This is an author postprint version of an article published in:
Biochemistry (**2013**) 52, 5620-5628. http://dx.doi.org/10.1021/bi4006072
Diffusion on institutional repository has been authorized by the Editor on October 24, 2016.

Multiple and cooperative binding of fluorescence light-up probe Thioflavin T with human telomere DNA G-quadruplex

Valérie Gabelica[‡¶,*], *Ryuichi Maeda*[#], *Takeshi Fujimoto*[#], *Hidenobu Yaku*[#,||,§], *Takashi Murashima*[#,||], *Naoki Sugimoto*[#,||], *Daisuke Miyoshi*[#,||,*]

[‡] Physical Chemistry and Mass Spectrometry Laboratory, Department of Chemistry, University of Liège, B-4000 Liège, Belgium, [#] Faculty of Frontiers of Innovative Research in Science and Technology (FIRST), Konan University, 7-1-20 Minatojima-minamimachi, Chuo-ku, Kobe 650-0047, Japan, [||] Frontier Institute for Biomolecular Engineering Research (FIBER), Konan University, 7-1-20 Minatojima-minamimachi, Chuo-ku, Kobe 650-0047, Japan, [§] Advanced Technology Research Laboratories, Panasonic Corporation, 3-4 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0237, Japan

[¶] Present addresses: (1) Univ. Bordeaux, IECB - ARNA laboratory, F-33600 Pessac, France, (2) INSERM, U869 - ARNA laboratory, F-33000 Bordeaux, France)

Abbreviations and Textual Footnotes:

† This work is supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan (MEXT) [24350089 and 24107525 to D.M.], the Scientific Research on Innovative Areas ‘‘Nanomedicine Molecular Science’’ (No. 2306), the ‘‘Strategic Research Foundation at Private Universities’’ (2009-2014) (to N.S.), the Kurata Grants, the Kurata Memorial Hitachi Science and Technology Foundation (to D.M.), the Fonds de la Recherche Scientifique-FNRS [FRFC 2.4528.11 to V.G.], and the Japan Society for the Promotion of Science [research fellowship to T.F.].

* To whom correspondence should be addressed. Tel: +81 78 303 1426; Fax: +81 78 303 1495; Email: miyoshi@center.konan-u.ac.jp. Correspondence can also be addressed to: v.gabelica@iecb.u-bordeaux.fr

Abbreviations. DNA, deoxyribonucleic acid; ThT, Thioflavin T; HPLC, high performance liquid chromatography; ESI-MS, electrospray ionization mass spectrometry; CD, circular dichroism; G-rich sequence, guanine-rich sequence.

ABSTRACT: Thioflavin T (ThT), a typical probe for protein fibrils, also binds human telomeric G-quadruplexes with a fluorescent light-up signal change and high specificity against DNA duplexes. Cell penetration and low cytotoxicity of fibril probes having been widely established, modifying ThT and other fibril probes is an attractive means of generating new G-quadruplex ligands. Thus, elucidating the binding mechanism is important for the design of new drugs and fluorescent probes based on ThT. Here, we investigated the binding mechanism of ThT with several variants of the human telomeric sequence in the presence of monovalent cations. Fluorescence titrations and electrospray ionization mass spectrometry (ESI-MS) analyses demonstrated that each G-quadruplex unit cooperatively binds to several ThT molecules. ThT brightly fluoresces when a single ligand is bound to the G-quadruplex and is quenched as ligand binding stoichiometry increases. Both the light-up signal and the dissociation constants are exquisitely sensitive to the base sequence and to the G-quadruplex structure. These results are crucial for the sensible design and interpretation of G-quadruplex detection assays using fluorescent ligands in general, and ThT in particular.

Some of the most frequently encountered repetitive DNA sequences throughout the genomes of most organisms are guanine (G)-rich sequences and their complementary cytosine (C)-rich strands (1). G-rich sequences can adopt a stable four-stranded structure known as the G-quadruplex (Figure 1A) (2-4). G-quadruplexes can be formed with four Hoogsteen-paired coplanar guanines called a G-quartet (Figure 1B) (2). The best known G-rich (G-quadruplex forming) sequence is found at the ends of the chromosomes, the telomeres (5,6). Moreover, bioinformatic studies have identified more than 300,000 G-rich sequences in the human genome (7,8). Interestingly, G-rich sequences are enriched in the untranslated regions of proto-oncogenes, indicating that G-quadruplexes are critical participants in the gene regulation system (9,10). However, the actual roles and functional mechanisms of G-quadruplexes remain unclear.

G-quadruplex ligands developed to investigate the functional roles of G-quadruplexes should have both high affinity for G-quadruplexes and also high specificity against DNA duplexes (11-14), because the most abundant DNA structure in a living cell is the canonical duplex. Moreover, a bright light-up signal upon G-quadruplex binding is ideal for the detection and imaging of G-quadruplexes (15). Thus, various G-quadruplex probes, which can transduce the binding event into a detectable signal, have been developed (16-21). However, the use of fluorescent probes for the detection of G-quadruplexes is still insufficient due to difficulties in designing fluorescent ligands that combine high specificity for DNA with significant changes in emission quantum yield upon binding. Moreover, such ligands should be able to penetrate cells and have low toxicity for cellular and in *vivo* imaging.

Here, we report on Thioflavin T, or "ThT" (4-(3,6-dimethyl-1,3-benzothiazol-3-ium-2-yl)-N,N-dimethylaniline) (Figure 1C), a new G-quadruplex ligand possessing very unique fluorescent properties. Mohanty and coworkers recently reported that ThT, which is a typical probe for protein fibrils, binds to the human telomeric G-quadruplex sequence 22AG (dAGGG(TTAGGG)$_3$) with a fluorescent light-up signal change and very low background, and almost no fluorescent response with DNA duplexes (22). Various ThT-DNA G-quadruplex binding modes have been discussed, but the ligand binding

stoichiometry and binding mechanisms remain unclear. Also, the origins of the specific fluorescence signal with human telomeric G-quadruplexes compared to DNA duplexes are unknown. We conducted detailed investigations of the ligand binding mode and fluorescence response upon ThT binding with human telomeric DNAs of various nucleotide sequences. Although the fluorescence titration curves of ThT by DNA (22) were confirmed to show a single saturation curve corresponding to a 1:1 complex, our fluorescence reverse titration curves (G-quadruplex by ThT) and electrospray mass spectrometry (ESI-MS) data reveal multiple ligand binding stoichiometries. We discuss in detail how the ligand binding mode influences the ThT fluorescence light-up response.

