

Organic & Biomolecular Chemistry

Accepted Manuscript



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 [Information for Authors](#).

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](#) 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

Optimization of Fluorescent 8-Heteroaryl-Guanine Probes for Monitoring Protein-Mediated Duplex→G-Quadruplex Exchange†

Kaila L. Fadock,^a Richard A. Manderville,^{*a} Purshotam Sharma^b and Stacey D. Wetmore^{*b}

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

DOI: 10.1039/x0xx00000x

www.rsc.org/

In this study, we describe the thermal and optical properties of the thrombin binding aptamer (TBA) that has been modified at *syn*-G-tetrad positions with fluorescent 8-heteroaryl-2'-deoxyguanosine derivatives consisting of pyrrolyl (^{Py}dG), furyl (^{Fur}dG), thienyl (ThdG), benzofuryl (^{Bfur}dG), indolyl (^{Ind}dG) and benzothienyl (^{Bth}dG). Insertion of the modified base into the *syn*-G₅ position of TBA decreases duplex stability, but enhances stability of the antiparallel G-quadruplex (GQ) structure produced by TBA in the presence of K⁺ ion and its molecular target, thrombin. The resulting modified TBA (mTBA) oligonucleotides have been employed in duplex→GQ exchange to monitor thrombin binding affinity and rates of GQ formation driven by thrombin binding. Our studies demonstrate that 8-heteroaryl-dG bases can be inserted into *syn*-G-tetrad positions of TBA without perturbing thrombin binding affinity and that the 8-thienyl-dG (ThdG) analog is particularly useful as an emissive probe for monitoring duplex→GQ exchange due to its heightened emissive sensitivity to change in DNA topology compared to the other 8-heteroaryl-dG analogs. The positional impact of a single ThdG probe versus multiple ThdG incorporation at *syn*-G sites of TBA highlight an advantage for di-substituted mTBA oligonucleotides for increased emission intensity and rates of duplex→GQ exchange that can be vital for diagnostics through aptamer detection strategies.

Introduction

Nucleic acid aptamers are single-stranded DNA/RNA-based ligands that bind molecular targets with high affinity and specificity for applications as therapeutic agents and diagnostic tools.^{1,2} They are generated by an *in vitro* selection method commonly named SELEX (systematic evolution of ligands by exponential enrichment).^{3,4} Aptamer modifications can enhance their utility, and it has been demonstrated that sugar and phosphate modifications promote nuclease resistance for improved therapeutic use.⁵⁻⁸ Additionally, base modifications increase chemical diversity for enhanced protein⁹⁻¹¹ and small molecule binding,^{12,13} and targeting cancer cells.¹⁴

Aptamer base modifications can also provide a diagnostic signal for target detection. For G-rich oligonucleotides that fold into G-quadruplex (GQ) structures, 8-aryl-2'-deoxyguanosine (8-aryl-dG) residues can replace normal G nucleobases within G-tetrads without disturbing H-bonding interactions and GQ folding.¹⁵⁻¹⁸ Furthermore, aryl ring attachment to the 8-site of dG extends the π -surface of the G nucleobase to afford a fluorescent G mimic.¹⁵⁻¹⁸ The G-G base-

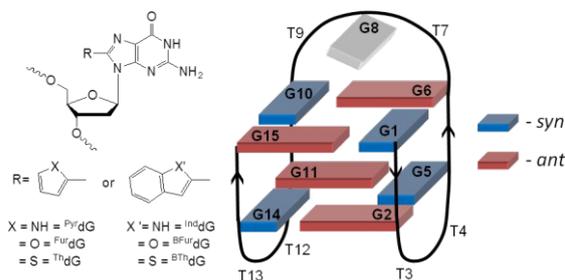


Fig. 1 Structures of 8-heteroaryl-dG residues and the GQ produced by TBA (5'-GGTTGGTGTGGTTGG-3'), *anti*-Gs are shown in red, *syn*-Gs are shown in blue.

stacking and restricted motions within GQ structures can amplify the emission of the 8-aryl-dG base compared to its emission in the single-strand or especially in the duplex.¹⁵ This fluorescence sensitivity to GQ folding should permit the application of duplex→GQ exchange¹⁹ as an effective strategy for detecting molecular target binding by the wide variety of nucleic acid aptamers that produce GQs upon ligand binding.^{20,21}

Recently, we established proof-of-concept for probe performance by inserting the fluorescent 8-furyl-dG (^{Fur}dG) base into the thrombin binding aptamer (TBA),²²⁻²⁴ which represents the most commonly employed GQ-based aptamer.²⁵ Native TBA folds into an intramolecular antiparallel GQ in the presence of certain metal cations and thrombin.²⁶ The GQ structure contains two G-tetrads with alternating *syn*- and *anti*-G residues and three loops.²⁷ In K⁺-solution, the TBA duplex is more stable than the GQ (duplex thermal melting

^a Department of Chemistry & Toxicology, University of Guelph, Guelph, ON, Canada N1G 2W1, E-mail: rmanderv@uoguelph.ca

^b Department of Chemistry & Biochemistry, University of Lethbridge, Lethbridge, AB, Canada T1K 3M4, E-mail: Stacey.Wetmore@uleth.ca

† Electronic supplementary information (ESI) available: Fig. S1–S4 described in the text, NMR spectra of synthetic samples, MS spectra of modified TBA oligonucleotides. See DOI: 10.1039/x0xx00000x

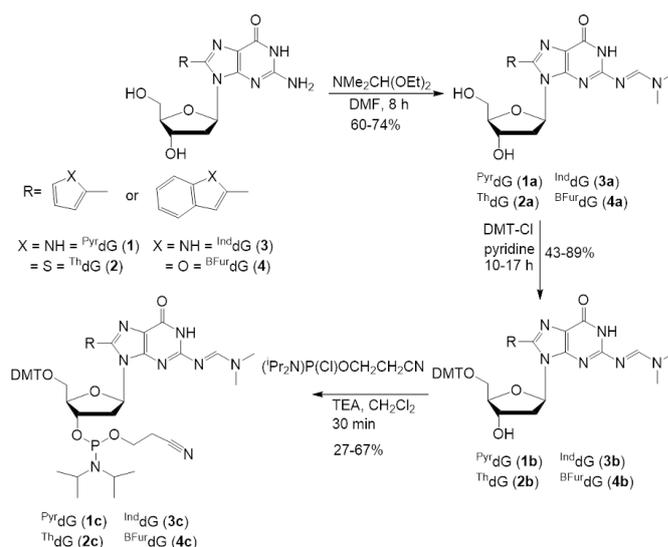
temperature (T_m) \sim 11 °C greater than the GQ) permitting GQ \rightarrow duplex exchange to be monitored by addition of the complementary strand to the GQ structure.²⁸ However, due to a *syn*-preference, the ^{Fur}dG modification strongly enhances GQ stability at *syn*-G₅ ($\Delta T_m \sim$ 9 °C) and strongly decreases duplex stability ($\Delta T_m \sim$ -7 °C), making the GQ more stable than the duplex (GQ $T_m \sim$ 4 °C greater than duplex).²² Furthermore, the ^{Fur}dG probe exhibits quenched fluorescence in the duplex, but lights-up in the GQ due to effective energy-transfer from the natural G bases.²² These characteristics permitted the use of ^{Fur}dG at *syn*-G₅ for monitoring K⁺ ion²² and thrombin binding²³ using duplex \rightarrow GQ exchange.

Although our studies with ^{Fur}dG at the *syn*-G₅ position of TBA demonstrated the potential utility of 8-heteroaryl-dG residues for monitoring GQ-folding-aptamer/protein interactions using fluorescence spectroscopy, several important factors must be explored in order to optimize probe performance. Specifically, alternative 8-heteroaryl-dG nucleosides have been reported and such analogues may be more effective than ^{Fur}dG at signaling thrombin binding by fluorescence spectroscopy. TBA also contains three other *syn*-G positions (G₁, G₁₀ and G₁₄, Fig. 1) that have yet to be tested for probe performance. Similarly, the impact of multiple probe incorporation has yet to be established. Furthermore, the thrombin binding affinity of the modified TBA (mTBA) containing ^{Fur}dG was not compared to the native 15mer and the binding studies were carried out in Na⁺-solution²³ despite TBA producing a more stable GQ structure with a K⁺ ion, which is also known to enhance the thrombin-inhibiting effect of the aptamer.²⁶ Thus, further studies are warranted to optimize 8-heteroaryl-dG probe performance at *syn*-G-tetrad positions within TBA, determine their impact on thrombin binding affinity compared to native TBA in K⁺-solution, and establish their impact on the rate of duplex \rightarrow GQ exchange driven by thrombin binding.

In the present study, we describe the thermal (UV-vis T_m values) and optical (circular dichroism (CD) and fluorescence) properties of mTBA duplex and GQ structures containing *syn*-G₅ 8-heteroaryl-dG residues (furyl (^{Fur}dG), pyrrolyl (^{Pyr}dG), thienyl (ThdG), benzofuryl (^{Bfur}dG), benzothienyl (^{BTh}dG) and indolyl (^{Ind}dG), Fig. 1). Thrombin binding by the mTBA samples in K⁺-solution using duplex \rightarrow GQ exchange established the 5-membered 8-thienyl derivative ThdG as the best performing turn-on emissive probe for thrombin detection. The ThdG analog was then selected for further studies to establish positional and multiple probe impact on the rate of duplex \rightarrow GQ exchange and thrombin binding affinity to the mTBA oligonucleotides. Our results reveal important new information concerning the utility of internal fluorescent DNA base surrogates for monitoring target binding by aptamers that produce antiparallel GQ structures.

