

# SCIENTIFIC REPORTS



OPEN

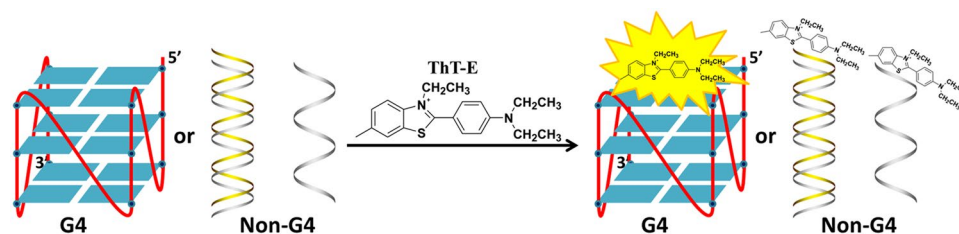
## Ethyl-substitutive Thioflavin T as a highly-specific fluorescence probe for detecting G-quadruplex structure

Ai-jiao Guan<sup>1</sup>, Xiu-Feng Zhang<sup>2</sup>, Xin Sun<sup>2</sup>, Qian Li<sup>1,3</sup>, Jun-Feng Xiang<sup>1,3</sup>, Li-Xia Wang<sup>1,3</sup>, Ling Lan<sup>2,3</sup>, Feng-Min Yang<sup>1,3</sup>, Shu-Juan Xu<sup>1,3</sup>, Xiao-Meng Guo<sup>1,3</sup> & Ya-Lin Tang<sup>1,3</sup>

G-quadruplex has attracted considerable attention due to their prevalent distribution in functional genomic regions and transcripts, which can importantly influence biological processes such as regulation of telomere maintenance, gene transcription and gene translation. Artificial receptor study has been developed for accurate identification of G-quadruplex from DNA species, since it is important for the G-quadruplex related basic research, clinical diagnosis, and therapy. Herein, fluorescent dye ThT-E, a derivative of the known fluorescence probe Thioflavin T (ThT), was designed and synthesized to effectively differentiate various G-quadruplex structures from other nucleic acid forms. Compared with methyl groups in ThT, three ethyl groups were introduced to ThT-E, which leads to strengthened affinity, selectivity and little inducing effect on the G-quadruplex formation. More importantly, ThT-E could be served as a visual tool to directly differentiate G-quadruplex solution even with naked eyes under illumination of ultraviolet light. Thus, this probe reported herein may hold great promise for high-throughput assay to screen G-quadruplex, which may widely apply to G-quadruplex-based potential diagnosis and therapy.

G-quadruplexes are non-canonical nucleic acids secondary structures which are consisted of guanine-rich nucleotide sequence via stacking of Hoogsteen hydrogen-bonded G-quartets in the presence of monovalent cations (usually  $K^+$  or  $Na^+$ )<sup>1,2</sup>. Computational studies have revealed that G-quadruplex could exist simultaneously in human genome and transcriptome, such as chromosomal telomere ends, promoter, mitochondria genome and untranslated regions (UTRs) of mRNA<sup>3–8</sup>. Importantly, these structures are associated with a series of significant genome functions, including transcription, replication, recombination and maintenance of chromosome stability<sup>9,10</sup>. For instance, some G-quadruplex in promoter regions such as c-myc, c-kit, K-ras and VEGF are involved in regulating gene expression<sup>11–13</sup>. G-quadruplex in mRNA untranslated regions (UTRs) could regulate gene at the translation level<sup>14–17</sup>. Therefore, accurate recognition of G-quadruplex from other nucleic acid forms became pretty significant. So far, many approaches have been developed to identify G-quadruplex structure, such as fluorescence resonance energy transfer (FRET), circular dichroism (CD), nuclear magnetic resonance (NMR), X-ray crystallography<sup>18–21</sup>. Besides these techniques, recognizing G-quadruplex by fluorescent chemical probes may provide an efficient way for G-quadruplex screening and detecting. To date, many fluorescence “light-up” probes have been successfully designed and developed to recognize G-quadruplex<sup>22–34</sup>. These probes provide useful tools for *in vitro* and *in vivo* detecting G-quadruplex structures and regulating G-quadruplex related biological function<sup>35–37</sup>. Besides, fluorescent probes targeting G-quadruplex also used in the emerging G-quadruplex based label-free luminescence detection techniques<sup>38–41</sup>. However, in the reported G-quadruplex probes, most of them are water-insoluble and require complicated multi-step synthesis, which would limit their application in some case. Therefore, there is an urgent need for novel fluorescent probes targeting G-quadruplex structures.

<sup>1</sup>National Laboratory for Molecular Sciences, Center for Molecular Sciences, State Key Laboratory for Structural Chemistry of Unstable and Stable Species, CAS Research/Education Center for Excellence in Molecular Sciences, Institute of Chemistry Chinese Academy of Sciences, Beijing, 100190, P. R. China. <sup>2</sup>College of Chemistry Engineering, North China University of Science and Technology, Tangshan, 063009, P. R. China. <sup>3</sup>University of the Chinese Academy of Sciences, Beijing, 100049, P. R. China. Correspondence and requests for materials should be addressed to Q.L. (email: [qianlee@iccas.ac.cn](mailto:qianlee@iccas.ac.cn)) or L.-X.W. (email: [wlx8825@iccas.ac.cn](mailto:wlx8825@iccas.ac.cn)) or Y.-L.T. (email: [tangyl@iccas.ac.cn](mailto:tangyl@iccas.ac.cn))



**Figure 1.** Schematic illustration for highly selective recognition of G-quadruplex based on the fluorescence light-up of ThT-E.

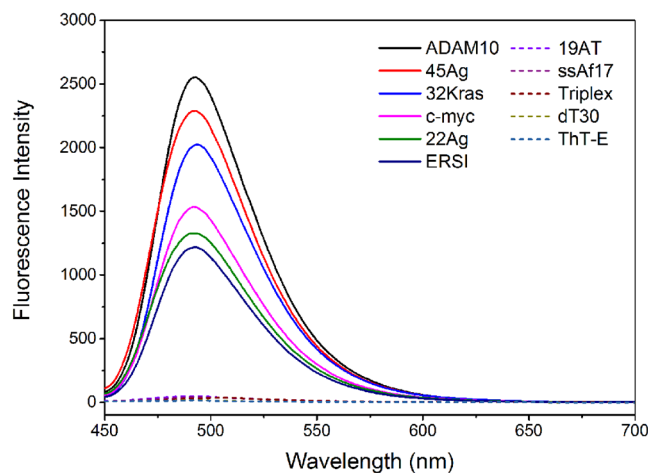
Thioflavin T (ThT), a commercially cationic benzothiazole dye, is served as a sensitive sensor for amyloid fibril and other tissue structures<sup>42</sup>. Recently, Mohanty *et al.* reported a detailed experimental and theoretical study of the interaction between ThT and human telomeric G-quadruplex, illustrating this probe could selectively recognize and stabilize G-quadruplex<sup>43</sup>. Furthermore, Mergny *et al.* expanded the studies to a series of oligonucleotides to further evidence that ThT could be used as a probe for efficient G-quadruplex sensing<sup>44</sup>. However, strong G-quadruplex-forming capacity induced by ThT can significantly change the topology of G-quadruplex, which may limit its application in dynamic detection of G-quadruplex structure formation. Consequently, it is highly desirable to design a new fluorescent probe that could recognize G-quadruplex with improved selectivity and little effect on the G-quadruplex formation. Recently, it has been witnessed that a larger side chain or a specific side chain of probes played important roles in improving the recognition of G-quadruplex through non-bond interactions, such as hydrogen bond, van der Waals force and others<sup>45,46</sup>. For example, the porphyrin derivatives with larger side chains have better selectivity towards G-quadruplex over duplex, which are more suitable for G-quadruplex sensing<sup>47</sup>. On the basis of the above information, we therefore conceived that a new ThT derivative with ethyl groups instead of methyl groups maybe better fit for G-quadruplex recognition. In this work, a new probe ThT-E was designed and synthesized via 1-step transformation starting from 2-(4-aminophenyl)-6-methylbenzothiazole (Fig. S1). Further studies indicated that the fluorescence of ThT-E enhanced significantly when interacting with various G-quadruplex (Fig. 1), which enables ThT-E to be used as a highly-specific fluorescence sensor for G-quadruplex detection.