Materials and Methods

Sample preparation. Oligodeoxynucleotides used in this study were purchased from Hokkaido System Science Co., Ltd. (Hokkaido, Japan) after purification by high pressure liquid chromatography (HPLC) or from Eurogentec (Seraing, Belgium) with RP-cartridge-Gold quality. ThT was purchased from Sigma-Aldrich (Tokyo, Japan) and used without further purification. The binding studies of ThT were carried out using G-quadruplex forming DNA sequences: 22AG [dA(GGGTTA)$_3$GGG], 21G [d(GGGTTA)$_3$GGG], 22GT [d(GGGTTA)$_3$GGGT], 46AG [dA(GGGTTA)$_7$GGG], and 45G [d(GGGTTA)$_7$GGG]. As controls, we also used the following DNA sequences: dsDNA [dAGTTCAAGGCGCCTTGAACT] as a self-complementary duplex, tsDNA [TCTTCTCTTTCT<u>TTTT</u>TCTTTCTCTTCT<u>TTTT</u>AGAAGAGAAAGA] (loop regions are underlined) as a triplex-forming sequence, and ssDNA [dT$_{20}$] as a single-stranded unstructured sequence. Single-strand concentrations of the DNA oligodeoxynucleotides were determined by measuring the absorbance at 260 nm at high temperature (23). ThT concentration was determined by measuring the absorbance at 412 nm (molar extinction coefficient ε = 32,000 mol^{-1} cm^{-1}) (24). All other chemical reagents were of reagent grade from Wako Pure Chemical Co., Ltd. (Osaka, Japan) or Sigma-Aldrich (Diegem, Belgium) and used without further purification.

Fluorescence spectroscopy. Fluorescence spectra of ThT were measured using Varioskan Flash (Thermo Fisher Scientific K.K., Yokohama, Japan) with a 96-well titer plate (SUMILON S-type plate, Sumitomo Bakelite Co. Ltd., Tokyo, Japan). The fluorescence spectra were obtained by taking the average of three scans made at 0.5-nm intervals from 450 to 700 nm. The excitation wavelength was 450 nm. Before measurement, the samples were heated at 90°C for 10 min, gently cooled at 0.5°C min^{-1}, and incubated at 25°C for 1 h.

DNA oligonucleotides were titrated with 0.85 μM ThT (ThT by DNA fluorescence titration). The fluorescence intensity at 485 nm (F_{485}) was plotted against the concentration of the DNA oligonucleotide. The data were fitted to Equation (1) based on a model that postulates *n* binding sites of ThT to estimate the apparent dissociation constant (K_D) at 25°C:

$$F_{485} = a \cdot [\mathrm{ThT}]^n / (K_D + [\mathrm{ThT}]^n) + b \qquad (1)$$

where *a* is the scaling factor, and *b* is the initial F_{485} value. The "ThT by DNA" fluorescence titration curves were fitted by KaleidaGraph (Synergy Software, PA).

Reverse titrations, "DNA by ThT", were also carried out. ThT was titrated with 5 μM DNA oligonucleotide at 25°C. The fluorescence intensity at 485 nm was plotted against the concentration of ThT.

Circular dichroism (CD) spectroscopy. The G-quadruplex structure has characteristic peaks in CD spectroscopy. Antiparallel G-quadruplexes show positive and negative peaks around 295 and 265 nm, respectively, whereas parallel G-quadruplex structures show positive and negative peaks around 260 and 240 nm, respectively (25). Moreover, hybrid-type G-quadruplexes show two positive peaks at 290 nm and 265 nm, and a negative peak at 240 nm (26). CD spectra of DNA oligonucleotides were

measured for 10 μM DNA total strand concentration using a J-820 spectropolarimeter (JASCO Co. Ltd., Hachioji, Japan) with a 0.1-cm path length quartz cell at 4°C. The CD spectrum was obtained by taking the average of three scans made from 200 to 350 nm. Before measurement, the DNA samples were heated at 90°C for 10 min, gently cooled at 0.5°C min^{-1}, and incubated at 25°C. The temperature of the cell holder was regulated by a PTC-348 temperature controller (JASCO), and the cuvette-holding chamber was flushed with a constant stream of dry N_2 gas to avoid water condensation on the exterior of the cuvette.

Electrospray ionization mass spectrometry (ESI-MS). All experiments were carried out on a Synapt HDMS mass spectrometer (Waters, Manchester, UK) operated in negative ion mode. The ion mobility mode was turned on, although the mobility data were not used in the present context. The DNA concentration was 5 μM in 100 mM NH_4OAc solution, and the concentration of ThT varied between 0 and 60 μM. Typical conditions were, for 22AG, 21G, and 22GT, a cone voltage = 40 V at a backing pressure of 3.3 mbar, and for the duplex and 45G, a cone voltage = 80 V at a backing pressure of 4.8 mbar. The ions were introduced in the IMS cell (N_2 at 0.532 mbar) with a bias of 15 V. We tested higher voltages as well, and observed no diminishing of the relative intensities of the complexes relative to the free DNA. Thus, we concluded that loss of the complexes by collision-induced dissociation does not occur in the chosen experimental conditions. The relative areas of the mass spectral peaks corresponding to each complex stoichiometry were extracted using MassLynx 4.0. Using the oligonucleotide dT_6 as an internal intensity standard (27), we confirmed that the relative responses of free and bound DNA were similar. In addition, the ratio between the sum of the intensities of all complexes and the intensity of the internal standard did not change by more than 30% in the titrations. Thus, it is not necessary to consider different mass spectral response factors of the different species. The equilibrium concentrations of each "DNA + *n*ThT" species (where *n* is the number of bound ThT in each complex) were determined from the peak areas $A_{(DNA + nThT)}$ as follows:

$$[DNA + nThT] = [DNA]_{tot} \cdot \frac{A_{(DNA+nThT)}}{\sum_{i=0}^{n_{tot}} A_{(DNA+iThT)}} \quad (2)$$

For 22AG, three independent titration experiments were performed using different source and IMS voltages. The standard errors of the relative intensities extracted from the mass spectra were very small (see Figure 3 below).