Results and discussion

Probe performance at *syn*-G₅. The 8-heteroaryl-dG nucleosides were synthesized as previously described,^{29,30} and converted into phosphoramidites (Scheme 1) for use on a DNA synthesizer (see ESI for NMR spectra of phosphoramidites and



Scheme 1 Synthesis of 8-heteroaryl-dG phosphoramidites

ESI-MS data (Table S1) and spectra of mTBA oligonucleotides) for 8-heteroaryl-dG probe incorporation into *syn*-G₅ of TBA. Full synthetic details for phosphoramidites derived from ^{Fur}dG³⁰ and ^{BTh}dG³¹ have been previously described.

Thermal melting studies (Table 1) were initially carried out to determine the impact of the 8-heteroaryl-G modification on duplex (T_m 's in Na⁺-solution) and GQ (T_m 's in K⁺-solution) stability. As noted previously for ^{Fur}dG at *syn*-G₅,²² the 8-heteroaryl-dG modifications decrease duplex stability, but enhance GQ stability. In general, it was noted that the X = S heteroatom (ThdG and ^{BTh}dG) had a greater destabilizing influence on duplex stability and a diminished influence on GQ stability (see $\Delta\Delta T_m$ values) compared to X = NH (^{Pyr}dG and ^{Ind}dG) and X = O (^{Fur}dG and ^{Bfur}dG). For each heteroatom it was also noted that the 5-membered derivatives tended to be less destabilizing in the duplex and more stabilizing in the GQ compared to their 8-benzoheteroaryl-dG counterparts.

Table 1 UV-thermal melting parameters for mTBA with *syn*-G₅ 8-heteroaryl-dG probes

Modification	T_m^a (ΔT_m) ^b _{Dup}	T_m^a (ΔT_m) ^c _{GQ}	$\Delta\Delta T_m^d$
Native	64.5	53.5	-11.0
^{Pyr} dG	58.0 (-6.5)	64.5 (11.0)	6.0
^{Fur} dG	58.0 (-6.5)	62.5 (9.0)	4.5
Th dG	56.5 (-8.0)	60.5 (7.0)	4.0
^{Ind} dG	56.5 (-8.0)	62.5 (9.0)	6.0
^{Bfur} dG	56.5 (-8.0)	61.5 (8.0)	5.0
^{BTh} dG	52.0 (-12.5)	55.0 (1.5)	3.0

^a T_m values in °C were determined from solutions of 6 μ M oligonucleotide in 100 mM M⁺ phosphate buffer pH 7.0 with 0.1 M MCl (M = Na⁺, duplex; M = K⁺, GQ). Samples were monitored at 260 nm (duplex) or 295 nm (GQ) over 5 ramps at a rate of 0.5 °C/min and are reproducible within 3%. ^b $\Delta T_m = T_m$ (mTBA duplex) - T_m (native TBA duplex). ^c $\Delta T_m = T_m$ (mTBA GQ) - T_m (native TBA GQ). All 8-heteroaryl-dG modifications of the mTBA samples are present at the *syn*-G₅ position. ^d $\Delta\Delta T_m = T_{mGQ} - T_{mDup}$.

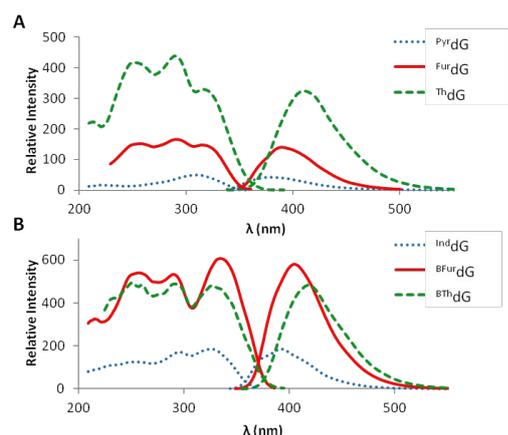


Fig. 2 Fluorescence excitation and emission spectra of mTBA GQ for **A** 5-membered ring series, Pyr dG (blue dotted), Fur dG (red solid), and Th dG (green dashed) and **B** benzoheteroaryl series, Ind dG (blue dotted), BFur dG (red solid), and BTh dG (green dashed). All spectra of mTBA oligonucleotides (6 μM) were recorded in 100 mM potassium phosphate buffer, pH 7, with 0.1 M KCl at 10 $^{\circ}\text{C}$.

Circular dichroism (CD) was utilized to confirm the duplex and antiparallel GQ structures of the mTBA samples. Typical duplex CD spectra were obtained with negative peaks at 240 nm and positive peaks at 260 nm. The GQ CD spectra confirmed an antiparallel structure with positive peaks at 245 and 290 nm and negative peaks at 260 nm (Fig. S1, ESI).

The emission and excitation spectra of the mTBA GQ structures highlight the relative emission wavelengths and intensities of the various probes (Fig. 2). For the 5-membered ring derivatives (Fig. 2A), the GQ of mTBA containing Th dG was the most emissive (dashed green trace), exhibiting an almost 3-fold increase in intensity compared to the mTBA GQ containing Fur dG (solid red trace, Fig. 2A). The excitation spectrum of the mTBA GQ containing the Th dG probe exhibited an energy transfer band at ~ 290 nm as a diagnostic feature of GQ formation.^{18,22,23} The Th dG probe also exhibited the greatest Stokes shift ($\Delta\lambda = 95$ nm, Table 2) with the most red-shifted emission wavelength ($\lambda_{\text{em}} = 411$ nm, Table 2) of the 5-membered derivatives. In contrast, the pyrrolyl derivative Pyr dG exhibited quenched emission in the GQ structure (dotted blue trace, Fig. 2A) with a weak energy transfer band in the excitation spectrum. All 5-membered probes displayed quenched emission in the duplex structures and lacked energy transfer bands in the excitation spectra (Fig. S2, ESI). Compared to the emission intensity of the probes in the GQ structures, the Th dG probe displayed a 10.8-fold increase in emission intensity (I_{rel} values, Table 2), compared to 6.2-fold for Fur dG and 1.7-fold for Pyr dG .

The benzoheteroaryl series of probes tended to be brighter in the GQ structure than the 5-membered ring counterparts and exhibited distinct 290 nm charge transfer bands in their excitation spectra (Fig. 2B). For this series, the most emissive derivative was BFur dG (red solid trace), although BTh dG showed similar emission intensity (dashed green trace) along with a more red-shifted emission wavelength (418 nm for BTh dG vs. 405 nm for BFur dG , Table 2). The Ind dG derivative exhibited quenched emission compared to BFur dG and BTh dG (dotted blue trace, Fig. 2B) and a blue-shifted λ_{em} (392 nm, Table 2). The

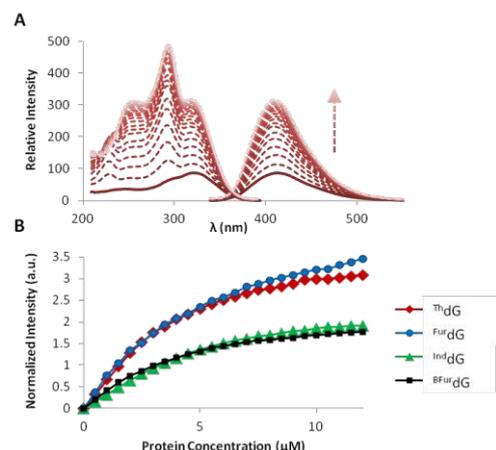


Fig. 3 Fluorescence titrations (6 μM mTBA) carried out with thrombin, (A) Th dG@5 , duplex \rightarrow GQ, initial trace of duplex depicted by solid line, while dashed traces depict GQ formation upon successive addition of thrombin. (B) Normalized fluorescence responses (initial fluorescence of the duplex set to $y = 0$) for the protein binding titrations of mTBAs@5 for Fur dG (blue circle), Th dG (red diamond), Ind dG (green triangle) and BFur dG (black square).

benzoheteroaryl probes were also more emissive than their 5-membered ring counterparts in the duplex structure (Fig. S2). Consequently, they exhibited smaller I_{rel} values than Fur dG and Th dG (4.8 for BFur dG , 3.0 for Ind dG and 1.3 for BTh dG , Table 2).

To estimate the thrombin binding affinity of the mTBA oligonucleotides, thrombin-aptamer fluorescence titrations were carried out (6 μM [mTBA]) in 100 mM potassium phosphate buffer pH 7.0 with 0.1 M KCl at 37 $^{\circ}\text{C}$. The mTBA samples were initially annealed to 1.1 equivalents of the complementary strand to produce the duplex and quench the emission of the 8-heteroaryl-dG probe. Addition of thrombin promoted GQ formation, which triggered an increase in emission intensity of the internal 8-heteroaryl-dG probe. Titrations were carried out with the four derivatives (Fur dG , Th dG , Ind dG and BFur dG) that provided the greatest I_{rel} values (Table 2). A representative titration for the mTBA sample containing Th dG at syn-G_5 is presented in Fig. 3A. Thrombin titrations of the other three derivatives are provided in Fig. S3 (ESI). Plots of the normalized fluorescent intensity (initial fluorescence of the duplex was set to $y = 0$, see Fig. 3B) versus [protein] (Fig. 3B) indicated a 1:1 mTBA/thrombin interaction and provided dissociation constants (K_d) ranging from 4.4-6.8 μM (Table 2). For native TBA binding in K^+ -solution at 37 $^{\circ}\text{C}$, a 5'-fluorescein (FAM)-labeled single-strand TBA sample was monitored for thrombin binding affinity using fluorescence polarization (FP),^{11,32} which provided a K_d value of 4.9 μM (Table 2). This value is around one order of magnitude larger than the K_d value for thrombin binding by unlabeled TBA at 25 $^{\circ}\text{C}$ using isothermal titration calorimetry.³³ Taking into account differences in methodology including instrumentation, temperature and protein origin, our titration data indicates that the internal 8-heteroaryl-dG probes do not perturb thrombin binding compared to the 5'-FAM-labeled TBA sample. It was also noted that the ability of the internal probe to enhance GQ stability (Table 1) did not provide enhanced binding affinity to the protein.