## Results and Discussion

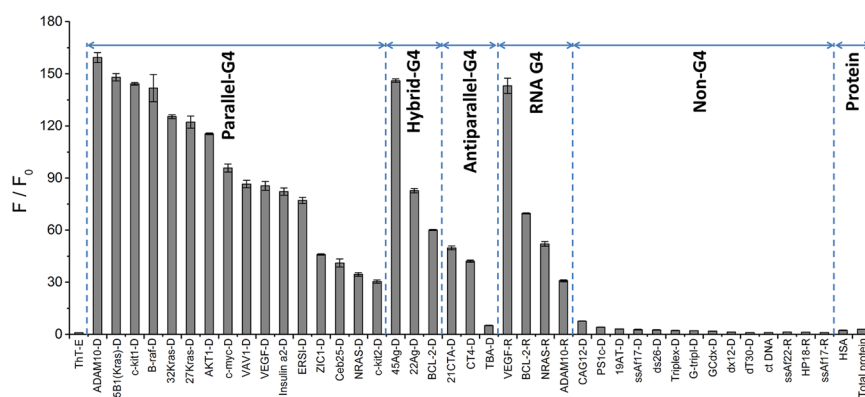
**Design and Characterization of ThT-E.** An attractive application of ThT is served as a sensor for G-quadruplex detection due to their selective fluorescence enhancement, which could be attributed to enforcing planarization and restricting rotation between benzothiazole and dimethylaminobenzene rings around the C-C bond<sup>48</sup>. Substituting methyl groups with ethyl of ThT may better for G-quadruplex recognition since larger side chains play important roles in improving the selectivity and affinity of ligand for G-quadruplex<sup>47</sup>. To further evaluate the effect of ethyl groups on binding strength for G-quadruplexes, the G-quadruplex structure formed from oncogene c-myc promoter regions was selected to carry out molecular docking with ThT or ThT-E using G4LDB system<sup>49</sup>, respectively. The values of  $\Delta G$  for ThT and ThT-E binding to the G-quadruplex structure were  $-27.57 \text{ kJ}\cdot\text{mol}^{-1}$  and  $-28.60 \text{ kJ}\cdot\text{mol}^{-1}$ , respectively (Table S1). In comparison with ThT, the ThT-E showed higher binding ability to the parallel G-quadruplex structure. Obviously, these results demonstrated that ThT-E has a stronger binding strength than ThT for G-quadruplexes, implying that ThT-E may better differentiate G-quadruplexes from other nucleic acid forms. In this study, the water-soluble fluorogenic dye ThT-E was available prepared via 1-step transformation starting from 2-(4-aminophenyl)-6-methylbenzothiazole and iodoethane with 61 percent productivity, indicating that the preparation is simple and inexpensive. The molar extinction coefficient of ThT-E is  $49342 \text{ M}^{-1}\text{cm}^{-1}$ , which was measured at  $\lambda = 418 \text{ nm}$  in water (see Supplementary Fig. S2).

**ThT-E selectively targeted G-quadruplex by fluorescence light-up.** To evaluate the feasibility of the ThT-E probe for G-quadruplex recognition, we compared the fluorescence intensity of ThT-E in the absence and presence of various DNA forms. As shown in Fig. 2, ThT-E alone in the Tris-HCl buffer exhibited negligible fluorescence emission. With the addition of G-quadruplexes, an emission peak at approximately  $\lambda = 492 \text{ nm}$  appeared and remarkably enhanced. The  $F/F_0$  was determined to be 159-, 146-, 125-, 83-, 77- and 96-fold for ADAM10, 45Ag, 32Kras, 22Ag, ERSI and c-myc, respectively. In contrast, the fluorescence of ThT-E enhanced less than 3 fold in the presence of other DNA formations, such as duplex (19AT), single-strand (dT30, ssAf17) and triplex. Thus, ThT-E is an efficient candidate as a fluorophore to differentiate various G-quadruplex structures from other nucleic acid forms by fluorescence enhancement.

To further certify the ability of selectivity recognition of ThT-E towards G-quadruplex, another 23 DNA sequences and 7 RNA sequences were introduced to interact with ThT-E. Moreover, the interactions between ThT-E and human serum albumin (HSA) or cell total protein (TP) were also inspected, respectively. As shown in Fig. 3, both DNA G-quadruplex and RNA G-quadruplex resulted in significant fluorescence enhancement of ThT-E. Interestingly, the B-raf bimolecular G-quadruplex induced remarkable fluorescence enhancement, indicating that ThT-E may have good recognition ability for high-order G-quadruplex. In contrast, much weaker fluorescence emission were observed when ThT-E interacting with TBA G-quadruplex and other nucleic acids forms. Importantly, ThT showed conspicuous fluorescence enhancement for some non-G-quadruplex structures, such as single-stranded (dT30), double-strands (PS1c) and triplex<sup>44</sup>. The fluorescence signals of ThT for these non-G-quadruplex structures even exceeded some G-quadruplex structures. This may lead to a false positive result when using the fluorescent



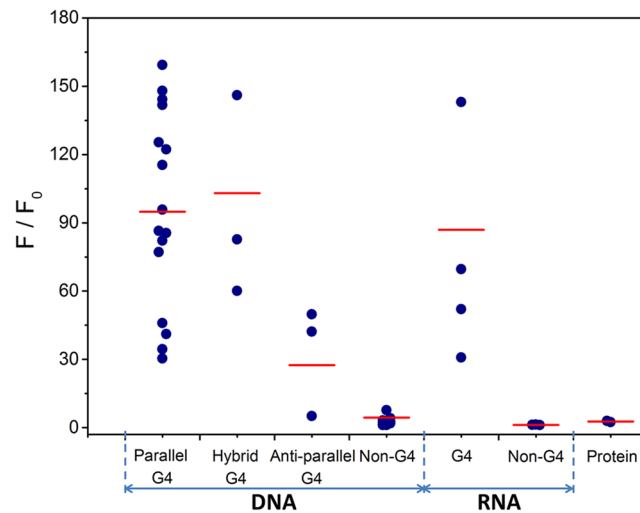
**Figure 2.** Fluorescence emission spectra of ThT-E (2  $\mu$ M) with various oligonucleotides (2  $\mu$ M) in a 20 mM Tris-HCl (40 mM  $K^+$ , pH 7.4) solution.



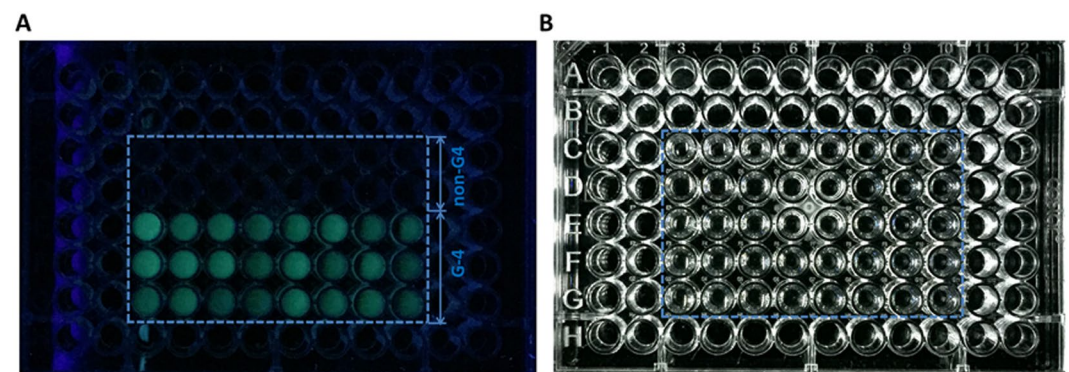
**Figure 3.** Diagrammatic bar representation of ThT-E (2  $\mu$ M) fluorescence enhancement at 492 nm for a variety of oligonucleotide sequences (2  $\mu$ M) or protein (HSA concentration is 2  $\mu$ M, total protein concentration set as 20 mg/L) in a 20 mM Tris-HCl (40 mM  $K^+$ , pH 7.4) solution. Error bars correspond to S.D.

probe ThT to recognize G-quadruplex structures. By contrast, the probe ThT-E does not emit significant fluorescence when interacting with the non-G-quadruplex structures. And it also emits strong fluorescence when interacting with G-quadruplex structures. As a result, the ThT-E could be used to accurately distinguish G-quadruplex and non-G-quadruplex structures. And ThT-E may not give such false positive results (dT30, PS1c and triplex) as the ThT does. Therefore, the fluorescent probe ThT-E shows much better ability for distinguish G-quadruplex and non-G-quadruplex structures than the ThT probe. On the other hand, HSA and TP only induced negligible fluorescence emission of ThT-E, implying the detection of G-quadruplex by ThT-E may not be affected by proteins. Owing to these results, we thus proposed that ThT-E could be served as a highly selective fluorescence “light-up” probe with promising recognition ability not merely for DNA G-quadruplex but for RNA G-quadruplex.

Quadruplexes are structurally polymorphic and could be classified into parallel, antiparallel and hybrid types according to the strand polarities and location of the loops<sup>19</sup>. To further study the relationship between fluorescence behavior of ThT-E and G-quadruplex structure formation of these G-quadruplexes, we next carried out CD measurement (see Supplementary Fig. S3). As depicted in Fig. 4, the statistical data revealed that the fluorescence of ThT-E enhanced about 94.9-, 103-, 27.4- and 87-fold in the presence of parallel DNA G-quadruplex, hybrid DNA G-quadruplex, antiparallel DNA G-quadruplex and RNA G-quadruplex, respectively. Obviously, the enhancement of ThT-E fluorescence intensity for antiparallel G-quadruplexes is weaker than parallel or hybrid G-quadruplex structures, including 21CTA, TBA and CT4 sequences. Compared with ThT-E, there is no distinct rule in the ThT interacting with antiparallel G-quadruplex, such as ThT showed significant fluorescence enhancement for 21CTA G-quadruplex<sup>44</sup>. On the other hand, ThT-E fluorescence enhanced about 4.4-, 1.2- and 2.7 times in the presence of other nucleic acids forms and proteins. Moreover, competitive fluorescence titration assay was performed to further ascertain the selective recognition capacity of ThT-E for G-quadruplex. In this assay, ThT-E was titrated with the 22Ag in the presence of a large excess of duplex DNA (ct DNA, 50  $\mu$ M bp). As shown in Fig. S4, the 22Ag induced similar fluorescence enhancement of ThT-E in the presence or absence of ct DNA, indicating that ThT-E holds great promise for G-quadruplex detection, even in a competitive biological environment.