Data fitting for multiple binding stoichiometries. With mass spectrometry, the equilibrium concentrations of all species were obtained, and it was therefore possible to determine each consecutive K_D separately. These K_D's are macroscopic, i.e. they reflect the total concentration of each complex *stoichiometry*, without distinction of the binding site. These K_{Dn}'s are defined as:

$$K_{Dn} = [\mathrm{DNA}+(n-1)\mathrm{ThT}]\,[\mathrm{ThT}] \,/\, [\mathrm{DNA}+n\mathrm{ThT}] \quad (3)$$

These parameters were determined simultaneously by fitting the ESI-MS titration data (concentration of each species at equilibrium as a function of total ThT concentration added) using DynaFit 3.0. (28). The DynaFit program computes the composition of complex mixtures at equilibrium by solving simultaneously the mass balance equations for the component species by using the multidimensional Newton-Raphson method. DynaFit 3.0 was also used to simulate equilibrium concentrations of each species upon titration, again using the K_{Dn} values as input parameters.

Results

ThT fluorescence light-up with various human telomeric G-quadruplexes. One previous report focused exclusively on the DNA sequence 22AG (22). Figure 2A shows fluorescence titration of 0.85 μM ThT by increasing concentrations of different variants of the human telomeric sequence, and with single-stranded, double-stranded, and triplex DNAs, in KCl. Significant fluorescence enhancement was

found with all G-quadruplexes but not for the other DNA structures. The intensity increased by more than 700-fold by addition of 50 μM 22AG. This significant increase is comparable with the fluorescence increase upon fibril protein binding (100-2000 fold) (29-31), and is in line with the 2100-fold fluorescence increase when 3 μM ThT is titrated in 50 mM KCl and 50 mM Tris (pH 7.2) (22). Notably, the fluorescence enhancements were only 6-fold, 12-fold, and 5-fold in the presence of 50 μM dsDNA, tsDNA, and ssDNA respectively, demonstrating the specific fluorescence change of ThT with G-quadruplexes. This drastic fluorescence enhancement allows us to detect the 22AG G-quadruplex with a limit of detection (LOD; $3\sigma/S$, where σ and S are standard deviation and slope of the linear calibration curve) of 5.2 nM with 30 μM ThT at 25°C (Supporting information Figure S1). Equilibrium binding constants with duplex, triplex, and single stranded DNAs could not be obtained from the fluorescence titration experiments. A competition experiment where ThT was titrated by 22AG in the presence of 100 μM dsDNA (Supporting information Figure S2) showed almost the same titration curve as in the absence of dsDNA. This strongly suggests that preferential ThT binding to 22AG compared to duplex DNA is the major reason for the selectivity of the fluorescence detection. The binding selectivity has been further confirmed by ESI-MS (see below).

The apparent dissociation constants (K_D) of ThT binding with 22AG and other telomeric G-quadruplex sequences were evaluated from the fluorescence titration curves (ThT by DNA) with an assumption that n ThT and one G-quadruplex form the complex. Table 1 summarizes the K_D values and Hill constants (n) deduced from fluorescence titration fittings using Equation 1. Since the number of ThT molecules binding to 22AG was evaluated to be 0.83 ~ 1.1, these results suggest a 1:1 binding in G-quadruplex-ThT complexes. However, the Hill constant does not directly reflect the number of ThTs bound to G-quadruplexes. Thus, the stoichiometry will be further discussed with ESI-MS analysis.

Both the fluorescence light-up signal and the K_D of ThT are influenced by the base sequence. This could be due to a specific fluorescence response in the vicinity of certain bases. For example, the K_D and fluorescence response are lower with 21G or 22GT than with 22AG (Figure 2A and Table 1). This

might suggest that the 5'-adenine is promoting ThT binding and fluorescence. However, the base sequence also has known effects on the G-quadruplex structure. For example, in the presence of KCl, 22AG forms a mixture of predominantly hybrid (3+1) structures (32), whereas 22GT forms a 2-tetrad antiparallel structure.

To understand the extent to which binding and fluorescence depend on the structure, we studied 22AG with different monovalent cations. The circular dichroism (CD) spectra of 10 μM 22AG with different cations (Supporting information Figure S3) are consistent with previous structural studies by NMR and X-ray crystallography and biophysical studies by spectroscopy on structural polymorphisms of 22AG that depend on the coexisting cation (32-34). In the presence of NaCl, 22AG adopts a well-defined three-tetrad antiparallel structure (34). In the presence of NH_4OAc (which was used for ESI-MS), the CD spectrum show positive peaks around 295 nm and 260 nm, although of lower intensity than in the presence of KCl. We therefore suppose that in NH_4OAc there is a mixture of structures that might include (3+1), 2-tetrad, and 3-tetrad antiparallel G-quadruplexes. In the presence of LiCl, 22AG did not form a stable G-quadruplex. Fluorescence titration curves of ThT by 22AG in different cations are shown in Figure 2B, and the K_D values are also summarized in Table 1. The order of K_D values is: $K^+ < Li^+ \approx NH_4^+ < Na^+$. The affinity is independent of the thermal stability of the G-quadruplex ($K^+ > Na^+ > NH_4^+ >> Li^+$), and these results suggest that ThT preferentially binds to pre-formed (3+1) G-quadruplex folds. In addition, it is interesting that ThT has higher affinity in the presence of LiCl than in the presence of NaCl, although 22AG does not form a stable G-quadruplex in the presence of LiCl. We also found that ThT did not bind to single-stranded ssDNA, as shown Figure 2A. Thus it is possible that ThT can induce 22AG to fold into a G-quadruplex in the presence of LiCl as shown by Mohranty et al. in Tris buffer alone (without cations).

Stoichiometry and dissociation constants of ThT binding to telomeric G-quadruplexes. The forward (ThT by DNA) fluorescence titration data (Figure 2) show evidence only for a 1:1 binding stoichiometry.

In contrast, reverse titration experiments (fixed DNA concentration titrated by increasing ThT concentrations as shown in Figure 3A) revealed bell-shaped curves, with fluorescence light-up at low ThT concentrations, then fluorescence quenching at higher ThT concentrations. The curves cannot be accounted for by a single binding site, and indicate a more complex binding mechanism.

