Spring 2013

Investigating G-quadruplex Formation and its Structural Perturbation by and Interactions with N-methyl mesoporphyrin IX

Steven P. Barrett , '13

Investigating G-quadruplex Formation and its Structural Perturbation by and Interactions with N-methyl mesoporphyrin IX

Presented as a Senior Honors Thesis
in Chemistry

Steven P. Barrett
April 29, 2013
Swarthmore College
Advisor: Liliya A. Yatsunyk

Acknowledgements

This work would not have been possible without the help and guidance of my research advisor, teacher, and friend, Professor Liliya Yatsunyk. I would also like to thank past and present members of the Yatsunyk Lab: Vienna Tran, Michelle Ferreira, Navin Sabharwal, Joshua Turek-Herman, Supriya Davis, Karan Ahluwalia, and especially, Jack Nicoludis, who served as an important mentor for me as I began my research at Swarthmore.

I would next like to thank my colleagues at the University of Pennsylvania, notably Professor Brad Johnson, who advised my research over the summer of 2011 and who constantly reminds me of the excitement of science and the beauty of research through our collaborations, and Jesse Platt, my direct mentor during my time at Penn.

Thanks to the professors who contributed to this project and had the patience to teach me for four years, specifically Stephen Miller, Alison Holliday, Paul Rablen, and Robert Paley. Also, I would like to thank my friends and peers in the chemistry department: Nina Kogekar, Tori Barber, Travis Mattingly, Alice Wong, Jake Tracy, Brett McLarney, Emma Spady, Liz Williams, Richard Chen, Vishaal Chhabria, Vuong Dang, Mariah Parker, Ahmad Ammous, and Di Yan.

Finally, I would like to thank my family and friends for all of their support and encouragement throughout my time at Swarthmore. You are more central to this work than you can imagine.

Abstract

G-quadruplexes (GQs) are noncanonical secondary structures of DNA that can adopt a multitude of different topologies. Here, we explore the factors that drive the formation of these structures and favor the specific topologies they adopt.

Our discussion of this topic begins with an investigation of the interactions of N-methyl mesoporphyrin IX (NMM), a small aromatic molecule, with Tel22, a 22-nucleotide DNA sequence that serves as a model for the putative quadruplex formed by the human telomeric repeat sequence. NMM is exceptionally selective for GQ DNA over all other types of nucleic acid secondary structure and specifically prefers the parallel GQ topology, inducing an isomerization in Tel22 from the mixed hybrid structure to the all-parallel form. Thus, we were interested in understanding the molecular basis of this specificity. Our recently obtained crystal structure of the NMM-Tel22 complex indicates that the *N*-methyl group plays a large role in this selectivity by distorting the porphyrin core of NMM, thereby optimizing its interactions with GQ DNA and precluding its association with double stranded B-form sequences. NMM's peripheral side chains are largely unresolved in the crystal structure while the *N*-methyl group has well-defined electron density, suggesting either a large degree of thermal motion or the binding of 4 distinct NMM regioisomers. Here, we detail the separation and spectroscopic charcterization of these isomers. Our preliminary data suggest that all four isomers are able to interact with Tel22, although there appears to be significant contamination of these samples.

In the next chapter of this work, we explore the bimolecular quadruplex formation between a strand of DNA and its complement. Using a variety of different tactics, we attempted to induce GQ formation in engineered sequences endowed with what is known as duplex derived interstrand quadruplex forming potential (ddiQFP). Our efforts were mostly unsuccessful likely due to the extremely high thermal stability of the competing duplex structure and the difficulty of interpreting our spectroscopic data.

In the final chapter of this work, we investigate quadruplex formation at mitochondrial DNA (mtDNA) sites associated with DNA deletions. Three of these sequences showed stable quadruplex formation under physiological conditions, indicating that GQs may be responsible for these breakpoints *in vivo*. In low salt conditions, the addition of lead (Pb^{2+}) to two of these sequences appears to select for a small population of stable GQs and even induce GQ formation in one sequence. This suggests an additional mechanism for the toxicological effects of lead on mitochondrial structure and function.

Because of the disparate nature of these topics, the conclusions and future directions will be discussed at the end of each chapter rather than together.

Table of Contents

Introduction

Background: DNA Composition and Non B-form DNA Structures

Deoxyribonucleic acid (DNA) is the genetic material contained within a cell that encodes all of an organism's proteins. DNA is a long polymer consisting of repeating units referred to as nucleotides. Each nucleotide has three chemical components: sugar (deoxyribose), phosphate, and a nitrogenous base, the composition of which determines identity of the nucleotide. The four bases of DNA are adenine (A), guanine (G), cytosine (C), and thymine (T). Typically, DNA is found as a double-stranded helix with hydrogen bonds forming between A and T, and C and G. B-form DNA is a right-handed helix with ~10 nucleotides per turn and is the most common DNA structure found *in vivo*.

