 2013, **4**, 19–27.
- O. Doluca, T. K. Hale, P. J. B. Edwards, C. Gonzalez and V. V. Filichev, *Chempluschem*, 2014, **79**, 58–66.
- B. L. Renard, R. Lartia and U. Asseline, *Org. Biomol. Chem.*, 2008, **6**, 4413–4425.
- K. R. Fox, *Curr. Med. Chem.*, 2000, **7**, 17–37.
- J. Bijapur, M. D. Keppler, S. Bergqvist, T. Brown and K. R. Fox, *Nucleic Acids Res.*, 1999, **27**, 1802–1809.
- O. A. Amosova and J. R. Fresco, *Nucleic Acids Res.*, 1999, **27**, 4632–4635.
- S. R. Gerrard, M. M. Edrees, I. Bouamaied, K. R. Fox and T. Brown, *Org. Biomol. Chem.*, 2010, **8**, 5087–5096.
- B. C. Froehler, S. Wadwani, T. J. Terhorst and S. R. Gerrard, *Tetrahedron Lett.*, 1992, **33**, 5307–5310.
- A. M. Soto, J. Loo and L. A. Marky, *J. Am. Chem. Soc.*, 2002, **124**, 14355–14363.
- J. S. Lee, M. L. Woodsworth, L. J. Latimer and A. R. Morgan, *Nucleic Acids Res.*, 1984, **12**, 6603–6614.
- E. Privat, T. Melvin, F. Mérola, G. Schweizer, S. Prodhomme, U. Asseline and P. Vigny, *Photochem. Photobiol.*, 2002, **75**, 201–210.



Journal Name

ARTICLE

- 40 J. Nygren, N. Svanvik and M. Kubista, *Biopolymers*, 1998, **46**, 39–51.
- 41 I. Lubitz, D. Zikich and A. Kotlyar, *Biochemistry*, 2010, **49**, 3567–3574.
- 42 S. Beater, P. Holzmeister, E. Pibiri, B. Lalkens and P. Tinnefeld, *Phys. Chem. Chem. Phys.*, 2014, **16**, 6990–6996.
- 43 P. J. Bates, H. S. Dosanjh, S. Kumar, T. C. Jenkins, C. A. Laughton and S. Neidle, *Nucleic Acids Res.*, 1995, **23**, 3627–3632.
- 44 R. Haug, M. Kramer and C. Richert, *Chem. - A Eur. J.*, 2013, **19**, 15822–15826.