ThT-DNA binding was therefore studied by ESI-MS, which is a powerful tool to evaluate binding affinity and stoichiometry (35-37). Since ESI-MS measurements can only be carried out in a buffer containing NH_4^+ as a coexisting cation, we confirmed that the fluorescence titration curve obtained in the presence of NH_4OAc (Figure 3A, black) was similar to that obtained in the presence of KCl (red), although the fluorescence intensity was different. Figure 3B shows ESI-MS spectra of 5 μM 22AG without ThT in NH_4OAc at 22°C (room temperature); the *m/z* peaks annotated $22AG^{6-}$, $22AG^{5-}$, and $22AG^{4-}$ correspond to 22AG alone with six, five, and four negative charges, respectively. Figure 3C shows ESI-MS spectra of 5 μM 22AG with 30 μM ThT. Free 22AG is still detected, and new peaks appear, corresponding to 22AG with ThT molecules simultaneously bound. Surprisingly, this spectrum demonstrates that up to four ThT molecules bind the 22AG G-quadruplex. Figure 3D shows the ESI-MS titration curve reconstructed from a series of ESI-MS spectra recorded at various total ThT concentrations. The data were fitted according to a binding model involving up to four consecutive ligand binding events, giving four consecutive macroscopic dissociation constants as defined in Equation 3 (see Materials and Methods). On the other hand, any microscopic binding model that assumed independent binding sites did not fit the data (data not shown), indicating that the binding model for the curve fitting procedure, assuming a cooperativity between the binding events, is reasonable for data analysis.

ESI-MS experiments were also carried out with dsDNA, 21G, 22GT, and 45G. Representative spectra are shown in Figure 4 for dsDNA and 45G, and all reconstructed titrations are shown in Supporting information Figure S4. In contrast to the G-quadruplexes, 5 μM duplex DNA with 30 μM ThT showed only a very weak signal corresponding to a 1:1 complex (Figure 4B), confirming the low binding

affinity of ThT for duplex DNA. All consecutive dissociation constants obtained by ESI-MS in the 100 mM NH_4OAc solution are summarized in Table 2. The binding affinities for the 1:1 complexes (K_{D1}) are consistent with those evaluated from the fluorescence titration experiments, indicating that the binding model for the curve fitting and the K_D values evaluated from the ESI-MS titrations are appropriate. ESI-MS titrations were not pursued to ThT concentrations higher than 60 μM to minimize the risk of detecting nonspecific binding. The highest K_D (156 μM), obtained for dsDNA, was therefore due either to specific binding or can be considered as the limit of nonspecific binding detection for ESI-MS in our experimental conditions.

At 60 μM ThT, we detected up to three ligands bound for 21G, up to four with 22AG and 22GT (the first three K_D's of 22GT being identical to those of 21G), and up to seven ThT ligands bound to 45G (up to 6 bound ThTs are visible in Figure 4D, at 30 μM ThT). 45G contains two G-quadruplex units (38-40). Overall, our results suggest that the binding stoichiometry is at least up to 3 or 4 ThT molecules per G-quadruplex unit depending on the sequence.

Cooperative binding of ThT. Positive cooperativity can be inferred by examining the values of the macroscopic dissociation constants and considering statistics. If all binding sites are independent (no cooperativity) and equivalent, each microscopic binding event has an identical microscopic dissociation constant (K_d). The probability of populating the *n*th site depends on the number of sites remaining available, and therefore decreases with *n*. As a consequence, the macroscopic K_{Dn} values increase as *n* increases, even though the microscopic constants (K_d) are identical for each binding event. If the sites are independent (no cooperativity) but non-equivalent, the most affine binding sites are more likely to be populated first, so the increase of macroscopic K_D with *n* is even larger. Therefore, any value of K_{Dn} smaller than $K_{D(n-1)}$ indicates positive cooperativity. With this reasoning, from Table 2, cooperativity is obvious for the second binding sites of 22AG and 45G, and for the fourth binding site of 22GT.

In all other cases, one has to take into account the number of available sites and binding statistics to

detect cooperativity from the consecutive macroscopic K_{Dn}'s. For example, ESI-MS shows that 22AG possesses at least four ligand binding sites. If the total number of available sites (X) is equal to 4, and if all four sites bind ThT with the same microscopic dissociation constant (K_d), independently of the fact that other sites are already occupied, the four macroscopic dissociation constants K_{Dn} can be expressed as: $K_{D1} = (1/4)K_d$, $K_{D2} = (2/3)K_d$, $K_{D3} = (3/2)K_d$, and $K_{D4} = 4K_d$ (41). So, if $K_{D2} < 8/3\ K_{D1}$ when $X = 4$, (or similarly $K_{D2} < 3\ K_{D1}$ when $X = 3$), there is positive cooperativity for the binding of the second ligand. The statistics therefore show that the second ligand binding event is cooperative for all G-quadruplexes. With similar reasoning, it can be deduced that the third ThT ligand binding event is also cooperative, although the differences between the four K_D values are small.

ThT-induced conformational changes in human telomeric G-quadruplex DNA. It was reported previously that ThT induces changes in the CD spectra of 22AG in the presence of 50 mM KCl, lowering the peak at 270 nm, and thereby indicating a transition towards more "antiparallel" strand arrangements (22). Such conformational transitions of G-quadruplexes induced by ligands have been observed. We carried out similar CD titration experiments by ThT on 5 μM 22AG in 100 mM KCl (Figure 5A) and in 100 mM NH_4OAc (Figure 5B), and found similar behavior in these conditions. As the latter conditions are similar to those for the ESI-MS titrations described above, each CD spectrum of Figure 5B can be interpreted in light of the concentration of each species in solution (Figure 3D).

In the presence of NH_4OAc, the 1:1 complex is already present at 10 μM ThT, but at that ThT concentration the CD spectrum of 22AG is not affected. However, the CD spectrum changes (with a lowering of the band at 265 nm) when 20 μM ThT is added, and according to ESI-MS the 2:1 complex is present in those conditions. At 40 μM and 60 μM ThT, the dip at 265 nm is even deeper. At those concentrations, the relative abundance of the 2:1 complex increases and the 3:1 complex also becomes present. At 60 μM, free 22AG represents less than 20% of the total concentration of 22AG. Thus, 22AG with two or more ligands bound seems to have a different conformation than free 22AG or 22AG with

one ligand bound. This suggests that ligand-induced conformational changes or re-organizations in the stacking and dipole interactions of bases might be responsible for cooperativity, and that conformational changes require two bound ThTs, although further detailed studies are required to reveal how ThT molecules might induce this. This conformational change induced by the higher stoichiometry can explain the cooperative binding of ThT molecules and quenching of fluorescence at higher ThT concentrations as discussed below.

Discussion

Forward and reverse titration experiments give different and complementary information. In the "forward" titration (ligand by DNA), the ligand concentration is smaller than K_{D1}, the major species is the free DNA, and the next most abundant species is the 1:1 complex. As confirmed by simulations (Supporting information Figure S5) using the macroscopic K_D's from Table 2, the abundance of other complexes with higher stoichiometries is minor in the conditions used for the "forward" titrations. The plateaus in the forward fluorescence binding curves therefore provide information on the relative response of the 1:1 stoichiometry. In contrast, the "reverse" titration (DNA by ligand) conditions are more favorable to detect the contribution of other complexes with higher stoichiometries. When the ligand concentration increases, the complexes become predominant compared to the free DNA. Such comparisons of both titrations are useful to investigate the ligand binding mechanism.