**Figure 4.** Scatter diagram of ThT-E ( $2\mu\text{M}$ ) fluorescence enhancement folds at 492 nm for a variety of oligonucleotide sequences ( $2\mu\text{M}$ ) or protein in a 20 mM Tris-HCl (40 mM  $\text{K}^+$ , pH 7.4) solution.



**Figure 5.** Photographs of  $2\mu\text{M}$  ThT-E (except for sample C4) with different oligonucleotides ( $2\mu\text{M}$ ) in 20 mM Tris-HCl (40 mM KCl, pH 7.4) solution. (A) Ultraviolet light and (B) Visible light. C3 ~ C10: ThT-E, Tris-HCl buffer, CAG12-D, PS1c-D, 19AT-D, ssAf17-D, ds26-D, Triplex-D; D3 ~ D10: G-tripl-D, GCdx-D, dx12-D, dT30-D, ct DNA, ssAf22-R, HP18-R, ssAf17-R; E3 ~ E10: ADAM10-D, 35B1(Kras)-D, 32kras-D, 27Kras-D, AKT1-D, c-myc-D, VAV1-D, VEGF-D; F3 ~ F10: Insulin a2-D, ERSI-D, ZIC1-D, Ceb25-D, 45Ag-D, c-kit1-D, 22Ag-D, NRAS-D; G3 ~ G10: BCL-2Mid-D, c-kit2-D, 21CTA-D, CT4-D, BCL-2-R, VEGF-R, NRAS-R and ADAM10-R, respectively.

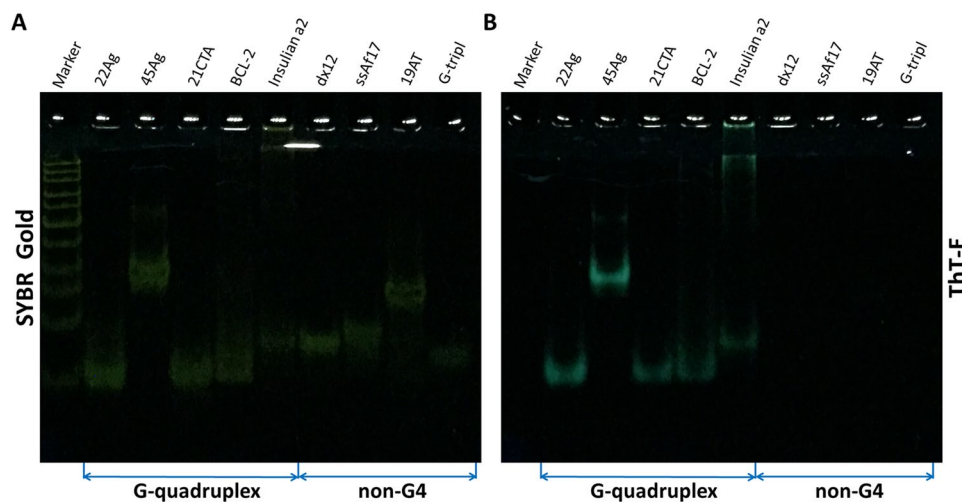
#### The fluorescent light-up effect can be detected even by naked eyes under ultraviolet light.

Based on the exact recognition of ThT-E to G-quadruplex under fluorescence spectra, we further explored the visual character of G-quadruplex in aqueous solution under ultraviolet light condition. As seen in Fig. 5, the difference in fluorescence resulting from the G-quadruplex and others can even be achieved with the naked eye under UV illumination. Almost all of the ThT-E solution in the presence of G-quadruplex exhibited an emission of greenish blue light with the aid of ultraviolet light (Fig. 5A). Meanwhile, the solution was transparent under other nucleic acid structures condition. On the other hand, there is no obvious distinction between G-quadruplex and non-G-quadruplex solutions under visible light, as shown in Fig. 5B. These results showed that the ThT-E could be an efficient fluorescent probe for high-throughput G-quadruplex detection and may be widely used in future G-quadruplex-based biomarker discovery.

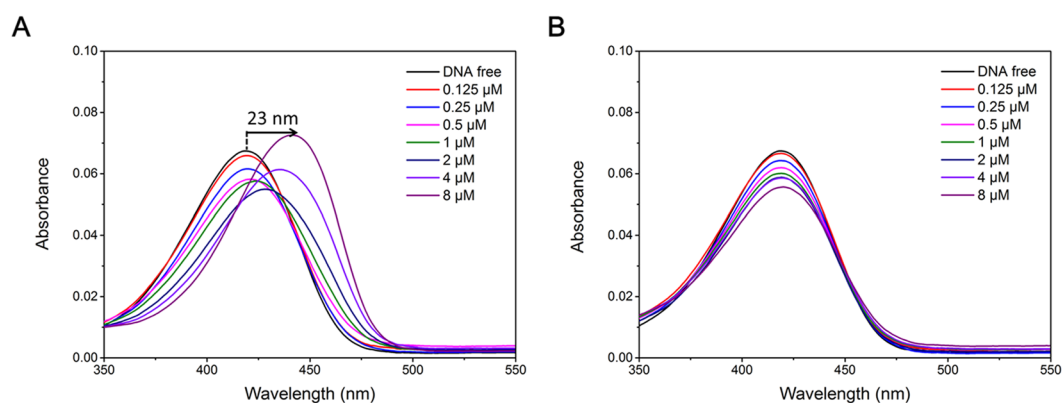
#### The ThT-E molecule can also selectively recognize G-quadruplex structures in the native polyacrylamide gel electrophoresis (PAGE) system.

The significant fluorescence enhancement of ThT-E when binding to G-quadruplex prompted us to investigate whether ThT-E could act as a G-quadruplex recognizer in electrophoresis gels. As shown in Fig. 6A, all bands could be detected after staining with SYBR Gold. In contrast, only the G-quadruplex 22Ag, 45Ag, 21CTA, BCL-2Mid and Insulin a2 could be detected after ThT-E staining (Fig. 6B). Compared to staining with SYBR Gold, the non-G-quadruplex dx12, ssAf17, 19AT and G-tri-plex did not show any band when staining with ThT-E. These results were in accord with the results from fluorescence and ultraviolet visualization experiments showing in Figs 3 and 5, implying ThT-E could also recognize G-quadruplex structures in the native gel system.





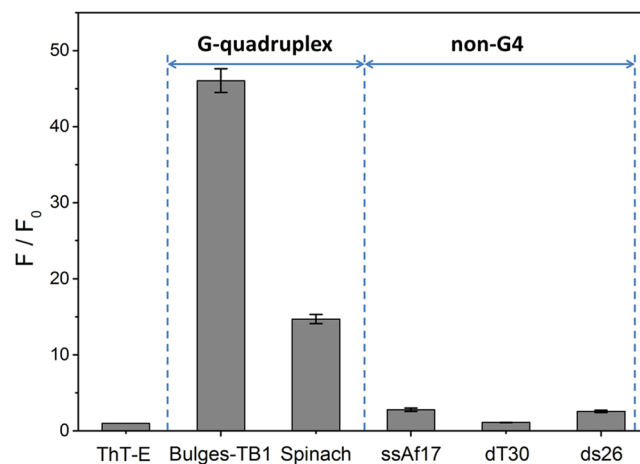
**Figure 6.** Selectivity assays on PAGE. DNA samples were prepared in a 20 mM Tris-HCl, pH 7.4, Loaded on a 5% stacking and 12% separating polyacrylamide gel. The gels were stained with (A) SYBR Gold, (B) 20  $\mu$ M ThT-E, and visualized on a ZF-90 UV viewing cabinet.