Table 2 Photophysical parameters of *syn*-G₅-modified mTBA and dissociation constants for thrombin binding

Mod.	λ_{exGQ} (nm)	λ_{emGQ} (nm)	$\Delta\nu_{\text{GQ}}$ (nm)	I_{rel}^a	K_d (μM) ^b
Native	NA	NA	NA	NA	(4.9 ± 0.1) ^c
^{Pyr} dG	310	379	69	1.7	n.d.
^{Fur} dG	315	378	63	6.2	5.7 ± 0.1 ^d
Th dG	316	411	95	10.8	4.6 ± 0.2 ^d
^{Ind} dG	325	392	67	3.0	6.8 ± 0.4 ^d
^{BFur} dG	335	405	70	4.8	4.4 ± 0.1 ^d
^{BTh} dG	330	418	88	1.3	n.d.

^a I_{rel} is the emission intensity (I) of $I_{\text{GQ}}/I_{\text{Dup}}$. ^b R^2 values for thrombin titrations were > 0.97. ^c Dissociation constant for 5'-FAM-labeled TBA determined via fluorescence polarization (FP). ^d Dissociation constants for internal *syn*-G₅ probes by fluorescence emission titrations upon duplex→GQ exchange.

Positional impact of ThdG. Given that the ThdG probe exhibited the greatest light-up emission upon duplex→GQ exchange (I_{rel} values Table 2) and had minimal effect on thrombin binding, it was selected for further studies to gauge the positional impact of an 8-heteroaryl-dG probe within alternate *syn*-G-tetrad positions of TBA. The ΔT_m values of mTBA duplex and GQ with ThdG incorporated into positions G₁, G₅, G₁₀, and G₁₄ are presented in Table 3. Incorporation of ThdG within the different *syn*-positions indicated that probe location had an effect on overall tertiary stability. As expected, the duplex structure was significantly destabilized by incorporating ThdG at internal G₅ or G₁₀ positions, with ΔT_m values of -8.0 °C and -9.0 °C, respectively. When located at G₁ or G₁₄, which are at or near the end of the duplex, the stability was less affected (ΔT_m values of -3.0 °C and -3.5 °C). In terms of fluorescence response, ThdG at G₁₀ provided the greatest I_{rel} value of 21.6, compared to 10.8 for G₅, 5.6 for G₁₄ and 3.1 for G₁. The G₁ modification at the end of the DNA strand had the lowest I_{rel} value due to its highly emissive nature in the duplex likely from decreased π -stacking in the helix (6-fold increase in emission for ThdG at G₁ compared to G₅ and G₁₀). For ThdG at the two internal duplex positions (G₅ and G₁₀), the mTBA samples exhibited similar quenched emission in the duplex, but the probe was ~ 2-fold more emissive in the GQ at G₁₀ versus G₅. In fact, the ThdG probe at G₅ within the GQ was the least emissive (I_{GQ} , Table 3).

In an effort to provide an explanation for the quenching of ThdG emission at G₅ relative to the other *syn*-positions, geometric features of mTBA GQ structures were examined using molecular dynamics (MD) simulations (Figure 4, ThdG probe in yellow; see also Table S2 (ESI) for relative free energies of ThdG mTBA GQ structures). The hydrogen-bonding and stacking interactions of ThdG at G₁ and G₁₄ were similar to

Table 3 Thermal melting parameters, photophysical properties and dissociation constants for ThdG-mTBA oligonucleotides

Site(s)	ΔT_m^a (°C) _{Dup}	I_{Dup}	ΔT_m^a (°C) _{GQ}	I_{GQ}	I_{rel}^b	K_d (FPK _d) ^c (μM)
1	-3.0	181	8.0	564	3.1	5.8 ± 0.1
5	-8.0	30.1	7.0	324	10.8	4.6 ± 0.2
10	-9.0	31.5	10.5	681	21.6	5.4 ± 0.2 (5.3 ± 1.4)
14	-3.5	127	8.0	707	5.6	4.9 ± 0.1
1, 10	-11.0	207	18.5	1362	6.6	5.9 ± 0.6 (8.3 ± 1.1)
5, 10	-18.5	76.5	18.5	734	9.6	8.0 ± 1.3 (5.9 ± 0.7)
1, 5, 10, 14	n.d.	n.d.	>36.5	822	n.d.	(6.7 ± 0.6)

^a ΔT_m values compared to native TBA duplex (65.0°C) and GQ (53.5°C). ^b I_{rel} is the emission intensity (I) of $I_{\text{GQ}}/I_{\text{Dup}}$. ^c Dissociation constants (K_d values) determined by fluorescence emission titrations upon duplex→GQ exchange or via FP (values in brackets). R^2 values were > 0.98.

native TBA. Although the guanine moiety in ThdG interacts with T₁₃ when the probe is at G₁₄, no additional interactions were observed with the thienyl group at G₁ and G₁₄. At G₁₀, ThdG forms nonspecific interactions with the sugar moieties of residues constituting the TGT loop (highlighted in green, Fig. 4C). However, none of the nucleobase moieties of the TGT loop stack with the thienyl moiety of ThdG. Nevertheless, the ThdG probe at G₁₀ possesses the largest interaction energy with the surrounding residues compared to all other positions (Table S2, ESI), which may explain why the mTBA GQ modified at G₁₀ has the highest T_m value of the mono-substituted mTBA oligonucleotides (Table 3). In contrast to the other *syn*-positions, additional stacking interactions occur between ThdG and T₄, and a nearly T-shaped interaction occurs between the thienyl moiety and T₇ when the probe is at G₅ (T₄ and T₇ in green, Fig. 4B). Thus, the MD simulations demonstrate that the ThdG probe at G₅ is present in a more crowded environment compared to all other probe locations.

The interactions between ThdG at G₅ and the T₄ and T₇ residues may hinder free rotation of the 8-thienyl group, which can quench fluorescence. For example, we previously demonstrated that 8-heteroaryl-dG nucleosides exhibit quenched emission with increased solvent rigidity.³⁰ The ThdG probe has a twisted ground-state structure that in non-viscous water achieves planarity in the excited-state, which accounts for its relatively large Stokes' shift ($\Delta\nu$). With increased viscosity, excited-states and ground-states have similar geometries,³⁴ which can quench fluorescence if the excited-state remains twisted and cannot achieve planarity.

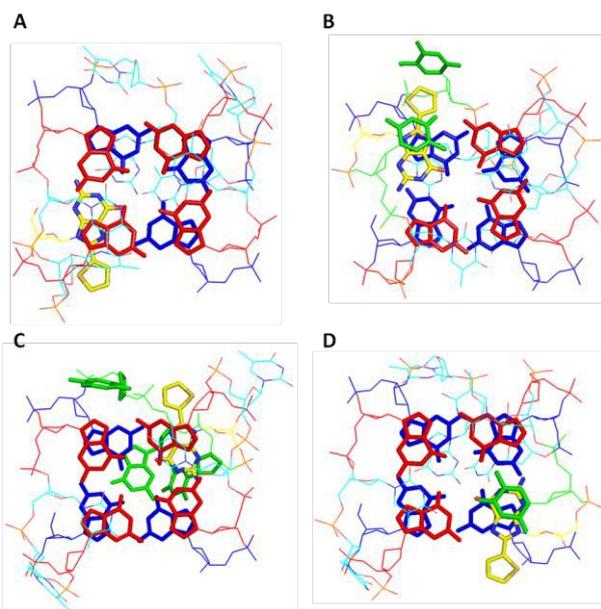


Fig. 4 A portion of the MD representative GQ structures showing the orientation of the ThdG probe (yellow) in positions G_1 (A), G_5 (B), G_{10} (C), and G_{14} (D), interacting with bases in the TGT or TT loops (green) and Gs in the G-tetrad (red and blue).

For multiple probe incorporation into TBA it was desirable to have one ThdG at position G_{10} because it provided the best emissive response to the change in DNA topology and the greatest influence on GQ stability (Table 3). For the disubstituted mTBA samples, internal probe placement ($G_5 + G_{10}$) had a greater destabilizing influence on duplex stability than mTBA with ThdG at $G_1 + G_{10}$. For the tetra-substituted ThdG mTBA oligonucleotide, a reliable T_m value for the duplex could not be established due to the stability of the GQ structure even in the presence of excess complementary strand. In the GQ topology, the ΔT_m values for the disubstituted and tetra-substituted mTBA samples were essentially the sum of the individual ΔT_m 's determined for the monosubstituted samples (i.e. G_1 (8.0) + G_{10} (10.5) = 18.5 °C).