Moreover, we utilized complementary techniques to obtain information on the ligand binding mode: (1) Separation techniques sensitive to the mass of the complex (e.g., ESI-MS) can define the stoichiometries formed and the macroscopic K_D values for each ligand binding stoichiometry when performed quantitatively. (2) Spectroscopic techniques sensitive to the ligand environment reveal how the ligand response varies with the ligand binding stoichiometry. In the case of ThT binding to human telomeric G-quadruplexes, fluorescence showed how the response of ThT changes with the total number of ligands bound to the G-quadruplexes. (3) Finally, spectroscopic techniques sensitive to the DNA

conformation (e.g., CD spectroscopy for G-quadruplexes) also provide important complementary information on the ligand binding mode, and the conformation of the complexes can be probed only when their abundance is predominant over that of the free DNA.

Here we combined three independent probing techniques (ESI-MS, fluorescence, and CD), and in the case of fluorescence, forward and reverse titrations, to decipher the ligand binding mode of ThT to human telomeric G-quadruplexes. Mass spectrometry reveals each stoichiometry more clearly than any other spectroscopic method. However, it can only be performed in ammonium buffers. Nevertheless, the reverse titrations performed by fluorescence and CD spectroscopy revealed similar behavior of the complexes in NH_4OAc and in KCl. Thus, it is reasonable to conclude that similar binding behavior is at work in both buffers, despite the fact that the K_D values differ.

ThT binding mode to G-quadruplexes. Although the mechanism of fluorescence light-up of ThT upon biomolecule (protein fibril) binding is still a matter of debate (30, 31, 42-48), it is usually thought that fluorescence light-up of ThT is caused when the rotational motion of the bond between the benzothiazole ring and the dimethylaminobenzene ring is restricted into a planar conformation in the ground and excited states upon binding (47). In addition, the stacking of aromatic amino acids has also been shown to enhance ThT fluorescence, even though ThT may not be constricted to a planar conformation (48).

Here, fluorescence titration experiments showed the influence of the nucleotide sequence on the fluorescence response: the height of the plateau is always lower with 21G and 22GT than with 22AG and the 2-unit G-quadruplexes (46AG and 45G). This observation is valid in both KCl (Figure 3A) and NH_4OAc (Figure S6). This highlights the role of the 5' adenine in enhancing the light-up fluorescence. The presence of the 5’-A, or of a full G-quadruplex unit in the neighborhood of ligand binding, could further restrict the torsional motions of the bond between the rings and be responsible for further enhancement of the fluorescence response with those sequences. Such a fluorescence intensity change

depending on the rotation of a bond in a G-quadruplex ligand has been proposed in an acetylene-bridged purine dimer as a fluorescent switch-on probe (20). Fluorescence light-up upon ThT stacking on planar G-quartets is fully consistent with this mechanism.

Moreover, the presence of a 5'-A, or of a full G-quadruplex unit and a TTA linker, increases not only the fluorescence response as discussed above but also the binding affinity of ThT. In contrast, 21G and 22GT give similar K_{D1} values and fluorescence responses, so one can infer that the 3'-T does not influence the first binding event, but only the fourth one (see Table 2). All these results indicate higher affinity of ThT for the 5'-end than for the 3'-end of the human telomeric sequence. A likely binding site for ThT molecules in human telomeric DNA is around the 5'-A (A1); the neighboring G-quartet plane of G1, G11, G15, and G21; and the neighboring loop composing T12, T13, and A14 (highlighted in Figure 1A). Overall, these results suggest that the stacking of nucleobases influences the fluorescence light-up signal of ThT, as proposed previously for β2-microglobulin fibrils (48).

The presence of the A1 also increases the cooperativity of the second and third binding events. In 22AG, the fourth ligand binding event ($K_D \approx 120$ μM) appears independent, and may occur on a remote site compared to the first three. Interestingly, a similar $K_D \approx 120$ μM is also observed for 22GT and 21G, for the third binding site, instead of the fourth site for 22AG. The only difference between 22GT and 21G is the presence of a fourth binding site for 22GT, which shows cooperative binding relative to the third site. A possible interpretation is that the presence of A1 causes three ThT ligands to gather on the 5'-end, whereas G-quadruplexes starting with a guanine bind only two ThT ligands on the 5'-end. Similarly, the presence of a 3'-T causes two ThT ligands to gather on the 3'-end, whereas sequences terminated by a guanine bind only one ThT ligand. Sequences containing several G-quadruplex units can bind even more ligands. ESI-MS titration experiments of ThT on 45G (which can be seen as one 22GT unit connected to a 22AG unit by a thymine linker) showed a larger number of ThT bindings, up to seven.

Multiple ligand binding to G-quadruplexes is rare for large ligands such as BRACO-19 and

naphthalene diimide-based compounds (49,50), but it is not unprecedented for smaller G-quadruplex ligands. For example, three daunomycin molecules can stack onto one G-quartet in a parallel G-quadruplex (51). In the case of the isoquinoline alkaloid berberine and its derivatives, up to 6 ligand molecules bind to one G-quadruplex with some binding modes (52). Although our DNA sequences, experimental conditions, and detection methods differ from previous reports, the results indicate that a smaller molecule can bind G-quadruplexes with higher stoichiometry. Since ThT is smaller than daunomycin (which has five six-membered rings), it is possible that up to three ThT molecules stack onto a single G-quartet.

Cooperativity can be due to proximal interactions between ligands, to ligand-induced conformational rearrangements, or most likely to both. Proximal interactions between ligands is supported by fluorescence quenching effects. Additional ThT binding events indeed changed the fluorescence of the ThT ligands that were already bound, and this is easier to understand if additional ThT molecules bind in close proximity to the ligands already in place. Ligand-induced conformational changes are also supported by CD spectroscopy. The CD titrations showed that ligand addition leads to structures having a lower proportion of guanine-guanine homo-stacking. This implies either rearrangement of the strand polarity or disruption of G-G homo-stacking by the ligand bindings.