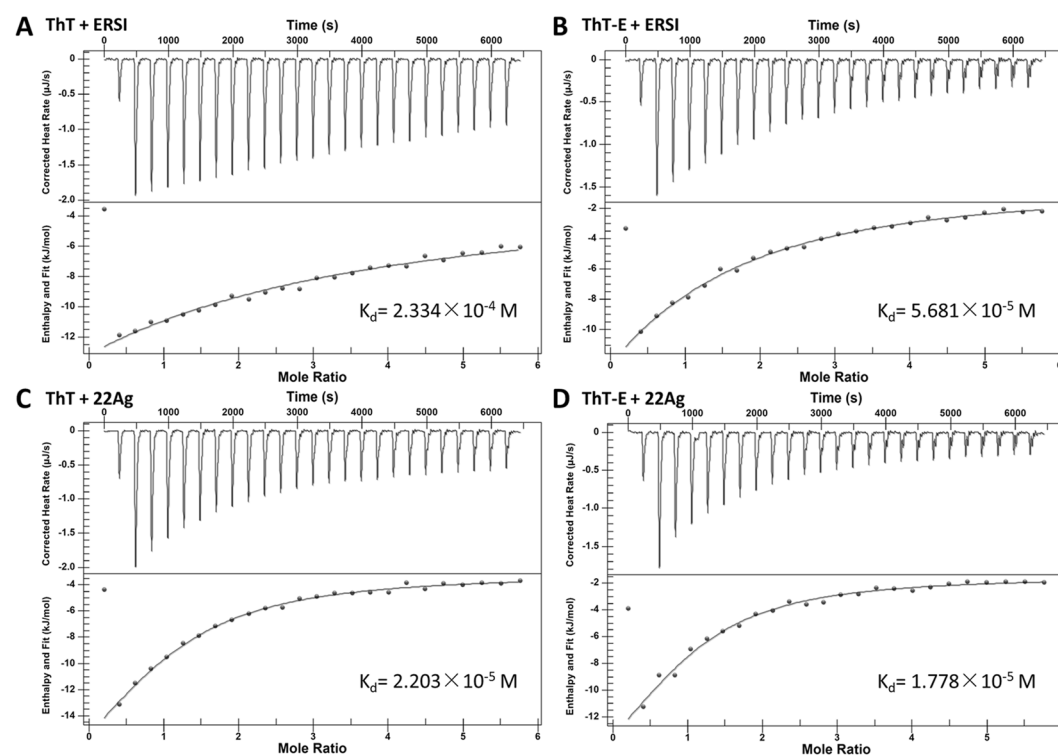
**Figure 7.** Absorbance spectra of ThT-E (2  $\mu$ M) with DNA G-quadruplex sequence (A) c-myc (B) PS1c at eight concentrations ( $\mu$ M): (1) 0, (2) 0.125, (3) 0.25, (4) 0.5, (5) 1, (6) 2, (7) 4, (8) 8.

**UV-vis spectra showed that the presence of G-quadruplex could significantly induced red-shift of the ThT-E molecule.** To get insight into the nature of the interactions between ThT-E and various nucleic acid sequences, absorption spectra was then applied to investigate the specific binding of ThT-E with G-quadruplex. Remarkably, ThT-E displayed distinguishing behaviors when interacting with different nucleic acid sequences. As shown in Fig. 7A, the absorption spectra showed a systematic red shift about 23 nm of the absorption band, accompanied a strong hypochromicity and hyperchromicity with gradual addition of the G-quadruplex DNA from 0 to 4 equivalent. In addition, as shown in Fig. 7B, the non-G-quadruplex (with PS1c duplex as an example) only showed moderate red shift and hypochromic effect, implying the weaker interaction between non-G-quadruplex and ThT-E. These distinctions of absorbance spectra between G-quadruplex and non-G-quadruplex may attribute to the different interaction modes between ThT-E and DNA. It has been reported that the interaction of organic dyes with ds-DNA or single strand DNA mainly through intercalative, electrostatic or groove binding modes<sup>50,51</sup>. However, the electron-rich aromatic core of ThT-E can interact with G-quartet through  $\pi$ - $\pi$  stacking mode at the end of G-quadruplex. Simultaneously, the side arm ethyl groups may also interact with loops or bind to the grooves of the G-quadruplex.

**The ThT-E molecule may also detect non-canonical G-quadruplex structures by fluorescence light-up effects.** The criteria for potential G-quadruplex sequence was restricted to:  $G_{3-5}N_{L1}G_{3-5}N_{L2}G_{3-5}N_{L3}G_{3-5}$ <sup>52</sup>. However, some sequences defying this standard notion could still form stable G-quadruplex structure, such as Bulges-TB1 and Spinach<sup>53,54</sup>. In addition, there are fewer reports about G-quadruplex probe to avaiably detect these non-canonical G-quadruplex structures. To explore the suitability of ThT-E for these G-quadruplexes, the non-canonical G-quadruplex Bulges-TB1 and Spinach were introduced to evaluate the recognition ability of ThT-E. As shown in Fig. 8, the fluorescence intensity of ThT-E increased about 46.05- and 14.71-fold in the presence of Bulges-TB1 and Spinach, respectively. On the contrary, single-strand DNA (ssAf17 and dT30),



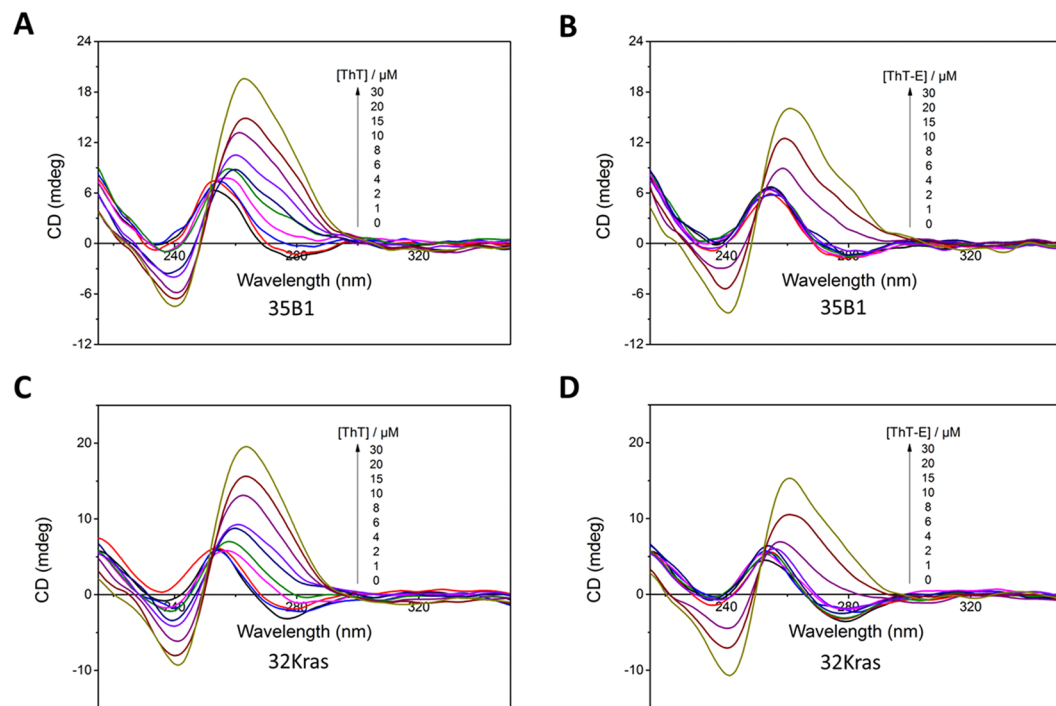
**Figure 8.** Dependence of ThT-E (2  $\mu$ M) fluorescence intensity on non-canonical G-quadruplex and non-G-quadruplex sequences in a 20 mM Tris-HCl (40 mM  $K^+$ , pH 7.4) solution.



**Figure 9.** Representative results of G-quadruplexes with ThT (A,C) or ThT-E (B,D). Calorimetric analysis of the interaction of the G-quadruplexes (22Ag and ERSI) with the ThT or ThT-E.

double-strand (ds26) only induced 2.77-, 1.11- and 2.56-fold, respectively. Obviously, the fluorescence induced by non-canonical G-quadruplex is much higher than non-G-quadruplex structures. From these results, we infer that ThT-E may be suitable for the extensive detection of non-canonical G-quadruplex structures.

**The ThT-E probe showed higher binding affinity with parallel G-quadruplex but similar binding affinity with hybrid G-quadruplex structure compared to ThT.** In order to design a fluorescent probe targeting G-quadruplex with higher selectivity and affinity, we rationally designed and synthesized the fluorescent probe ThT-E. The isothermal titration calorimetry (ITC) was utilized to investigate the binding affinity between G-quadruplexes and ThT or ThT-E. As to the ThT-E molecule, a 1:1 binding model was performed, which has been confirmed by a Job's plot analysis in this study (see Supplementary Fig. S5). As shown in Fig. 9, the dissociation constants of ThT and ThT-E for parallel-G-quadruplex ERSI are  $2.334 \times 10^{-4}$  M and  $5.681 \times 10^{-5}$  M, respectively. Meanwhile, ThT and ThT-E showing a dissociation constant for parallel-G-quadruplex VAV1

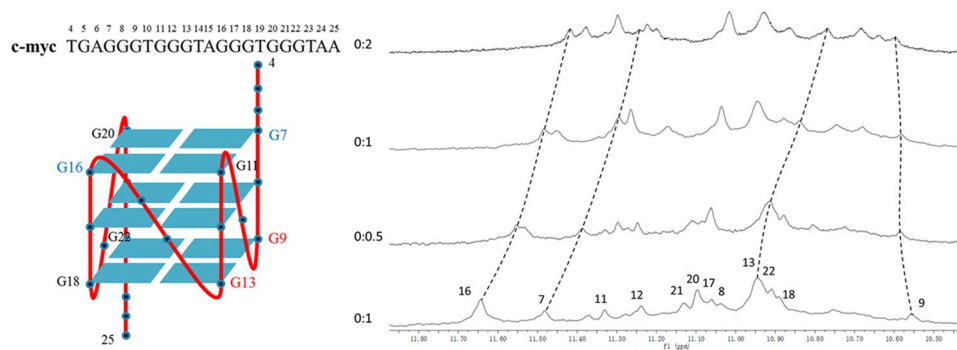


**Figure 10.** Circular dichroism spectra of 35B1 (A,B) and 32Kras (C,D) DNA sequences (2  $\mu$ M) at 20 mM Tris-HCl (pH 7.4) solution with ThT (A,C) or ThT-E (B,D) of concentrations ( $\mu$ M): 0, 1, 2, 4, 6, 8, 10, 15, 20, 30.