Increasing the number of probes within mTBA also increased the absorbance intensity for the ThdG probe at 320 nm in the UV-vis spectra of the mTBA oligonucleotides in an additive manner (Fig. 5A). The fluorescence emission intensity of the ThdG probes at ~ 415 nm in the duplex structure also increased in an additive manner following excitation at 320 nm (I_{Dup} , Table 3). For example, the emission intensity of the disubstituted mTBA duplex ($G_1 + G_{10}$) was 207, while summing the contributions from the individual ThdG probes is 212.5 (i.e. 181 + 31.5 = 212.5). Within the antiparallel GQ structure, the CD spectra of the tetra-substituted and the di-substituted ($G_5 + G_{10}$) mTBA samples displayed positive induced CD (ICD) bands at 320 nm (Fig. 5B) with the tetra-substituted sample exhibiting the strongest ICD band (black trace, Fig. 5A). Interestingly, the di-substituted mTBA GQ with the probes in the same G-tetrad ($G_1 + G_{10}$, blue trace, Fig. 5B) failed to exhibit an ICD band. Mono-substituted mTBA samples also lack ICD bands in their CD spectra (red trace, Fig. 5B, see also Fig. S1, ESI) suggesting a requirement for at least two ThdG

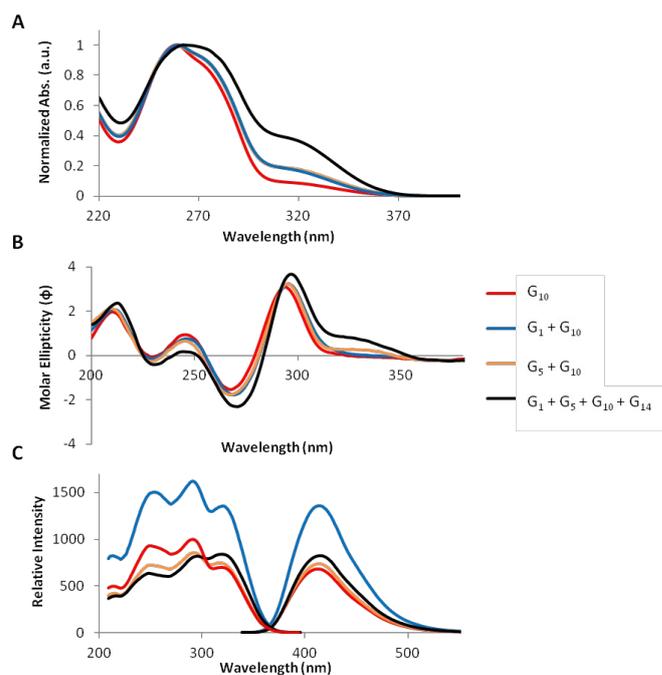


Fig. 5. UV-vis (A), CD (B) and fluorescence (C) spectra of mTBA GQs (6 μ M) with ThdG in position G_{10} (red traces), $G_1 + G_{10}$ (blue traces), $G_5 + G_{10}$ (orange traces) and $G_1 + G_5 + G_{10} + G_{14}$ (black traces).

probes present in separate G-tetrads in order to observe a positive ICD band.

In terms of fluorescence response, the di-substituted ($G_1 + G_{10}$) mTBA GQ (blue trace, Fig. 5C) with the probes in the same G-tetrad was the most emissive. In fact, the emission intensity of the di-substituted derivative ($I_{GQ} = 1362$, Table 3) was slightly greater than the sum of the individual contributions from $G_1 + G_{10}$ (i.e. 564 + 681 = 1245). In contrast, the di-substituted mTBA GQ ($G_5 + G_{10}$, orange trace, Fig. 5C) and especially the tetra-substituted mTBA GQ (black trace, Fig. 5C) exhibited quenched emission compared to what would be expected from summing the intensities from the individual probes (i.e. calculated for tetra-substituted mTBA GQ; $I_{GQ} = 2276$, observed $I_{GQ} = 822$). One factor that could play a contributing role in the observed emission intensities is the efficiency of energy-transfer within the GQ structures. We previously utilized the ratio of excitation intensity at 290 nm (charge transfer band) versus the intensity at 320 nm (direct probe excitation) as an indicator for charge-transfer efficiency for the ^{Fur}dG probe in a GQ.¹⁷ For the mono-substituted G_{10} -mTBA GQ (red trace, Fig. 5C) the I_{290}/I_{320} ratio is 1.42, which drops to 1.21 for the di-substituted ($G_1 + G_{10}$) GQ with the probes in the same G-tetrad. However, in the di-substituted ($G_5 + G_{10}$) mTBA GQ with the probes in stacked G-tetrads the ratio drops to 1.15, while in the tetra-substituted analog the ratio further drops to 0.97. These results suggested that the energy-transfer efficiency decreases when 8-heteroaryl-dG probes are placed in separate G-tetrads that undergo π -stacking. However, di-incorporation of 8-heteroaryl-dG probes in the same G-tetrad of antiparallel GQ structures can enhance emission intensity for fluorescence detection of target molecules.

ARTICLE

Journal Name

Thrombin binding titrations of the ThdG-mTBA oligonucleotides (see Fig. S4, ESI for representative titrations) were carried out at 37 °C in K⁺-solution using both duplex→GQ exchange and FP (see K_d values, Table 3). Site-specific studies indicated that the *syn*-location of the ThdG modification in mTBA does not greatly affect the aptamers ability to bind to its molecular target, thrombin (at least compared to binding by 5'-FAM-labeled TBA). Furthermore, multiple incorporations did not strongly hinder protein binding affinity. Overall, the results established that 8-heteroaryl-dG probes can be placed in multiple *syn*-G-tetrad positions of TBA with minimal impact on protein binding affinity. Interestingly, the ability of the ThdG modification to greatly increase GQ stability at *syn*-G positions did not impact the affinity of thrombin for the mTBA oligonucleotide. This observation may be due to the fact that thrombin interacts with the loops of TBA, not the bases in the G-tetrad.²⁶

Rates of duplex→GQ exchange. For native TBA, Mendoza and coworkers recently determined rates of GQ→duplex by adding FAM-labeled complementary strand to the preformed GQ.²⁸ Duplex production quenched the FAM emission, permitting the rate of duplex formation to be monitored as function of time. At 37 °C in 100 mM K⁺-solution, the unfolding of the GQ structure took place rapidly with a half-life of ~ 5 min (based on predictions from normalized emission intensity traces for TBA²⁸). For the mono- and di-substituted ThdG-mTBA oligonucleotides, rates of GQ formation as a function of time could be monitored by measuring the increase in ThdG fluorescence intensity driven by addition of thrombin to the preformed duplex in K⁺-solution (100 mM) at 37 °C. For these experiments, the mTBA oligonucleotides (2 μM) were annealed to 1.1 equiv. complementary strand and samples were monitored at λ_{ex} 290 nm (the GQ energy transfer band) and λ_{em} 415 nm. A baseline reading was obtained over a 30 second period, followed by manual injection of 2 equiv. of thrombin protein with less than 2 seconds of dead time before collection. Rate experiments with the tetra-substituted mTBA sample were not carried out due to the inability to produce the duplex in K⁺-solution given the enhanced stability of the GQ structure (Table 3).

Fluorescence emission traces as a function of time (Fig. 6) showed the anticipated increase in emission intensity at 415 nm, due to production of the GQ structure required for thrombin binding. Apparent first-order rate constants (k_{obs} , min⁻¹) and half-lives ($t_{1/2}$, min) for GQ formation (i.e. thrombin binding) are given in Table 4. For the mono-substituted derivatives, the rate of GQ formation was particularly slow for the G₁-mTBA oligonucleotide, while placement of the probe at G₁₀ provided the fastest rate of GQ formation ($t_{1/2}$ = 8.36 min). Significant increases in reaction rate were observed for the di-substituted mTBA samples, especially for the G₅ + G₁₀ derivative. The thrombin binding rates could be correlated with the difference in T_m for the duplex and GQ structures (ΔT_m values, Table 4.) These experiments demonstrate a real advantage for the di-substituted mTBA oligonucleotides, as increased GQ emission intensities (IGQ, Table 3) coupled with

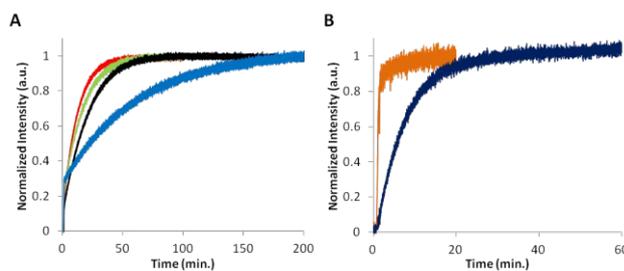


Fig. 6 Normalized emission intensity traces upon GQ formation driven by thrombin binding to ThdG-mTBA for (A) mono-substituted derivatives, position G₁ (blue), G₅ (green), G₁₀ (red), and G₁₄ (black) and (B) di-substituted derivatives, positions G₁ + G₁₀ (dark blue), and G₅ + G₁₀ (orange).

Table 4. Change in duplex and GQ T_m values and calculated reaction rate constants and half-lives for GQ formation by mTBA duplexes driven by thrombin addition

Site(s)	ΔT_m^a (Dup-Quad)	k_{obs} (min ⁻¹) ^b	$t_{1/2}$ (min) ^c
Native	11.0	-	-
1	0.5	0.016 ± 0.001	42.3
5	-4.0	0.070 ± 0.010	9.95
10	-8.0	0.083 ± 0.006	8.36
14	0.0	0.046 ± 0.008	15.1
1+10	-18.0	0.136 ± 0.009	5.10
5+10	-26.0	0.794 ± 0.060	0.87

^a ΔT_m values of 6 μM TBA duplex in Na⁺-solution – 6 μM TBA GQ in K⁺-solution. ^b Determined in 100 mM potassium phosphate buffer pH 7, μ = 0.1 M KCl, at 37 °C starting from 2 μM mTBA duplex. Reaction initiated by addition of 4 μM thrombin. Apparent rate constants were averaged from two independent repetitions. ^c $t_{1/2}$ = 0.693/ k_{obs} .

increased rates of protein binding can be vital for diagnostic applications.