Implications for specific fluorescence light-up detection of G-quadruplexes using ThT. Our results show that the fluorescence response of ThT is the highest in the 1:1 complex, and quenched in higher ligand binding stoichiometries. As a consequence, quantitative fluorescence light-up detection assays of G-quadruplex presence are better carried out at low ThT concentrations in which the 1:1 complex is the major form of bound ThT. This should also be kept in mind when detecting G-quadruplexes in the presence of other G4-binding molecules, or for the quantitative interpretation of ligand-induced fluorescence displacement assays (19,53). In addition, the macroscopic K_D values of 45G, which includes two G-quadruplex units (38-40), are significantly smaller than any value found for 21G, 22GT,

or 22AG, which are composed of only one unit each. These results suggest that ThT could have higher affinity for longer telomeric DNA sequences, and be used for their fluorescence detection with a lower detection limit than for single-unit G-quadruplexes. These results are useful for further rational design of G-quadruplex ligands with higher affinity and sensitivity based on ThT.

Conclusion

In this study, we combined three independent probing techniques (ESI-MS, fluorescence, and CD), and in the case of fluorescence, forward and reverse titrations, to decipher the ThT binding mechanism to human telomeric DNA G-quadruplexes. The mechanism of ThT binding to human telomeric G-quadruplexes can be summarized as follows. Multiple ThT molecules bind to the 5' G-quartet cooperatively. The first bound ThT occurs in a binding pocket formed with the A1 and 5' G-quartet, and produces the brightest fluorescence light-up signal. When present, A1, or in the case of 2-unit sequences, a full G-quadruplex, stacks on ThT, further restraining the rotational motion between the benzothiazole ring and the dimethylaminobenzene ring. The cooperative binding starts with the binding of the second ThT molecule. It is coupled to a fine structural or base rearrangement of the G-quadruplex, and also to changes in the ThT environment that cause a decline of fluorescence response. Since the cell penetration ability and low toxicity of ThT have been confirmed, this type of information is useful to set optimal conditions for cellular imaging by use of ThT. The stoichiometry-dependent fluorescence of ThT will also help current investigations in our laboratory, in which we are designing future G-quadruplex ligands similar to how new fibril probes have been developed based on a ThT scaffold (54). The affinity of ThT with G-quadruplexes (micromolar range) is relatively low compared with G-quadruplex ligands reported previously. On the other hand, the affinity of ThT to fibril proteins is within the same order of magnitude, and this is sensitive enough for fibril detection in test tubes and living cells. This is partly because the background fluorescence of ThT is exceptionally low, leading to a high signal-to-noise ratio. In addition, G-quadruplexes show many similarities with protein fibrils in regards to the light-up

mechanism of ThT upon binding. Further structural analysis of ThT-G-quadruplex complexes might therefore give new insights into the fluorescence light-up mechanism of ThT upon binding to protein fibrils. As shown for ThT, probes that bind to a broad spectrum of biomolecules can potentially be applied to G-quadruplexes, leading to new G-quadruplex probes and ligands with new structural and mechanistic features.

Supplemental materials

Supplementary Figures: Fluorescence intensity of 30 μM ThT with various concentrations (0 to 0.1 μM) of 22AG (S1); Fluorescence intensity of ThT at 485 nm with various concentrations of 22AG in the presence and absence of dsDNA (S2); CD spectra of 22AG (S3); ESI-MS titrations curves of 5 μM DNA with various concentrations of ThT (S4); Simulation of stoichiometric forms, demonstrating that the 1:1 complex is the main contributor to the fluorescence signal in our fluorescence "forward" titration conditions (S5); Fluorescence intensity of ThT in the presence of DNA in 100 mM CH_3COONH_4 buffer (S6). This material is available free of charge via the internet at http://pubs.acs.org.

Acknowledgement

The authors thank Ms. Aya Yamamoto for fluorescence measurements.

Tables and Figures

Table 1. Dissociation constants (K_D) at 25°C and Hill constants (n) of ThT-G-quadruplex binding. These parameters were evaluated from fluorescence titration (λ_{ex} = 450 nm, λ_{em} = 485 nm) of 0.85 μM ThT by various telomeric DNA sequences in buffers with different cations. All cation concentrations are 100 mM.

Sequence	Cation	K_D (μM)	n
22AG	K^+	5.9 ± 0.4	0.98 ± 0.04
22AG	NH_4^+	40 ± 12	0.87 ± 0.09
22AG	Na^+	96 ± 7	1.1 ± 0.1
22AG	Li^+	36 ± 3	0.9 ± 0.1
21G	K^+	45 ± 17	0.90 ± 0.16
21G	NH_4^+	46 ± 5	1.0 ± 0.05
22GT	K^+	22 ± 3	0.93 ± 0.19
22GT	NH_4^+	59 ± 15	0.90 ± 0.08
45G	K^+	7.0 ± 2.6	0.83 ± 0.12
45G	NH_4^+	25 ± 9	0.73 ± 0.09
46AG	K^+	2.7 ± 0.9	0.97 ± 0.05
46AG	NH_4^+	8.8 ± 2.6	0.83 ± 0.09

Table 2. Macroscopic, consecutive dissociation constants [$K_{D,n}$ (μM)] determined from ESI-MS in 100 mM NH_4OAc solution at 22°C.

Sequence	K_{D1}	K_{D2}	K_{D3}	K_{D4}	K_{D5}	K_{D6}	K_{D7}
45G	15 ± 0.5	8.0 ± 0.3	11 ± 0.4	15 ± 0.6	30 ± 1.2	45 ± 3	73 ± 8
21G	55 ± 2.1	86 ± 7	120 ± 4	n.d.[a]	n.d.	n.d.	n.d.
22GT	58 ± 0.8	76 ± 2.2	110 ± 7	80 ± 10	n.d.	n.d.	n.d.
22AG	37 ± 0.6	32 ± 0.7	56 ± 1.6	130 ± 9	n.d.	n.d.	n.d.
dsDNA	160 ± 3	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.

[a] "n.d." indicates complex not detected in the mixture of 5 μM DNA and 60 μM ThT.