Depending on the particular sequence and conditions, DNA can adopt non B-form structures. Examples of these structures are Z-DNA, a left-handed double helix whose sequence motifs have tracts of alternating pyrimidines and purines, and A-DNA, a right-handed double helix, similar to that of B-DNA, but with a more compact helical structure that is most commonly found under dehydrating conditions (1). Other non-cannonical structures include cruciforms, formed by sequence tracts of inverted repeats, and triplexes, three-stranded DNA structures. Finally, G-quaduplexes (GQs) are four-stranded structures formed by G-rich sequences of DNA. These unusual structures are the focus of this work.

G-quadruplex structure and heterogeneity

The idea of higher-order structures formed by G-rich sequences was conceived in 1910 when Bang observed gel formation of guanylic acid at high concentrations (2). In 1962, using X-ray diffraction data of guanylic acid crystals, Gellert *et al*. proposed a

tetrameric structure in which the guanines are related to each other via a four-fold rotational axis with two hydrogen bonds per base. Gellert asserts that these planar structures could form aggregates arranged by stacking the tetramers on top of each other, leaving a channel in the center. Indeed, Gellert's prescient interpretation of the data was extremely accurate, and today these aggregates are known as G-quadruplexes and each of the planar tetramers are called G-tetrads or G-quartets (see Figure 1A) (3). In addition, monovalent cations such as sodium and potassium are located in the central channel and are coordinated in a square antiprismatic arrangement to the nearby carbonyl groups of the surrounding guanines.

GQs are not only formed by free guanine bases floating in solution but also by

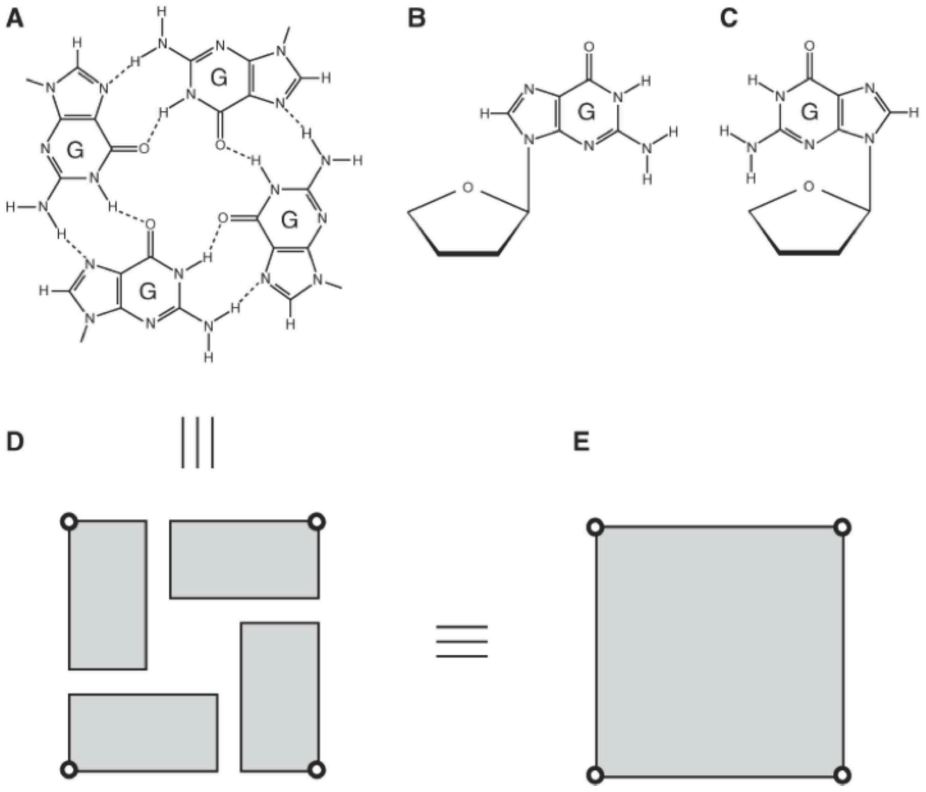


Figure 1. (A) G-tetrad alignment. **(B,C)** Guanine in **(B)** anti and **(C)** syn glycosidic conformations. **(D,E)** Schematic presentation of a G-tetrad used in Figure 2 (6).

guanine bases contained within a given strand or on different strands of DNA. An intramolecular quadruplex is formed by one continuous strand of DNA with the following sequence motif: $G_{\geq 2}N_xG_{\geq 2}N_yG_{\geq 2}N_zG_{\geq 2}$. N specifies arbitrary nucleotides (usually A and T) that make up loop regions normally 2-7 nucleotides long. The most stable quadruplexes are formed when the guanine tracts are longer and the loops are shorter (4). Bimolecular and tetramolecular quadruplexes consisting of two and four separate strands of DNA, respectively, can also be formed. These structures are gaining importance in nanotechnology and materials sciences (5).