Conclusions

The current work has uncovered key information that permits optimization of probe performance for monitoring target binding by aptamers. Specifically, among the series of 8-heteroaryl-dG probes consisting of (furyl (^{Fur}dG), pyrrolyl (^{Pyr}dG), thienyl (ThdG), benzofuryl (^{Bfur}dG), benzothienyl (^{BTh}dG) and indolyl (^{Ind}dG), the 8-thienyl-dG derivative (ThdG) possesses the best optical properties for monitoring duplex→GQ formation, as determined at the *syn*-G₅ position of the thrombin binding aptamer (TBA). The probe exhibits quenched emission in the duplex that shows a 10.2-fold increase in emission intensity upon GQ formation and does not perturb thrombin binding to the mTBA oligonucleotide. Of the four *syn*-G-tetrad positions of TBA, placement of the ThdG probe at G₁₀ produces the most stable GQ structure and provides the greatest increase in emission intensity (I_{rel} = 21.6) compared to the duplex emission. The ThdG modification minimally perturbs thrombin binding affinity to the mTBA sample regardless of ThdG probe *syn*-G-tetrad position and number of ThdG probes incorporated. Although increasing the number of ThdG probes within *syn*-G positions of mTBA

increases the duplex emission intensity in an additive manner (intensity \approx summation of intensities from the individual ThdG probes), the best emission response occurs with two ThdG probes incorporated into the same G-tetrad in the GQ structure, while placement in separate stacking G-tetrads quenches emission due to diminished charge-transfer efficiency. Finally, increasing the number of ThdG probes within TBA dramatically increases the rate of duplex \rightarrow GQ exchange driven by thrombin binding, which has important implications for diagnostics through aptamer detection strategies. These findings highlight the utility of ThdG as an internal base modification with excellent fluorescent switching properties, which can now be applied towards aptasensor development of other antiparallel G-quadruplex containing systems.

Experimental Section

Materials and Methods

Boronic acids, Pd(OAc)₂, 3,3',3''-phosphinidynetris-(benzenesulfonic acid) trisodium salt (TPPTS), and other commercial compounds were used as received. NMR spectra were recorded on Bruker 300 or 400 MHz spectrometers in either DMSO-d₆ or CDCl₃ referenced to the respective solvent. All UV-Vis and fluorescence spectra were recorded with baseline correction, stirring, and temperature control in 10 mm light path quartz cells. Water for buffers or spectroscopic solutions was of high purity obtained from a filtration system (18.2 M Ω). High-resolution mass spectra were recorded on an Agilent Q-ToF instrument using electrospray ionization.

Oligonucleotide synthesis

Oligonucleotide synthesis of the 8-heteroaryl-dG mTBA samples were carried out on a 1 μ mol scale on a BioAutomation Corp. MerMade 12 automatic DNA synthesizer. The oligonucleotides were cleaved from the solid support and deprotected using 1 mL of 30% ammonia hydroxide solution at 55 $^{\circ}$ C for 6 hours and purified by reverse phased HPLC. Oligonucleotides were quantified using extinction coefficients obtained from <http://www.idtdna.com/analyzer/applications/oligoanalyzer> with mTBA assumed to have the same extinction coefficient as native TBA. Samples were then resuspended in Milli-Q water and purified using an Agilent HPLC instrument equipped with an autosampler, a diode array detector, fluorescence detector and autocollector, as previously outlined in detail.³⁵ MS experiments for identification of the mTBA oligonucleotides were conducted on a Bruker amaZon quadrupole ion trap SL spectrometer in the negative electrospray ionization mode (see ESI for MS spectra). mTBA samples were prepared in 90% Milli-Q filtered water/10% methanol containing 0.1 mM ammonium acetate.

Thermal denaturation and CD Studies

All melting temperatures (T_m 's) of mTBA samples were measured on a Cary 300-Bio UV-Vis spectrophotometer at a concentration of 6.0 μ M mTBA in 100 mM M⁺-phosphate buffer pH 7.0 with 0.1 M M⁺Cl where M⁺ = K⁺ or Na⁺. Duplex

studies were prepared with 1.1 equivalents of complementary strand purchased from Sigma-Aldrich and used without further purification. The UV absorption was monitored as a function of temperature at either 295 nm for GQ, or 260 nm for duplex, and consisted of forward-reverse scans from 10 to 90 $^{\circ}$ C at a heating rate of 0.5 $^{\circ}$ C/min, and was repeated at least three times. The T_m values were calculated by determining the first derivative of the melting curve through the Varian Thermal software. CD spectra were performed on a Jasco J-815 CD spectrophotometer equipped with a thermal controlled 1 x 6 multicell block. The annealed samples obtained from the T_m studies were measured at 10 $^{\circ}$ C in quartz cells (110-QS) with a light path of 1 mm and monitored between 200 and 400 nm at a bandwidth of 1 nm and scanning speed of 100 nm/min. All samples were scanned a minimum of 6 times and corrected for background.

Fluorescence measurements and protein titrations

The fluorescence of the annealed mTBA samples were measured on a Cary Eclipse Fluorescence spectrophotometer as both excitation and emission spectra in quartz cells (108.002F-QS) with a path length of 10 mm at 10 $^{\circ}$ C. Excitation and emission slit widths were kept constant at 5 nm, with the exception of the multiply-incorporated ThdG strands, which were obtained with 2.5 nm excitation and 5 nm emission slit widths. Protein titrations were performed on duplex samples at a concentration of 6.0 μ M in 100 mM potassium phosphate buffer pH 7.0 with 0.1 M KCl with 1.1 equivalents of complementary strand. A Hellma Analytics 119.004F-QS quartz cell with pathlength of 10 x 2 mm was used with stirring and temperature controlled at 37 $^{\circ}$ C. Titrations proceeded according to a previously published protocol²³ with additions of 5 μ L of a 100 μ M thrombin protein solution in 100 mM sodium phosphate buffer pH 7.0 with 0.1 M NaCl. Scans were taken 10 minutes after addition of the protein, until a final concentration of two equivalents of protein had been added. A plot of the fraction of oligonucleotide bound versus [thrombin] generated a binding isotherm that was analyzed with SigmaPlot 13.0 to obtain K_d values.

Fluorescence polarization

Samples of single-stranded mTBA (6 μ M) were prepared in 100 mM potassium phosphate buffer pH 7.0 with 0.1 M KCl in a 1 mL quartz cell (114F-QS). Solutions were run on a PTI QuantaMaster Fluorimeter System at 37 $^{\circ}$ C and measured at the λ_{ex} and λ_{em} of each respective modification. Time-based polarization method was used and the anisotropy was averaged over a period of 30 seconds using Felix 32 software. Protein titrations were carried out the same as the fluorescence protein titrations, 24 additions of 5 μ L of a 100 μ M thrombin protein solution to reach 2 equivalents, with only 5 minutes of manual mix time required after each addition.

Kinetic measurements

Solutions of mTBA annealed duplex (2 μ M) were prepared in a Hellma Analytics 119.004F-QS quartz cell with pathlength of 10 x 2 mm with stirring and temperature controlled at 37 $^{\circ}$ C.

Samples were monitored at λ_{ex} 290 nm, the energy transfer band, and λ_{em} 415 nm. A baseline reading was obtained over a period of 30 seconds, followed by manual injection of 2 equivalents of thrombin prepared as a 200 μM stock in 100 mM sodium phosphate buffer pH 7.0 with 0.1 M NaCl with 2 seconds or less of dead time before collection. Kinetics were calculated with first order rate equation using the Marquardt non-linear regression analysis on the Cary Eclipse Kinetics Software V 1.1(133). The kinetic runs were performed twice and data was averaged.

Computational Details

Following the approach used in a previous study of modified TBA (mTBA),²² the NMR-based structure of the non-modified 15mer thrombin binding aptamer (PDB code: 148D, eighth frame) was used as a starting structure for molecular dynamics (MD) simulations. Structures of TBA containing ThdG at the *syn*-G₁, G₅, G₁₀ or G₁₄ position were built by carrying out the appropriate C8 modification using Gaussview (Version 5.0).³⁶ Initial structures were built with the *syn*-conformations of ThdG with two different orientations about the linkage between dG and the thienyl moiety (defined as the dihedral angle (θ) about the thienyl-dG bond with respect to N9 and S). However, the conformations with $\theta \sim 0^\circ$ is more favourable than $\theta \sim 180^\circ$ (Table S2), and therefore only the structure with $\theta \sim 0^\circ$ are discussed in the main text. The parmbsc0³⁷ modification to the parm99³⁸ force field was used for simulating the natural nucleosides (dA, dC, dG and dG), while GAFF³⁹ parameters were used for ThdG. Partial charges for ThdG were calculated using the RED v.III.8⁴⁰ program interfaced with Gaussian 09.

The DNA was immersed in a 15 Å octahedral box of TIP3P water molecules using the TLEAP module of AMBER 11.⁴¹ Since previous studies suggest that a single potassium ion bonds the 15mer TBA,⁴² the folded TBA was stabilized by placing a potassium cation in the geometrical center of the GQ stem. In synchrony with previous MD simulations on mTBA,²² the initial ion coordinates were determined by averaging the Cartesian coordinates of the O6 atoms of the guanines forming the GQ stem (i.e., G1, G2, G5, G6, G10, G11 and G15). The system was further neutralized by placing 13 sodium ions at random positions inside the water box, in accordance with the protocol implemented in the TLEAP module of AMBER11.⁴¹

Minimization of the solvent and ions was initially performed for 500 steps using the steepest descent algorithm followed by another 500 steps using the conjugate gradient minimization algorithm, with the DNA held fixed using a force constant of 500 kcal mol⁻¹ Å⁻². Subsequently, 2500 steps of minimization (1000 steps of steepest descent followed by 1500 steps of conjugate gradient) were performed on the entire system with all constraints on DNA removed. The system was subsequently heated from 0 to 300 K with the DNA restrained using a force constant of 10 kcal mol⁻¹ Å⁻² for 20 ps under constant volume. Finally, 20 ns of unrestrained MD simulations were carried out for each system at constant temperature (300 K) and pressure (1 atm) using AMBER 12.⁴³ The bonds involving hydrogen atoms were constrained using SHAKE and a 2 fs time step was used throughout the simulation.