are  $3.344 \times 10^{-4}$  M and  $7.473 \times 10^{-5}$  M, respectively (Fig. S6). Based on the above information, we therefore concluded that the ThT-E probe may be more favorable to bind to parallel G-quadruplex than ThT. On the other hand, ThT and ThT-E exhibited similar binding affinity for hybrid- G-quadruplex 22Ag with dissociation constants of  $2.203 \times 10^{-5}$  M and  $1.778 \times 10^{-5}$  M, respectively. The above results were highly in agreement with the results from molecular docking data and significant enhancement of fluorescence upon the ThT-E interacting with parallel and hybrid G-quadruplexes. In previous study, it has been confirmed that ThT could selectively target G-quadruplex with hybrid structures through its fluorescence light-up<sup>48</sup>. As a result, we infer that the ThT-E may be more favorable to bind to parallel G-quadruplex than ThT and may be used as a better fluorescent probe for G-quadruplex sensing.

#### The ThT-E molecule showed much less ability to induce the putative G-quadruplex sequences to form G-quadruplex structures than ThT in the absent of monovalent cations.

ThT as an efficiently selective fluorescence probe for G-quadruplex versus other DNA structure holds great promise for potential diagnostic applications. However, the ThT molecule could significantly induce the putative G-quadruplex sequences to form G-quadruplex structure even in the absence of monovalent cations. That means the ThT molecule itself may affect the sequence's structure and may limit its application in dynamic detection of G-quadruplex structure formation. In order to evaluate the probe's effect on formation of G-quadruplex, ThT and ThT-E were introduced to interact with two putative G-quadruplex sequences (35B1 and 32Kras) in the absence of metal ions, respectively. As shown in Fig. 10, both 35B1 and 32Kras displayed a positive band at 255 nm in 20 mM Tris-HCl (pH 7.4) buffer solution, corresponding to the CD feature of unfolded structure<sup>43</sup>. With increasing the concentration of ThT from 0 to 30  $\mu$ M, the CD spectra gradually changed to a negative band at 240 nm and a positive band at 265 nm, corresponding to the CD features of parallel G-quadruplex. Again, ThT had been confirmed that it could induce the putative G-quadruplex sequences to form G-quadruplex structures in the absent of salt. However, there was no significant change of CD bands when the concentration of ThT-E is less than 10  $\mu$ M, implying a weaker G-quadruplex-inducing ability of ThT-E probe. That is, ThT-E has no effect on the structure of G-quadruplex when the concentration of ThT-E is lower than 10  $\mu$ M. Furthermore, UV illumination measurement was then applied to study the inducing ability of ThT or ThT-E for potential G-quadruplex formation sequences in the absence of monovalent cations. As seen in Fig. S7, the solutions of 32Kras and 35B1 exhibited transparent in the presence of ThT-E, confirming that the oligonucleotides still maintained its unfolding state. That is, ThT-E has a weaker ability to induce the structure's change of nucleic acids. In contrast, both 32Kras and 35B1 solutions exhibited greenish blue in the presence of ThT condition under UV illumination, which may attribute to the transition of DNA formation from random unfolded state to G-quadruplex structure. Based on above results, we therefore proposed that ThT-E has much less structure's effect of putative G-quadruplex sequences than ThT. Therefore, the ThT-E probe is more suitable for dynamically detecting the G-quadruplex's folding states. Overall, these results would facilitate to understand the behavior of ThT-based probe interacting with G-quadruplexes and provide a good structural thought and basis to design or develop G-quadruplex probes and drugs.



**Figure 11.** The schematic illustration of the c-myc G-quadruplex (left). And the imino proton regions of the  $^1\text{H}$  NMR titration spectra (10.5–11.8) of c-myc in pH = 7.4 40 mM  $\text{K}^+$  PBS solution with varying the [ThT-E]: [c-myc G-quadruplex] ratio (right).

**$^1\text{H}$ -NMR results showed that the ThT-E molecule most probably binding to the G-quadruplex at the 5'-end G-tetrad.** To gain further insights into the interaction of G-quadruplex with ThT-E, 1D  $^1\text{H}$  NMR titration experiment were performed with various concentrations of c-myc. As shown in Fig. 11, there are 12 well-resolved imino proton peaks around 10.5–11.8 ppm, which comes from the tetrad-guanines of c-myc. All assignments of the imino proton peaks have been done according to Yang's work<sup>55</sup>. With addition of ThT-E, these peaks showed significant changes in the shape and position, indicating the interaction occurred between ThT-E and G-tetrads. As shown in Fig. 11, the guanines G7, G11, G16 and G20 constitute the 5'-end G-tetrad of c-myc G-quadruplex, and G9, G13, G18, G22 make up the G-tetrad which locates in the 3'-end of G-quadruplex. Further analysis revealed that the change of chemical shift of G7 and G16 (locate in the 5'-end G-tetrad) are more prominent than G9 and G13 (locate in the 3'-end G-tetrad) (Table S2), implying the ThT-E may stack to the 5'-end G-tetrad of the G-quadruplex.

## Conclusions

To conclude, we designed and synthesized a new fluorescent dye ThT-E via 1-step reaction starting from 2-(4-aminophenyl)-6-methylbenzothiazole. The ThT-E was particularly attractive for distinguishing G-quadruplex (DNA and RNA, canonical and non-canonical) from other nucleic acid forms with high selectivity and sensitivity. The G-quadruplex solution could also be visualized even by naked eyes under ultraviolet. Compared with ThT, ThT-E showed higher binding affinity, selectivity for G-quadruplex and little effect on the structure of G-quadruplex forming sequences, which is more suitable for G-quadruplex detection. Overall, the synthesis of ThT-E and assay in this study may be the first step for the G-quadruplex based bio-analysis, further studies for broadening the application of this probe are now underway.

## Materials and Methods

**Molecular mechanism screening.** Molecular docking was carried out using G4LDB system (<http://www.g4ldb.org>). The “compute” function was chosen to predict the binding strength between G-quadruplexes and ThT or ThT-E. The molecular docking and binding evaluation of four intra-G-quadruplexes with ThT or ThT-E were performed automatically after designing the ligand and submitting to G4LDB server.

**Oligonucleotides and proteins.** Oligonucleotide sequences were listed in Table 1. The DNA sequences were purchased from Zixi Bio Tech Co., Ltd (Beijing, China) and RNA sequences were synthesized by Ribo Bio Co., Ltd (Guangzhou, Guangdong, China). Human serum albumin (HSA) and calf thymus DNA (ct-DNA) were purchased from sigma and used without further purification. Total protein (TP) was extracted from HL60 cell by protein extraction kit (BestBio. Co. Ltd. Shanghai, China) and then quantified via BCA protein assay kit (BestBio. Co. Ltd. Shanghai, China). Oligonucleotides and proteins were dissolved directly in 20 mM Tris-HCl buffer (40 mM KCl, pH = 7.4). Before testing, all oligonucleotides were heated to 90 °C for 5 min following with slowly cooling to room temperature. Ultrapure water was prepared with deionized water purified by Milli-Q Gradient ultrapure water system (Millipore) and used throughout all experiments.

**Synthesis of the ThT-E.** To a sealed tube that contained a stir bar were added 2-(4-aminophenyl)-6-methylbenzothiazole (1.2 g, 5 mM), iodoethane (2.34 g, 15 mM) and potassium carbonate (1.38 g, 10 mM). The resulting mixture was stirred and heated at 140 °C for 14 h. Upon completion of the reaction, the mixture was cooled to room temperature and concentrated with the aid of a rotary evaporator. And the residue was purified by column chromatography on silica gel (the percent of methanol in dichloromethane from 2% to 20%). The pure product was isolated as a yellow solid (1.38 g, 61 percent). 2-(4-Aminophenyl)-6-methylbenzothiazole and iodoethane were obtained from J&K Company, Ltd. Dichloromethane and methanol were obtained from Beijing Chemical Plant (Beijing, China).  $^1\text{H}$  NMR (400 MHz,  $\text{CD}_3\text{OD}$ ): 7.98–7.96 (d, 1H), 7.93 (s, 1H), 7.69–7.6 (d, 2H), 7.63–7.61 (d, 1H), 6.92–6.90 (d, 2H), 4.72–4.66 (q, 4H), 3.51–3.46 (q, 4H), 2.50 (s, 3H), 1.65 (t, 3H), 1.17 (t, 6H).  $^{13}\text{C}$  NMR (500 MHz,  $\text{CD}_3\text{OD}$ ):  $\delta$  174.95, 153.83, 141.64, 140.57, 133.41, 132.38, 130.42, 124.49, 117.40,