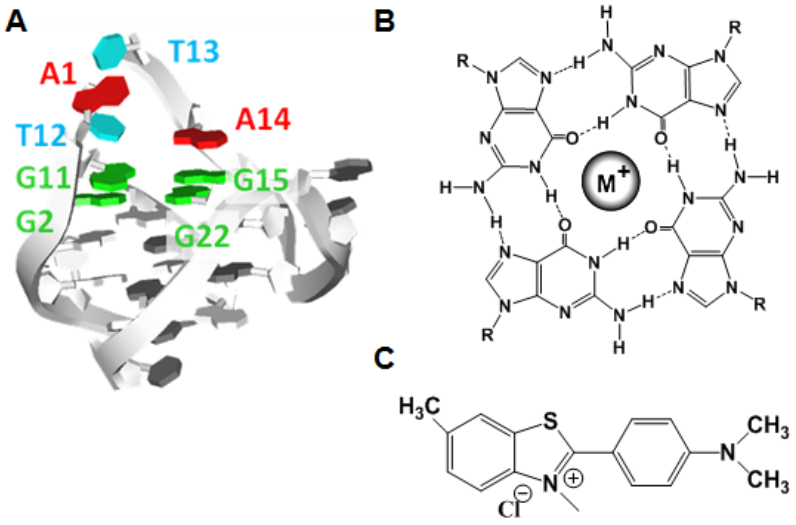


Figure 1. (A) Solution structure of human telomere DNA (4). (B) Chemical structure of a G-quartet. M^+ indicates a monovalent cation coordinating guanine O6. (C) Chemical structure of ThT.

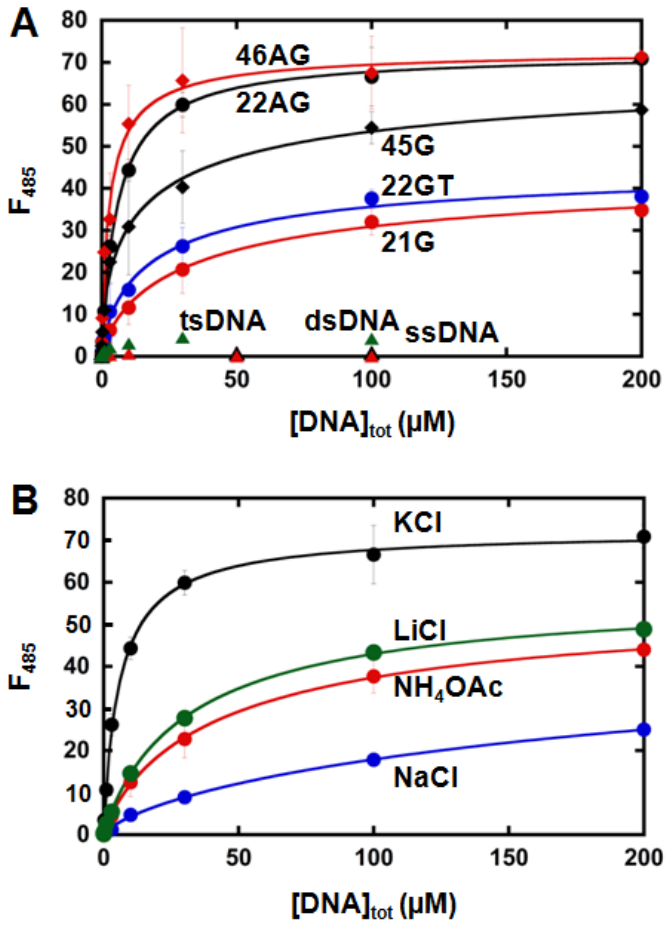


Figure 2. (A) Fluorescence intensity of ThT at 485 nm with various concentrations of 22AG (black circles), 21G (red circles), 22GT (blue circles), 46AG (red diamonds), 45G (black diamonds), ssDNA (black triangles), dsDNA (red triangles), or triplex DNA (green triangles) in 100 mM KCl buffer. (B) Fluorescence intensity at 485 nm of ThT with various concentrations of 22AG in buffers containing 100 mM KCl (black), CH_3COONH_4 (red), NaCl (blue), or LiCl (green). All measurements were carried out with 0.85 μM ThT at 25°C. Excitation wavelength is 450 nm.