The stability of the MD simulations was confirmed through analysis of root mean square derivations (RMSD) in the backbone residues over the course of the simulation, which indicated that no large fluctuations in the structures were observed over the MD trajectory. A free energy analysis was carried out using representative snapshots from the MD trajectories after removing the water and ions. Specifically, the molecular mechanics energy (which is the sum of the energy terms stemming from deviations in the bond lengths, bond angles and dihedral angles from their equilibrium values, as well as the van der Waals interaction potential and the electrostatic energy contributions) and the solvation free energy (which is the sum of the electrostatic solvation energy and the nonpolar contribution to the solvation free energy) were calculated using the Poisson-Boltzmann method, while the entropy term was estimated from normal mode calculations using the molecular mechanics Poisson-Boltzmann surface area calculations (MM-PBSA)⁴⁴ method as implemented in AMBER 11 and 12. The molecular mechanics and solvation free energy terms were calculated using 1000 snapshots (i.e., 1 snapshot every 20 ps), while the entropy term was calculated using 100 snapshots (i.e., 1 snapshot every 200 ps).

Phosphoramidite synthesis

General procedures for the three steps in the phosphoramidite synthesis (Scheme 1) have been previously detailed.⁴⁵ For compound numbering see Scheme 1.

N²-(Dimethylformamidyl)-8-(2''-pyrrolyl)-2'-dG 1a. Afforded 0.28 g of a grey powder (59% yield). ¹H NMR (400 MHz, DMSO-d₆) δ = 11.67 (s, 1H), 11.45 (s, 1H), 8.47 (s, 1H), 6.92 (s, 1H), 6.49 (s, 1H), 6.40 (t, J=7.2 Hz, 1H), 6.19 (s, 1H), 5.28 (s, 1H), 4.96-4.93 (m, 1H), 4.47 (bs, 1H), 3.83 (bs, 1H), 3.66 (m, 1H), 3.54 (m, 1H), 3.24 (m, 1H), 3.13 (s, 3H), 3.03 (s, 3H), 2.13-2.11 (m, 1H). ¹³C NMR (400 MHz, DMSO-d₆) δ = 158.5, 157.8, 156.9, 150.8, 142.9, 121.4, 121.0, 120.4, 110.4, 109.4, 88.0, 85.0, 71.5, 62.3, 41.3, 37.5, 35.0. HRMS Calcd for C₁₇H₂₀N₇O₄⁻ [M-H⁻]:386.1582; found 386.1584.

N²-(Dimethylformamidyl)-8-(2''-thienyl)-2'-dG 2a. Synthesis afforded 1.02 g of a white powder (63% yield). ¹H NMR (400 MHz, DMSO-d₆) δ = 11.48 (bs, 1H), 8.49 (s, 1H), 7.76 (s, 1H), 7.49 (s, 1H), 7.21 (s, 1H), 6.30 (t, J=7.2 Hz, 1H), 5.27 (bs, 1H), 4.86 (bs, 1H), 4.47 (bs, 1H), 3.82 (bs, 1H), 3.64 (m, 1H), 3.53 (m, 1H), 3.31 (m, 1H), 3.14 (s, 3H), 3.04 (s, 3H), 2.15-2.11 (m, 1H). ¹³C NMR (400 MHz, DMSO-d₆) δ = 158.1, 157.2, 156.8, 150.7, 142.2, 131.7, 128.8, 128.2, 127.9, 120.1, 87.6, 84.5, 70.8, 61.8, 40.8, 36.8, 34.6. HRMS Calcd for C₁₇H₂₁N₆O₄S⁺ [M+H⁺]:405.1340; found 405.1341.

N²-(Dimethylformamidyl)-8-(2''-indolyl)-2'-dG 3a. Synthesis generated 1.14 g of a beige powder (74% yield). ¹H NMR (400 MHz, DMSO-d₆) δ = 11.84 (s, 1H), 11.53 (s, 1H), 8.51 (s, 1H), 7.63 (d, J=7.6 Hz, 1H), 7.43 (d, J=8.0 Hz, 1H), 7.15 (m, 1H), 7.02 (m, 1H), 6.89 (s, 1H), 6.54 (m, 1H), 5.32 (d, J=4.0 Hz, 1H), 4.96 (t, J= 5.4 Hz, 1H), 4.51 (bs, 1H), 3.87 (m, 1H), 3.67 (m, 1H), 3.56 (m, 1H), 3.30 (m, 1H), 3.15 (s, 3H), 3.04 (s, 3H), 2.17 (m, 1H). ¹³C NMR (400 MHz, DMSO-d₆) δ = 158.7, 157.9, 157.3, 151.3, 142.2, 137.1, 128.2, 127.3, 123.3, 121.3, 120.7, 112.3, 103.4,

88.4, 85.3, 71.4, 62.3, 41.6, 37.6, 35.1. HRMS Calcd for $C_{21}H_{24}N_7O_4^+$ [M+H⁺]: 438.1834; found 438.1882.

***N*²-(Dimethylformamidy)-8-(2''-benzofuryl)-2'-dG 4a.**

Afforded 0.97 g of a beige powder (65% yield). ¹H NMR (300 MHz, DMSO-*d*₆) δ = 11.56 (s, 1H), 8.53 (s, 1H), 7.78 (d, J=7.5 Hz, 1H), 7.69 (d, J=8.4 Hz, 1H), 7.46 (s, 1H), 7.41-7.38 (m, 1H), 7.35-7.32 (m, 1H), 6.56 (t, J=7.2 Hz, 1H), 5.28 (d, J=4.5 Hz, 1H), 4.86 (s, 1H), 4.51 (bs, 1H), 3.84 (m, 1H), 3.66 (m, 1H), 3.55 (m, 1H), 3.23 (m, 1H), 3.16 (s, 3H), 3.05 (s, 3H), 2.23-2.20 (m, 1H). ¹³C NMR (400 MHz, DMSO-*d*₆) δ = 158.2, 157.3, 154.2, 150.7, 145.7, 138.3, 127.5, 125.7, 123.6, 121.8, 120.7, 111.4, 108.2, 87.6, 84.4, 70.7, 61.7, 40.8, 37.5, 34.6. HRMS Calcd for $C_{21}H_{23}N_6O_5^+$ [M+H⁺]: 439.1724; found 439.1722.

5'-O-(4,4'-Dimethoxytrityl)-*N*²-(dimethylformamidy)-8-(2''-pyrrolyl)-2'-dG 1b. Synthesis afforded product as a pale yellow powder (0.43 g, 89%). ¹H NMR (300 MHz, DMSO-*d*₆) δ = 11.67 (bs, 1H), 11.40 (bs, 1H), 8.33 (s, 1H), 7.24 (m, 2H), 7.14-7.11 (m, 7H), 6.94 (m, 1H), 6.75 (m, 1H), 6.68 (dd, J=8.7, 7.5 Hz, 4H), 6.39 (m, 1H), 6.21 (m, 1H), 5.34 (bs, 1H), 4.81 (m, 1H), 3.93 (m, 1H), 3.66 (s, 6H), 3.44 (m, 3H), 3.02 (s, 3H), 3.00 (s, 3H), 2.26 (m, 1H). ¹³C NMR (300 MHz, DMSO-*d*₆) δ = 157.9, 157.8, 157.4, 157.3, 155.9, 150.0, 145.0, 142.9, 135.7, 135.6, 129.8, 127.9, 126.3, 121.0, 119.8, 113.3, 110.3, 108.7, 85.4, 85.0, 84.3, 70.4, 63.6, 55.2, 40.9, 37.2, 34.9. HRMS Calcd for $C_{38}H_{40}N_7O_6^+$ [M+H⁺]: 690.3035; found 690.3045.

5'-O-(4,4'-Dimethoxytrityl)-*N*²-(dimethylformamidy)-8-(2''-thienyl)-2'-dG 2b. Synthesis generated product as a white powder (1.29 g, 74%). ¹H NMR (300 MHz, DMSO-*d*₆) δ = 11.50 (bs, 1H), 8.36 (s, 1H), 7.75 (m, 1H), 7.71 (m, 1H), 7.22 (m, 3H), 7.11-7.09 (m, 7H), 6.68 (dd, J=9.0, 4.8 Hz, 4H), 6.34 (m, 1H), 5.35 (d, J=4.8 Hz, 1H), 4.84 (m, 1H), 3.92 (m, 1H), 3.67 (s, 6H), 3.52-3.47 (m, 2H), 3.13 (m, 1H), 3.03 (s, 6H), 2.29-2.24 (m, 1H). ¹³C NMR (400 MHz, DMSO-*d*₆) δ = 158.3, 158.3, 158.2, 157.8, 156.9, 150.8, 145.42, 143.2, 136.1, 136.1, 133.0, 129.9, 129.3, 128.6, 128.3, 128.1, 128.0, 126.9, 120.5, 113.4, 85.9, 85.5, 84.6, 70.8, 63.7, 41.4, 39.3, 35.2. HRMS Calcd for $C_{38}H_{37}N_6O_6S^+$ [M-H]: 705.2501; found 705.2495.