Name	Type/origin	Sequence (from 5' to 3')	Ref
22Ag-D	G4-HumanDNA telomere	AG <sub>3</sub> TTAG <sub>3</sub> TTAG <sub>3</sub> TTAG <sub>3</sub>	43
45Ag-D	G4-Human DNA telomere	G <sub>3</sub> TTAG <sub>3</sub> TTAG <sub>3</sub> TTAG <sub>3</sub> TTAG <sub>3</sub> TTAG <sub>3</sub> TTAG <sub>3</sub> TTAG <sub>3</sub>	
21CTA-D	G4-Human telomere variant	G <sub>3</sub> CTAG <sub>3</sub> CTAG <sub>3</sub> CTAG <sub>3</sub>	56
c-myc-D	G4-Promoter	TGAG <sub>3</sub> TG <sub>3</sub> TAG <sub>3</sub> TG <sub>3</sub> TAA	57
c-kit1-D	G4-Promoter	G <sub>3</sub> AG <sub>3</sub> CGCTG <sub>3</sub> AGGAG <sub>3</sub>	58
B-raf-D	G4-Promoter	G <sub>3</sub> CG <sub>4</sub> AG <sub>5</sub> AAG <sub>3</sub> A	59
c-kit2-D	G4-Promoter	G <sub>3</sub> CG <sub>3</sub> CGCGAG <sub>3</sub> AGG <sub>3</sub>	
TBA-D	G4-Aptamer	G <sub>2</sub> TTG <sub>2</sub> TGTG <sub>2</sub> TTG <sub>2</sub>	
BCL-2Mid-D	G4-Promoter	G <sub>3</sub> CGCG <sub>3</sub> AG <sub>2</sub> AATTG <sub>3</sub> CG <sub>3</sub>	60
Insulin a2-D	G4-Promoter	ACAG <sub>4</sub> TGTG <sub>4</sub> ACAG <sub>4</sub> TGTG <sub>4</sub>	61
AKT1-D	G4-Promoter	G <sub>3</sub> CG <sub>3</sub> CGGCTCCG <sub>3</sub> CGCG <sub>3</sub>	
VEGF-D	G4-Promoter	G <sub>3</sub> AG <sub>3</sub> TTG <sub>4</sub> TG <sub>3</sub>	62
VAV1-D	G4-Promoter	G <sub>3</sub> CAG <sub>3</sub> AG <sub>3</sub> AACTG <sub>3</sub>	
35B1(Kras)-D	G4-Promoter	AG <sub>3</sub> CGGTGTG <sub>3</sub> AAGAG <sub>3</sub> AAGAG <sub>5</sub> AG <sub>2</sub> CAG	
32 Kras-D	G4-Promoter	AG <sub>3</sub> CGGTGTG <sub>3</sub> AAGAG <sub>3</sub> AAGAG <sub>5</sub> AGG	63
27 Kras-D	G4-Promoter	G <sub>3</sub> CG <sub>2</sub> TGTG <sub>3</sub> AAGAG <sub>3</sub> AAGAG <sub>4</sub>	
Ceb25-D	G4-Minisatellites	AG <sub>3</sub> TG <sub>3</sub> TGTAAGTGTG <sub>3</sub> TG <sub>3</sub> T	
CT4-D	G4/Mixed quartets	G <sub>3</sub> CT <sub>4</sub> G <sub>3</sub> C	
ZIC1-D	G4-5'-UTR	G <sub>3</sub> TG <sub>8</sub> CG <sub>5</sub> AG <sub>2</sub> CCG <sub>4</sub>	
NRAS-D	G4-5'-UTR	G <sub>3</sub> AG <sub>4</sub> CG <sub>3</sub> TCTG <sub>3</sub>	
ADAM10-D	G4-5'-UTR	G <sub>5</sub> ACG <sub>3</sub> TAG <sub>4</sub> CG <sub>3</sub> AG <sub>2</sub> TAG <sub>4</sub>	64
ERSI-D	G4-5'-UTR	G <sub>3</sub> TAG <sub>4</sub> CAAAG <sub>4</sub> CTG <sub>4</sub>	
ssAf17-D	Single strand	CTGAGTTGTATATATTCG	
NRAS-R	G4-5'-UTR	G <sub>3</sub> AG <sub>4</sub> CG <sub>3</sub> UCUG <sub>3</sub>	65
Bulges-TB1	Non-canonical G4	TTGTG <sub>2</sub> TG <sub>3</sub> TG <sub>3</sub> T	53
Spinach	Non-canonical G4	GCAGCCG <sub>2</sub> CTTGTGAGTAGAGTGTGAGCT	54
		CCGTAACG <sub>2</sub> TCGCGTC	
VEGF-R	G4-Promoter	G <sub>2</sub> AG <sub>2</sub> AG <sub>5</sub> AG <sub>2</sub> AG <sub>2</sub> A	66
NRAS-R	G4-5'-UTR	G <sub>3</sub> AG <sub>4</sub> CG <sub>3</sub> UCUG <sub>3</sub>	67
ADAM10-R	G4-5'-UTR	G <sub>5</sub> ACG <sub>3</sub> UAG <sub>4</sub> CG <sub>3</sub> AG <sub>2</sub> UAG <sub>4</sub>	64
BCL-2-R	G4-5'-UTR	AG <sub>5</sub> CCGUG <sub>4</sub> UG <sub>3</sub> AGCUG <sub>4</sub>	68
CAG12-D	Trinucleotide	(CAG) <sub>12</sub>	69
ds26-D	Duplex	CAATCG <sub>2</sub> ATCGAATTCGATCCGATTG	
G-tripl-D	Single strand	GGTTGGTGTGG	
		(a) TTTTTTTTTTTTTTTTTTTT	
Triplex-D	Triplex	(b) AAAAAAAAAAAAAAAAAAAAAA	
		(c) TTTTTTTTTTTTTTTTTTTT	
dT30-D	Single strand	T <sub>30</sub>	
GCdx-D	Stem-loop	GCGCGCGCT <sub>4</sub> GCGCGCGC	
PS1c-D	Parallel-duplex	(a) TTTTTTTTTTATTAAAATTTATAA	
		(b) AAAAAAAAAATAATTTTAAATATT	
19AT-D	Duplex	(a) ACGTCGATTATAGACGAGC	
		(b) GCTCGTCTATAATCGACGT	
dx12-D	Duplex	(a) GCGTGAGTTCGG	
		(b) CCGAACTCACGC	
ssAf22-R	Single strand	UGAGCUUAAUUGUAUUAUUCG	
HP18-R	Hairpin	CAGUACAGAUCUGUACUG	70
ss-Af17-R	Single strand	CUGAGUUGUAUUAUUCG	

**Table 1.** Oligonucleotides used in this work.

113.38, 111.90, 47.21, 45.93, 21.53, 14.85, 12.80; ESI-MS (m/z): [M]<sup>+</sup> calcd. for C<sub>20</sub>H<sub>25</sub>N<sub>2</sub>S<sup>+</sup>, 325.17; found, 325.2. Analysis (calcd., found for C<sub>20</sub>H<sub>25</sub>N<sub>2</sub>S<sup>+</sup>): C (53.10, 52.26), H (5.57, 5.57), N (6.19, 5.91).

**Fluorescence spectroscopy.** Fluorescence spectra were recorded with a Hitachi F-4600 spectrophotometer in a 1-cm path-length quartz cell with excitation of 420 nm and emission wavelength of 492 nm at room temperature. Both excitation and emission slits were 5 nm and voltage was 700 V and the scan speed was set as 1200 nm/min.

**UV-vis measurements.** The absorption spectra were performed on an Agilent 8453 UV-vis spectrophotometer with a 10 mm light path cuvette at room temperature. The absorption titration experiments were carried out by increasing the concentration of oligonucleotide sequences from 0.125 to 8  $\mu\text{M}$ , and ThT-E concentration was fixed at 2  $\mu\text{M}$ .

**Circular dichroism spectroscopy.** CD spectra were carried out on a Jasco-815 spectrometer equipped with a 10-mm path-length quartz cuvette at 298 K in the wavelength range of 200–500 nm. The scan speed was set as 500 nm/min and the response time was 0.5 ns. Each spectrum was the average of three scans.

**Visual observation.** Photographs of ThT-E in the presence of series oligonucleotide sequences were captured in ZF-90 UV viewing cabinet with ultraviolet light or visible light. The concentrations both oligonucleotide and ThT-E of solution were fixed at 2  $\mu\text{M}$ .

**Polyacrylamide gel electrophoresis.** Gel electrophoresis was performed with a 5% stacking and 12% separating polyacrylamide gel. 2.5  $\mu\text{L}$  various oligonucleotides (30  $\mu\text{M}$  for ThT-E, 5  $\mu\text{M}$  for SYBR Gold) were run in  $2 \times$  TBE buffer (0.18 M Tris-boric acid and 4 mM EDTA). Electrophoresis was run at 65 V for 90 min at room temperature. After electrophoresis, the gels were post-stained with a solution of  $1 \times$  SYBR Gold or 20  $\mu\text{M}$  ThT-E in aqueous solution for 10 min, and then briefly washed in water before visualization on a ZF-90 UV viewing cabinet. The polyacrylamide gels were photographed by a digital camera under the aid of ultraviolet light and visible lights.