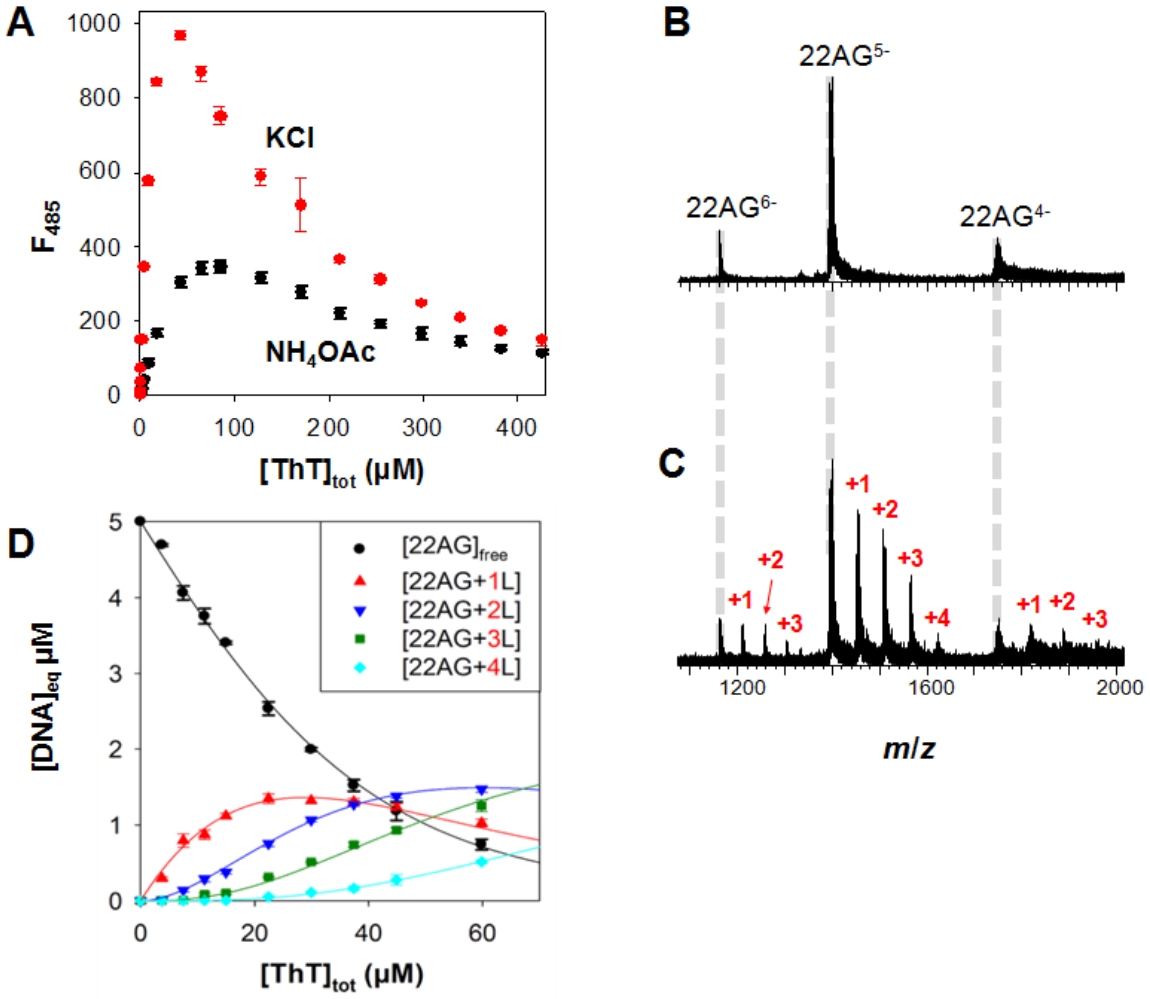


Figure 3. Titration of 5 μM 22AG with ThT. (A) Fluorescence titration curve in 100 mM KCl (red) or NH_4OAc (black). (B and C) ESI-MS spectra of 22AG with (B) 0 μM ThT and (C) 30 μM ThT, showing up to four ligands bound per 22AG molecule. (D) ESI-MS titration curve with equilibrium concentrations of each form deduced from the relative mass spectral peak areas. The lines correspond to fitting using sequential macroscopic binding events (K_{Dn} values are summarized in Table 2). Error bars in (A) and (D) represent standard error of triplicate measurements.

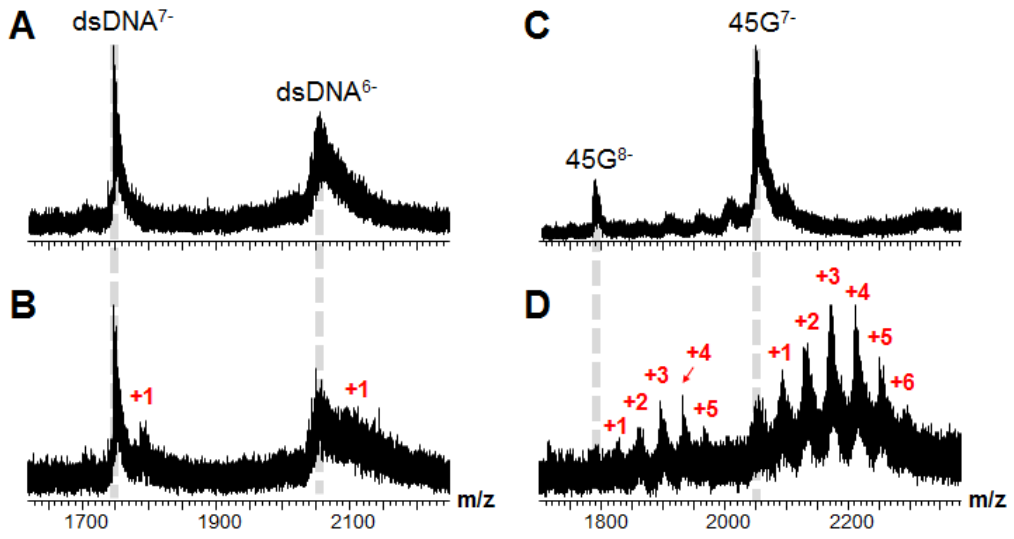


Figure 4. ESI-MS spectra of 5 μM dsDNA in the absence (A) and presence (B) of 30 μM ThT, and of 5 μM 45G in the absence (C) and presence (D) of 30 μM ThT. All experiments were carried out in 100 mM NH_4OAc solution at 22°C (room temperature).

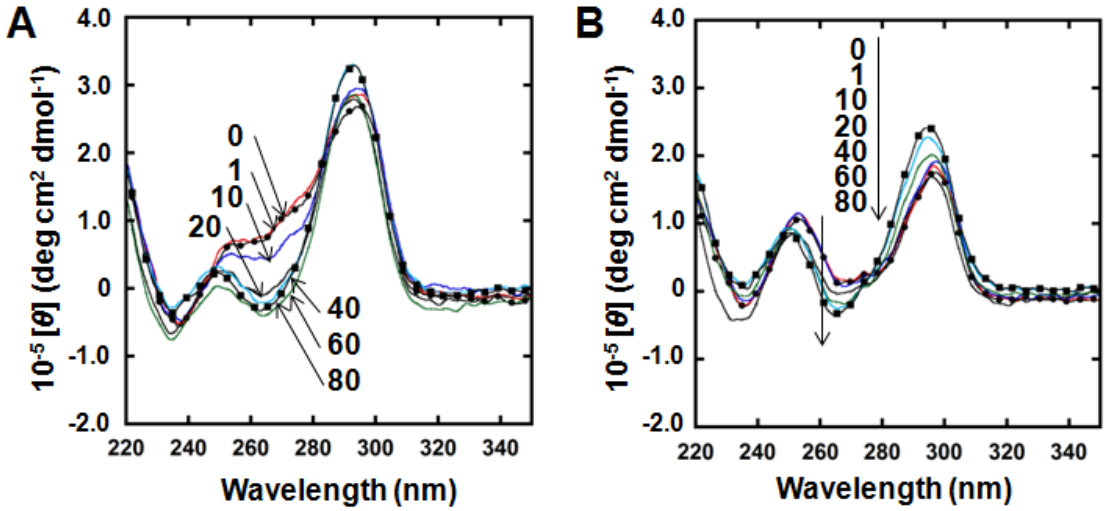
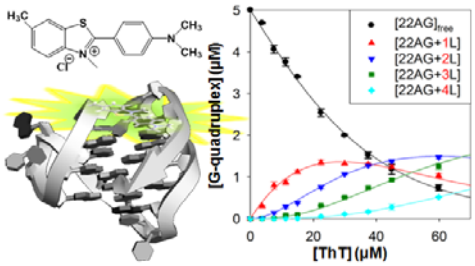


Figure 5. (A) CD spectra of 10 μM 22AG with various concentrations of ThT in KCl buffer. (B) CD spectra of 10 μM 22AG with various concentrations of ThT in NH_4OAc buffer. ThT concentrations are 0 μM (black with circles), 1 μM (red), 10 μM (blue), 20 μM (purple), 40 μM (light blue), 60 μM (green), and 80 μM (black with squares).

Table of Contents

For Table of Contents Use Only