5'-O-(4,4'-Dimethoxytrityl)-*N*²-(dimethylformamidy)-8-(2''-indolyl)-2'-dG 3b. Synthesis afforded product as an off white powder (0.50 g, 53%). ¹H NMR (400 MHz, DMSO-*d*₆) δ = 11.90 (s, 1H), 11.56 (s, 1H), 8.40 (s, 1H), 7.65 (d, J=8.0 Hz, 1H), 7.45 (d, J=8.0 Hz, 1H), 7.23-7.02 (m, 12H), 6.65-6.63 (m, 4H), 6.55 (m, 1H), 5.40 (d, J=5.2 Hz, 1H), 4.92 (m, 1H), 3.98 (m, 1H), 3.62 (s, 3H), 3.60 (s, 3H), 3.56 (m, 1H), 3.16 (m, 2H), 3.05 (s, 6H), 2.34 (m, 1H). ¹³C NMR (400 MHz, DMSO-*d*₆) δ = 157.8, 157.7, 157.5, 156.4, 150.4, 145.1, 142.8, 136.8, 135.7, 129.5, 128.0, 127.6, 127.2, 126.5, 122.7, 120.8, 120.2, 119.6, 118.1, 111.8, 102.7, 85.5, 85.0, 84.3, 70.5, 63.3, 55.0, 41.1, 37.3, 34.7. HRMS Calcd for $C_{42}H_{42}N_7O_6^+$ [M+H⁺]: 740.3191; found 740.3189.

5'-O-(4,4'-Dimethoxytrityl)-*N*²-(dimethylformamidy)-8-(2''-benzofuryl)-2'-dG 4b. Synthesis afforded product as a beige powder (0.18 g, 43%). ¹H NMR (300 MHz, DMSO-*d*₆) δ = 11.57 (bs, 1H), 8.40 (s, 1H), 7.75 (d, J=7.2 Hz, 1H), 7.63 (d, J=7.5 Hz, 1H), 7.55 (s, 1H), 7.41 (t, J= 8.1 Hz, 1H), 7.35 (t, J= 7.5 Hz, 1H), 7.20 (m, 2H), 7.10 (m, 7H), 6.68 (m, 4H), 6.60 (m, 1H), 5.38 (d, J= 5.1 Hz, 1H), 4.77 (m, 1H), 3.96 (m, 1H), 3.65 (s, 6H), 3.43 (m, 3H), 3.05 (s, 3H), 3.04 (s, 3H), 2.32 (m, 1H). ¹³C NMR (300 MHz,

DMSO-*d*₆) δ = 158.3, 158.3, 158.0, 157.4, 154.7, 150.8, 146.5, 145.4, 139.3, 136.0, 130.0, 130.0, 128.2, 128.2, 127.0, 124.1, 122.4, 121.1, 113.4, 111.8, 108.4, 86.0, 85.5, 84.6, 70.8, 63.9, 55.5, 41.6, 38.2, 35.4. HRMS Calcd for $C_{42}H_{41}N_6O_7^+$ [M+H⁺]: 741.3031; found 741.3037.

3'-O-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT-*N*²-(dimethylformamidy)-8-(2''-pyrrolyl)-2'-dG 1c. The amidite was obtained as yellow-green foam (0.25 g, 62%). ¹H NMR (400 MHz, CDCl₃) δ = 9.99 (bs, 1H), 9.67 (bs, 1H), 8.40, 8.35 (s, 1H), 7.33-7.09 (m, 9H), 6.90-6.88 (m, 2H), 6.67-6.62 (m, 4H), 6.45-6.39 (m, 1H), 6.26-6.24 (m, 1H), 5.25-5.10 (m, 1H), 4.18-4.17 (m, 1H), 3.69 (s, 3H), 3.68 (s, 3H), 3.55-3.51 (m, 3H), 3.38-3.27 (m, 2H), 3.01-2.97 (m, 6H), 2.90 (s, 1H), 2.84 (s, 1H), 2.41-2.38 (m, 1H), 1.14-1.12 (m, 12H). ¹³C NMR (400 MHz, CDCl₃) δ = 158.2, 157.7, 157.5, 155.3, 150.4, 144.7, 143.9, 135.6, 130.0, 127.9, 127.6, 126.6, 121.4, 121.3, 120.9, 120.2, 117.7, 113.0, 110.6, 109.8, 85.9, 84.8, 84.3, 84.0, 74.3, 74.1, 73.3, 73.1, 63.2, 62.6, 58.3, 58.2, 58.1, 55.1, 43.2, 41.2, 36.7, 35.1, 24.5, 24.3, 20.3, 20.2. ³¹P NMR (300 MHz, CDCl₃) δ = 148.7, 148.6. HRMS Calcd for $C_{47}H_{57}N_9O_7P^+$ [M+H⁺]: 890.4113; found 890.4100.

3'-O-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT-*N*²-(dimethylformamidy)-8-(2''-thienyl)-2'-dG 2c. The amidite was obtained as a pale green foam (0.17 g, 43%). ¹H NMR (300 MHz, CDCl₃) δ = 9.06 (bs, 1H), 8.35-8.29 (s, 1H), 7.60 (m, 1H), 7.38-7.36 (m, 1H), 7.31-7.29 (m, 2H), 7.21-7.00 (m, 9H), 6.66-6.60 (m, 4H), 6.30-6.26 (dd, J= 4.2 Hz, 1H), 5.14-5.00 (m, 1H), 4.13-4.11 (m, 1H), 3.67-3.66 (m, 7H), 3.46-3.27 (m, 6H), 2.99 (s, 3H), 2.91 (s, 3H), 2.29 (m, 1H), 1.21-1.17 (m, 4H), 1.11-0.98 (m, 9H). ¹³C NMR (400 MHz, CDCl₃) δ = 158.4, 157.7, 157.5, 155.7, 150.8, 144.8, 144.2, 135.9, 132.2, 130.0, 128.8, 128.4, 128.1, 127.7, 127.5, 126.8, 121.0, 117.6, 113.0, 86.1, 84.9, 84.3, 73.7, 63.6, 58.3, 55.2, 43.4, 43.3, 41.2, 37.0, 35.2, 24.4, 20.2. ³¹P NMR (300 MHz, CDCl₃) δ = 149.0, 148.8, 14.3, 13.8. HRMS Calcd for $C_{47}H_{56}N_8O_7P^+$ [M+H⁺]: 907.3725; found 907.3717.

3'-O-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT-*N*²-(dimethylformamidy)-8-(2''-indolyl)-2'-dG 3c. The amidite product was obtained as a white foam (0.17 g, 27%). ¹H NMR (400 MHz, CDCl₃) δ = 9.57 (bs, 1H), 8.61 (bs, 1H), 8.45-8.37 (s, 1H), 7.61 (m, 1H), 7.38-7.36 (m, 3H), 7.19-7.09 (m, 10H), 6.65-6.56 (m, 5H), 5.12 (m, 1H), 4.27-4.25 (m, 1H), 3.67-3.66 (m, 6H), 3.51-3.50 (m, 3H), 3.46-3.41 (m, 2H), 3.05 (s, 6H), 3.02 (s, 1H), 2.99 (s, 1H), 2.35 (m, 1H), 1.12-1.09 (m, 12H). ¹³C NMR (400 MHz, CDCl₃) δ = 158.4, 157.8, 157.7, 155.6, 150.9, 144.7, 143.4, 136.4, 135.7, 130.1, 128.6, 126.8, 126.7, 123.7, 120.8, 120.2, 117.6, 117.0, 111.3, 103.9, 86.0, 84.6, 84.4, 73.5, 63.3, 62.8, 58.2, 46.0, 43.4, 41.3, 34.5, 24.7, 24.4, 20.2, 20.1. ³¹P NMR (300 MHz, CDCl₃) δ = 148.5, 148.3. HRMS Calcd for $C_{51}H_{59}N_9O_7P^+$ [M+H⁺]: 940.4270; found 940.4259.

3'-O-[(2-Cyanoethoxy)(diisopropylamino)phosphino]-5'-O-DMT-*N*²-(dimethylformamidy)-8-(2''-benzofuryl)-2'-dG 4c. The amidite was obtained as a white foam (0.27 g, 67%). ¹H NMR (400 MHz, CDCl₃) δ = 8.59 (bs, 1H), 8.43-8.38 (s, 1H), 7.63-7.61 (m, 1H), 7.53-7.45 (m, 2H), 7.35-7.33 (m, 3H), 7.23-7.21 (m, 5H), 7.16-7.14 (m, 3H), 6.79-6.77 (m, 3H), 6.72-6.66 (m, 4H), 4.98 (m, 1H), 4.23-4.22 (m, 1H), 3.75-3.73 (m, 6H),

3.56-3.53 (m, 3H), 3.38-3.35 (m, 2H), 3.07-3.06 (m, 4H), 3.00-2.99 (m, 4H), 2.37-2.34 (m, 1H), 1.18-1.04 (m, 12H). ¹³C NMR (400 MHz, DMSO-d₆) δ = 157.9, 157.7, 157.4, 157.1, 154.3, 150.3, 145.9, 144.8, 138.8, 135.4, 129.5, 129.4, 127.9, 127.7, 126.5, 125.8, 123.6, 121.9, 120.7, 118.9, 112.8, 111.3, 108.0, 85.2, 84.1, 73.5, 63.0, 62.5, 58.3, 54.9, 46.3, 40.2, 34.7, 33.2, 24.2, 19.8, 19.1, 18.7. ³¹P NMR (300 MHz, CDCl₃) δ = 148.5, 148.3. HRMS Calcd for C₅₁H₅₈N₈O₈P⁺ [M+H⁺]: 941.4110; found 941.4098.