**Isothermal titration calorimetry.** Calorimetric measurements were carried out using Nano ITC 2G Isothermal Titration Calorimete. The DNA, ThT and ThT-E stock solution were prepared by dissolving the with 20 mM Tris-HCl buffer (40 mM  $\text{K}^+$ , pH = 7.4). In each experiment, volumes of 10  $\mu\text{L}$  of a 500  $\mu\text{M}$  ThT (ThT-E) solution were injected in a 25  $\mu\text{M}$  DNA solution, with a spacing of 250 s between each injection.

**NMR experiments.** The oligonucleotide was dissolved in phosphate buffer (40 mM  $\text{K}^+$ , 90%  $\text{H}_2\text{O}$ /10% deuterioxide). The ThT-E was firstly dissolved in water as a 5 mM stock solution, and then the measured samples were obtained by adding ThT-E stock solution into c-myc solution. The final concentration of c-myc was 1 mM. All NMR spectra were carried out on a Bruker Avance 600 spectrometer equipped with a 5 mm BBI probe at 25  $^\circ\text{C}$ .

## References

- Todd, A. K., Johnston, M. & Neidle, S. Highly prevalent putative quadruplex sequence motifs in human DNA. *Nuc. Acids Res.* **33**, 2901–2907 (2005).
- Henderson, E., Hardin, C. C., Walk, S. K., Tinoco, I. & Blackburn, E. H. Telomeric DNA Oligonucleotides Form Novel Intramolecular Structures Containing Guanine Guanine Base-Pairs. *Cell*. **51**, 899–908 (1987).
- Huppert, J. L. & Balasubramanian, S. Prevalence of quadruplexes in the human genome. *Nuc. Acids Res.* **33**, 2908–2916 (2005).
- Neidle, S. & Parkinson, G. N. The structure of telomeric DNA. *Curr. Opin. Struc. Biol.* **13**, 275–283 (2003).
- Dexheimer, T. S., Sun, D. & Hurley, L. H. Deconvoluting the structural and drug-recognition complexity of the G-quadruplex-forming region upstream of the bcl-2 P1 promoter. *J. Am. Chem. Soc.* **128**, 5404–5415 (2006).
- Eddy, J. & Maizels, N. Conserved elements with potential to form polymorphic G-quadruplex structures in the first intron of human genes. *Nuc. Acids Res.* **36**, 1321–1333 (2008).
- Huppert, J. L., Bugaut, A., Kumari, S. & Balasubramanian, S. G-quadruplexes: the beginning and end of UTRs. *Nuc. Acids Res.* **36**, 6260–6268 (2008).
- Huang, W. C. *et al.* Direct evidence of mitochondrial G-quadruplex DNA by using fluorescent anti-cancer agents. *Nuc. Acids Res.* **43**, 10102–10113 (2015).
- Chen, Y. & Yang, D. Sequence, Stability, and Structure of G-Quadruplexes and Their Interactions with Drugs. *Curr. Protoc. Nucleic Acid Chem.* **17**, 17.5.1–17.5.17 (2012).
- Hansel-Hertsch, R., Di Antonio, M. & Balasubramanian, S. DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential. *Nat. Rev. Mol. Cell. Bio.* **18**, 279–284 (2017).
- Balasubramanian, S., Hurley, L. H. & Neidle, S. Targeting G-quadruplexes in gene promoters: a novel anticancer strategy? *Nat. Rev. Drug Discov.* **10**, 261–275 (2011).
- Yang, D. & Hurley, L. H. Structure of the biologically relevant G-quadruplex in the c-MYC promoter. *Nucleos. Nucleot. Nucl.* **25**, 951–968 (2006).
- Patel, D. J., Phan, A. T. & Kuryavyi, V. Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics. *Nuc. Acids Res.* **35**, 7429–7455 (2007).
- Kumari, S., Bugaut, A., Huppert, J. L. & Balasubramanian, S. An RNA G-quadruplex in the 5' UTR of the NRAS proto-oncogene modulates translation. *Nat. Chem. Biol.* **3**, 218–221 (2007).
- Wolfe, A. L. *et al.* RNA G-quadruplexes cause eIF4A-dependent oncogene translation in cancer. *Nature*. **513**, 65–70 (2014).
- Beaudoin, J. D. & Perreault, J. P. 5'-UTR G-quadruplex structures acting as translational repressors. *Nuc. Acids Res.* **38**, 7022–7036 (2010).
- Lavrado, J. *et al.* KRAS oncogene repression in colon cancer cell lines by G-quadruplex binding indolo[3,2-c]quinolines. *Sci. Rep.* **5**, 9696, <https://doi.org/10.1038/srep09696> (2015).
- Simonsson, T. & Sjoback, R. DNA tetraplex formation studied with fluorescence resonance energy transfer. *J. Biol. Chem.* **274**, 17379–17383 (1999).
- Paramasivan, S., Rujan, I. & Bolton, P. H. Circular dichroism of quadruplex DNAs: Applications to structure, cation effects and ligand binding. *Methods*. **43**, 324–331 (2007).
- Wang, Y. & Patel, D. J. Solution Structure of the Tetrahymena Telomeric Repeat D(T(2)G(4))(4) G-Tetraplex. *Structure*. **2**, 1141–1156 (1994).
- Laughlan, G. *et al.* The High-Resolution Crystal-Structure of a Parallel-Stranded Guanine Tetraplex. *Science*. **265**, 520–524 (1994).
- Tan, W. B., Bhambhani, A., Duff, M. R., Rodger, A. & Kumar, C. V. Spectroscopic identification of binding modes of anthracene probes and DNA sequence recognition. *Photochem. Photobiol.* **82**, 20–30 (2006).
- Chang, C. C. *et al.* A fluorescent carbazole derivative: High sensitivity for quadruplex DNA. *Anal. Chem.* **75**, 6177–6183 (2003).
- Xie, X. *et al.* Asymmetric Distyrylpyridinium Dyes as Red-Emitting Fluorescent Probes for Quadruplex DNA. *Chem-Eur. J.* **19**, 1214–1226 (2013).

25. Monchaud, D., Allain, C. & Teulade-Fichou, M. P. Thiazole orange: A useful probe for fluorescence sensing of G-quadruplex-ligand interactions. *Nucleo. Nucleot. Nucl.* **26**, 1585–1588 (2007).
26. Lin, Y. H. *et al.* Selective recognition and stabilization of new ligands targeting the potassium form of the human telomeric G-quadruplex DNA. *Sci. Rep.* **6**, 31019, <https://doi.org/10.1038/srep31019> (2016).
27. Teulade-Fichou, M. P. *et al.* Selective recognition of G-quadruplex telomeric DNA by a bis(quinacridine) macrocycle. *J. Am. Chem. Soc.* **125**, 4732–4740 (2003).
28. Jin, B. *et al.* Fluorescence Light-Up Probe for Parallel G-Quadruplexes. *Anal. Chem.* **86**, 943–952 (2014).
29. Zagotto, G. *et al.* Amide bond direction modulates G-quadruplex recognition and telomerase inhibition by 2,6 and 2,7 bis-substituted anthracenedione derivatives. *Biorg. Med. Chem.* **16**, 354–361 (2008).
30. Xu, S. J. *et al.* Directly lighting up RNA G-quadruplexes from test tubes to living human cells. *Nuc. Acids Res.* **43**, 9575–9586 (2015).
31. Ma, D. L. *et al.* Recent Developments in G-Quadruplex Probes. *Chem. Biol.* **22**, 812–828 (2015).
32. Wang, M. D. *et al.* Conjugating a groove-binding motif to an Ir(III) complex for the enhancement of G-quadruplex probe behavior. *Chem. Sci.* **7**, 2516–2523 (2016).
33. Tseng, T. Y., Chang, C. C., Lin, J. J. & Chang, T. C. A Fluorescent Anti-Cancer Agent, 3,6-bis(1-methyl-4-vinylpyridinium) Carbazole Diiiodide, Stains G-Quadruplexes in Cells and Inhibits Tumor Growth. *Curr. Top. Med. Chem.* **15**, 1964–1970 (2015).
34. Tseng, T. Y. *et al.* Fluorescent probe for visualizing guanine-quadruplex DNA by fluorescence lifetime imaging microscopy. *J. Biomed. Opt.* **18**, 101309, <https://doi.org/10.1117/1.JBO.18.10.101309> (2013).
35. Rhodes, D. & Lipps, H. J. G-quadruplexes and their regulatory roles in biology. *Nucleic. Acids. Res.* **43**, 8627–8637 (2015).
36. Ma, D. L., Wang, M., Lin, S., Han, Q. B. & Leung, C. H. Recent Development of G-Quadruplex Probes for Cellular Imaging. *Curr. Top. Med. Chem.* **15**, 1957–1963 (2015).
37. Wang, Z. F., Li, M. H., Chen, W. W., Hsu, S. T. D. & Chang, T. C. A novel transition pathway of ligand-induced topological conversion from hybrid forms to parallel forms of human telomeric G-quadruplexes. *Nuc. Acids Res.* **44**, 3958–3968 (2016).
38. Leung, K. H. *et al.* Label-free luminescence switch-on detection of hepatitis C virus NS3 helicase activity using a G-quadruplex-selective probe. *Chem. Sci.* **6**, 2166–2171 (2015).
39. Lin, S. *et al.* Luminescence switch-on detection of protein tyrosine kinase-7 using a G-quadruplex-selective probe. *Chem. Sci.* **6**, 4284–4290 (2015).
40. Lu, L. H., Mao, Z. F., Kang, T. S., Leung, C. H. & Ma, D. L. A versatile nanomachine for the sensitive detection of platelet-derived growth factor-BB utilizing a G-quadruplex-selective iridium(III) complex. *Bioelectron.* **85**, 300–309 (2016).
41. Lu, L. H., Zhong, H. J., He, B. Y., Leung, C. H. & Ma, D. L. Development of a luminescent G-quadruplex-selective iridium(III) complex for the label-free detection of adenosine. *Sci. Rep.* **6**, 19368, <https://doi.org/10.1038/srep19368> (2016).
42. Saeed, S. M. & Fine, G. Thioflavin-T for Amyloid Detection. *Am. J. Clin. Pathol.* **47**, 588–593 (1967).
43. Mohanty, J. *et al.* Thioflavin T as an Efficient Inducer and Selective Fluorescent Sensor for the Human Telomeric G-Quadruplex DNA. *J. Am. Chem. Soc.* **135**, 367–376 (2013).
44. de la Faverie, A. R., Guedin, A., Bedrat, A., Yatsunyk, L. A. & Mergny, J. L. Thioflavin T as a fluorescence light-up probe for G4 formation. *Nuc. Acids Res.* **42** (2014).
45. Jain, A. K. & Bhattacharya, S. Interaction of G-Quadruplexes with Nonintercalating Duplex-DNA Minor Groove Binding Ligands. *Bioconjugate. Chem.* **22**, 2355–2368 (2011).
46. Zhu, L. N., Zhao, S. J., Wu, B., Li, X. Z. & Kong, D. M. A New Cationic Porphyrin Derivative (TMPiPpEOP) with Large Side Arm Substituents: A Highly Selective G-Quadruplex Optical Probe. *Plos One.* **7**, e35586, <https://doi.org/10.1371/journal.pone.0035586> (2012).
47. Dixon, I. M. *et al.* A G-quadruplex ligand with 10000-fold selectivity over duplex DNA. *J. Am. Chem. Soc.* **129**, 1502–1503 (2007).
48. Liu, L. L. *et al.* Molecular Rotor-Based Fluorescent Probe for Selective Recognition of Hybrid G-Quadruplex and as a K<sup>+</sup> Sensor. *Anal. Chem.* **86**, 1622–1631 (2014).
49. Li, Q. *et al.* G4LDB: a database for discovering and studying G-quadruplex ligands. *Nuc. Acids Res.* **41**, D1115–D1123, <https://doi.org/10.1093/nar/gks1101> (2013).
50. Berman, H. M. & Young, P. R. The Interaction of Intercalating Drugs with Nucleic-Acids. *Annu. Rev. Biophys. Bio.* **10**, 87–114 (1981).
51. Bhasikuttan, A. C., Mohanty, J. & Pal, H. Interaction of malachite green with guanine-rich single-stranded DNA: preferential binding to a G-quadruplex. *Angew. Chem. Int. Edit.* **46**, 9305–9307 (2007).
52. Burge, S., Parkinson, G. N., Hazel, P., Todd, A. K. & Neidle, S. Quadruplex DNA: sequence, topology and structure. *Nuc. Acids Res.* **34**, 5402–5415 (2006).
53. Mukundan, V. T. & Phan, A. T. Bulges in G-Quadruplexes: Broadening the Definition of G-Quadruplex-Forming Sequences. *J. Am. Chem. Soc.* **135**, 5017–5028 (2013).
54. Warner, K. D. *et al.* Structural basis for activity of highly efficient RNA mimics of green fluorescent protein. *Struct. Mol. Biol.* **21**, 658–663 (2014).
55. Ambrus, A., Chen, D., Dai, J. X., Jones, R. A. & Yang, D. Z. Solution structure of the biologically relevant g-quadruplex element in the human c-MYC promoter. implications for g-quadruplex stabilization. *Biochemistry.* **44**, 2048–2058 (2005).
56. Largy, E., Marchand, A., Amrane, S., Gabelica, V. & Mergny, J. L. Quadruplex Turncoats: Cation-Dependent Folding and Stability of Quadruplex-DNA Double Switches. *J. Am. Chem. Soc.* **138**, 2780–2792 (2016).
57. Phan, A. T., Modi, Y. S. & Patel, D. J. Propeller-type parallel-stranded g-quadruplexes in the human c-myc promoter. *J. Am. Chem. Soc.* **126**, 8710–8716 (2004).
58. Jantos, K., Rodriguez, R., Ladame, S., Shirude, P. S. & Balasubramanian, S. Oxazole-based peptide macrocycles: A new class of G-quadruplex binding ligands. *J. Am. Chem. Soc.* **128**, 13662–13663 (2006).
59. Wei, D. G. *et al.* Crystal Structure of a Promoter Sequence in the B-raf Gene Reveals an Intertwined Dimer Quadruplex. *J. Am. Chem. Soc.* **135**, 19319–19329 (2013).
60. Dai, J. X., Chen, D., Jones, R. A., Hurley, L. H. & Yang, D. Z. NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region. *Nuc. Acids Res.* **34**, 5133–5144 (2006).
61. Wan, Z. L., Xu, B., Chu, Y. C., Katsoyannis, P. G. & Weiss, M. A. Crystal structure of allo-Ile(A2)-insulin, an inactive chiral analogue: Implications for the mechanism of receptor binding. *Biochemistry.* **42**, 12770–12783 (2003).
62. Sun, D. Y., Guo, K. X., Rusche, J. J. & Hurley, L. H. Facilitation of a structural transition in the polypurine/polypyrimidine tract within the proximal promoter region of the human VEGF gene by the presence of potassium and G-quadruplex-interactive agents. *Nuc. Acids Res.* **33**, 6070–6080 (2005).
63. Cogoi, S., Paramasivam, M., Spolaore, B. & Xodo, L. E. Structural polymorphism within a regulatory element of the human KRAS promoter: formation of G4-DNA recognized by nuclear proteins. *Nuc. Acids Res.* **36**, 3765–3780 (2008).
64. Lammich, S. *et al.* Translational Repression of the Disintegrin and Metalloprotease ADAM10 by a Stable G-quadruplex Secondary Structure in Its 5'-Untranslated Region. *J. Biol. Chem.* **286**, 45063–45072 (2011).
65. Holman, G. G., Zewail-Foote, M., Smith, A. R., Johnson, K. A. & Iverson, B. L. A sequence-specific threading tetra-intercalator with an extremely slow dissociation rate constant. *Nat. Chem.* **3**, 875–881 (2011).
66. Morris, M. J., Negishi, Y., Pazsint, C., Schonhoft, J. D. & Basu, S. An RNA G-Quadruplex Is Essential for Cap-Independent Translation Initiation in Human VEGF IRES. *J. Am. Chem. Soc.* **132**, 17831–17839 (2010).
67. Balkwill, G. D. *et al.* Repression of Translation of Human Estrogen Receptor alpha by G-Quadruplex Formation. *Biochemistry.* **48**, 11487–11495 (2009).

68. Shahid, R., Bugaut, A. & Balasubramanian, S. The BCL-2 5' Untranslated Region Contains an RNA G-Quadruplex-Forming Motif That Modulates Protein Expression. *Biochemistry*. **49**, 8300–8306 (2010).
69. Vorlickova, M., Zimulova, M., Kovanda, J., Fojtik, P. & Kypr, J. Conformational properties of DNA dodecamers containing four tandem repeats of the CNG triplets. *Nuc. Acids Res.* **26**, 2679–2685 (1998).
70. Biffi, G., Di Antonio, M., Tannahill, D. & Balasubramanian, S. Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells. *Nat. Chem.* **6**, 75–80 (2014).

### Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant Number 21778058), the Beijing Natural Science Foundation (Grant Number 7172248), the National Natural Science Foundation of China (Grant Numbers: 21472197, U1432250, 21675162), the Major National Basic Research Projects (973) (Grant Number 2013CB733701) and the “Strategic Priority Research Program” of the Chinese Academy of Sciences (Grant Number XDA09030307). Funding for open access charge: The National Natural Science Foundation of China (21472197).

### Author Contributions

Y.T., Q.L. and L.W. supervised the project and conceived the research strategies. A.G., X.Z., X.S., Q.L., L.W., J.X., F.Y., S.X. and X.G. designed and performed the experiments, analyzed the data and wrote the paper. A.G. revised the manuscript. Y.T., Q.L., L.W., X.Z., X.S., J.X., L.L., F.Y., S.X. and X.G. discussed the results and implications and commented on the manuscript at all stages. All authors have approved the final version of the manuscript.

### Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-20960-7>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018