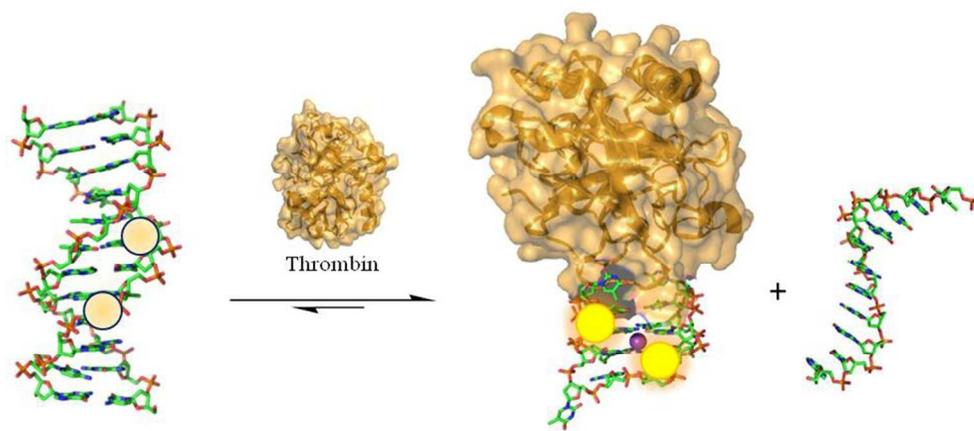
Acknowledgements

We gratefully acknowledge the Natural Sciences and Engineering Research Council of Canada [Discovery 311600-2013 to RAM and 249598-07 to S.D.W.]; Canada Research Chair Program [950-228175 to S.D.W.] and the Canadian Foundation of Innovation [10679 to R.A.M. and 22770 to S.D.W.] for their generous support. Computational resources provided by Westgrid and Compute/Calcul Canada are greatly appreciated. K.L.F and R.A.M also acknowledge Michael Wickens and Adam Gabara for carrying out preliminary experiments.

Notes and references

- 1 D. H. J. Bunka and P. G. Stockley, *Nat. Rev.* 2006, **4**, 588-596.
- 2 A. Ozer, J. M. Pagano and J. T. Lis, *Mol. Ther.- Nucleic Acids*, 2014, **3**, e183.
- 3 A. D. Ellington and J. W. Szostak, *Nature*, 1990, **346**, 818-822.
- 4 C. Tuerk and L. Gold, *Science*, 1990, **249**, 505-510.
- 5 M. Kuwahara, in *Chemical Biology of Nucleic Acids*, RNA Technologies, eds. V. A. Erdmann, W. T. Markiewicz and J. Barciszewski, Springer-Verlag Berlin Heidelberg, 2014, 243-270.
- 6 S. Diafa and M. Hollenstein, *Molecules*, 2015, **20**, 16643-16671.
- 7 E. W. Ng, D. T. Shima, P. Calias, E. T. Cunningham, Jr., D. R. Guyer and A. P. Adamis, *Nat. Rev. Drug Discovery*, 2006, **5**, 123-132.
- 8 I. K. Astakhova and J. Wengel, *Acc. Chem. Res.*, 2014, **47**, 1768-1777.
- 9 J. D. Vaught, C. Bock, J. Carter, T. Fitzwater, M. Otis, D. Schneider, J. Rolando, S. Waugh, S. K. Wilcox and B. E. Eaton, *J. Am. Chem. Soc.*, 2010, **132**, 4141-4151.
- 10 M. Kimoto, R. Yamashige, K. Matsunaga, S. Yokoyama and I. Hirao, *Nat. Biotechnol.*, 2013, **31**, 453-457.
- 11 V. B. Tsvetkov, A. M. Varizhuk, G. E. Pozmogova, I. P. Smirnov, N. A. Kolganova and E. N. Timofeev, 2015, *Sci. Rep.*, **5**, 16337.
- 12 Y. Imaizumi, Y. Kasahara, H. Fujita, S. Kitadume, H. Ozaki, T. Endoh, M. Kuwahara and N. Sugimoto, *J. Am. Chem. Soc.*, 2013, **135**, 9412-9419.
- 13 J. Zhao, S. Katsube, J. Yamamoto, K. Yamasaki, M. Miyagishi and S. Iwai, *Analyst*, 2015, **140**, 5881-5884.
- 14 K. Sefah, Z. Yang, K. M. Bradley, S. Hoshika, E. Jiménez, L. Zhang, G. Zhu, S. Shanker, F. Yu, D. Turek, W. Tan and S. A Benner, *Proc. Natl. Acad. Sci. U.S.A.*, 2014, **111**, 1449-1454.
- 15 A. Dumas and N. W. Luedtke, *J. Am. Chem. Soc.*, 2010, **132**, 18004-18007.
- 16 A. Dumas and N. W. Luedtke, *Nucleic Acids Res.*, 2011, **39**, 6825-6834.
- 17 M. Sproviero, K. L. Fadock, A. A. Witham and R. A. Manderville, *ACS Chem. Biol.*, 2015, **10**, 1311-1318.
- 18 R. A. Manderville and S. D. Wetmore, *Chem. Sci.* 2016, DOI:10.1039/c6sc00053c.
- 19 P. Alberti and J.-L. Mergny, *Proc. Natl. Acad. Sci. U.S.A.*, 2003, **100**, 1569-1573.
- 20 G. W. Collie and G. N. Parkinson, *Chem. Soc. Rev.*, 2011, **40**, 5867-5892.
- 21 W. O. Tucker, K. T. Shum and J. A. Tanner, *Curr. Pharm. Des.*, 2012, **18**, 2014-2026.
- 22 M. Sproviero, K. L. Fadock, A. A. Witham, R. A. Manderville, P. Sharma and S. D. Wetmore, *Chem. Sci.*, 2014, **5**, 788-796.
- 23 M. Sproviero and R. A. Manderville, *Chem. Commun.*, 2014, **50**, 3097-3099.
- 24 D. J. M. Blanchard, T. Z. Cservenyi and R. A. Manderville, *Chem. Commun.* 2015, **51**, 16829-16831.
- 25 B. Deng, Y. Lin, C. Wang, F. Li, Z. Wang and H. Zhang, *Anal. Chim. Acta*, 2014, **837**, 1-15.
- 26 I. R. Krauss, A. Merlino, A. Randazzo, E. Novellino, L. Mazzarella and F. Sica, *Nucleic Acids Res.*, 2012, **40**, 8119-8128.
- 27 P. Schultze, R. F. Macaya and J. Feigon, *J. Mol. Biol.*, 1994, **235**, 1532-1547.
- 28 O. Mendoza, J. Elezgaray and J.-L. Mergny, *Biochimie*, 2015, **118**, 225-233.
- 29 K. M. Schlitt, A. L. Millen, S. D. Wetmore and R. A. Manderville, *Org. Biomol. Chem.*, 2011, **24**, 1694-1709.
- 30 K. M. Rankin, M. Sproviero, K. Rankin, P. Sharma, S. D. Wetmore and R. A. Manderville, *J. Org. Chem.*, 2012, **77**, 10498-10508.
- 31 M. Sproviero, A. M. R. Verwey, A. A. Witham, R. A. Manderville, P. Sharma and S. D. Wetmore, *Chem. Res. Toxicol.*, 2015, **28**, 1647-1658.
- 32 D. M. Jameson and J. A. Ross, *Chem. Rev.*, 2010, **110**, 2685-2708.
- 33 B. Pagano, L. Martino, A. Randazzo and C. Giancola, *Biophys. J.*, 2008, **94**, 562-569.
- 34 S. Sharafy and K. A. Muszkat, *J. Am. Chem. Soc.*, 1971, **93**, 4119-4125.
- 35 D. J. M. Blanchard, K. L. Fadock, M. Sproviero, P. Deore, T. Z. Cservenyi, R. A. Manderville, P. Sharma and S. D. Wetmore, *J. Mater. Chem. C*, 2016, DOI:10.1039/c5tc03354c.
- 36 R. Dennington, T. Keith, J. Millam, GaussView, Version 5, 2009 Semichem Inc., Shawnee Mission, KS.
- 37 A. Perez, I. Marchán, D. Svozil, J. Šponer, T. E. Cheatham III, C. A. Loughton and M. Orozco, *Biophys. J.*, 2007, **92**, 3817-3829.
- 38 T. E. Cheatham III, P. Cieplak and P. A. Kollman, *J. Biomol. Struct. Dyn.*, 1999, **16**, 845-862.
- 39 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157-1174.
- 40 F.-Y. Dupradeau, A. Pigache, T. Zaffran, C. Savineau, R. Lelong, N. Grivel, D. Lelong, W. Rosanski and P. Cieplak, *Phys. Chem. Chem. Phys.*, 2010, **12**, 7821-7839.
- 41 D. A. Case, T. A. Darden, T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. E., R. Luo, M. Crowley, R. C. Walker, W. Zhang, K. M. Merz, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, D. H. Mathews, M. G. Seetin, C. Sagui, V. Babin and P. A. Kollman, *AMBER 11*, 2010, University of California, San Francisco, CA.

- 42 N. V. Hud, F. W. Smith, F. A. Anet and J. Feigon, *Biochemistry*, 1996, **35**, 15383–15390.
- 43 D.A. Case, T.A. Darden, T.E. Cheatham, III, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, K.M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A.W. Götz, I. Kolossváry, K.F. Wong, F. Paesani, J. Vanicek, R.M. Wolf, J. Liu, X. Wu, S.R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D.R. Roe, D.H. Mathews, M.G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko and P.A. Kollman, *AMBER 12*, 2012, University of California, San Francisco.
- 44 A. W. Goetz, M. J. Williamson, D. Xu, D. Poole, S. L. Grand and R. C. Walker, *J. Chem. Theory Comput.*, 2012, **8**, 1542-1555.
- 45 M. Sproviero, K. M. Rankin, A. A. Witham and R. A. Manderville, *J. Org. Chem.*, 2014, **79**, 692-699.



169x76mm (150 x 150 DPI)