All GQs can be classified on the basis of strand polarity, or the relative orientations of the four strands composing the GQ. There are three distinct categories (Figure 2): 1. All strands can point in the same direction, known as a parallel structure (Figure 2A); 2. Three strands can point in one orientation and one strand in the opposite orientation, known as a (3+1) mixed-hybrid structure (Figure 2B); 3. Two strands can be oriented in one direction while the other two are oriented in the opposite direction, known as an antiparallel structure (Figure 2C-D). Category three can be further divided into two categories depending on the relative orientations of neighboring strands. That is, in an antiparallel arrangement, one neighboring strand can be parallel, leading to an up-up-down-down arrangement (Figure 2C), or both neighboring strands can be antiparallel, leading to an up-down-up-down arrangement, (Figure 2D). The glycosidic conformation of the guanine (i.e., whether it is *syn* or *anti* with respect to the pentose ring) is associated with the 3 classes of strand topologies mentioned above (Figure 1B-C). In a parallel arrangement, guanines are either all *syn* or all *anti*. In a (3+1) mixed hybrid structure, three of the guanines in each G-quartet are *syn* with

the fourth being *anti* or three are *anti* with the fourth *syn*. Finally in an antiparallel structure, two guanines are *syn* and two are *anti* in each G-quartet.

The arrangement obtained by a given GQ sequence can depend strongly on the conditions. Often the choice of monovalent cation and its concentration can influence the topology. In addition, the presence of crowding or dehydrating agents, ligands, or proteins can have an impact on the strand orientations.

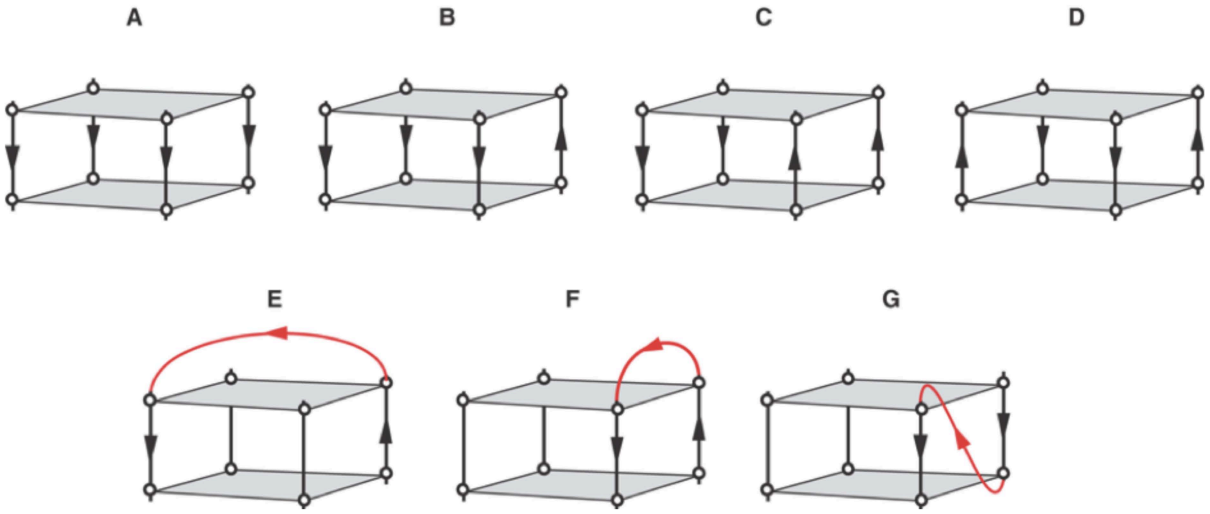


Figure 2. (A–D) Four types of G-tetrad cores: **(A)** parallel, **(B)** (3 + 1) mixed-hybrid, and **(C)** antiparallel (up–up–down–down) and **(D)** antiparallel G-tetrad core (up–down–up–down).
(E–G) Three types of loops (colored red): **(E)** diagonal loop, **(F)** lateral loop and **(G)** propeller loop. Arrows indicate the strand orientations, from 5' to 3' direction. (6)

In intramolecular quadruplexes, there are three main types of loops that dictate the strand topologies listed above (Figure 2E-G). The first is a propeller loop, which connects two adjacent parallel strands; second, a lateral loop that connects two adjacent antiparallel strands; and third, a diagonal loop that connects two antiparallel strands at opposite ends of the GQ (6).

Circular dichroism (CD) spectroscopy produces a signature signal that can be used to distinguish these different GQ topologies, since the structures have an overall right-handed twist (Figure 3). Although CD spectroscopy can only distinguish between homopolar (guanines in adjacent tetrads in the same orientation) and heteropolar stacking (guanines in adjacent tetrads in opposite orientations) of the GQs, these conformations are correlated to the overall topology forms described above (7). Homopolar stacking is correlated to parallel GQ formation, and heteropolar stacking is correlated to antiparallel GQ formation. However, caution must be exercised, as there have been instances in which circular dichroism spectra and other structural methods conflict. The most prominent example of this discrepancy is the HIV-1 integrase aptamer (8,9).

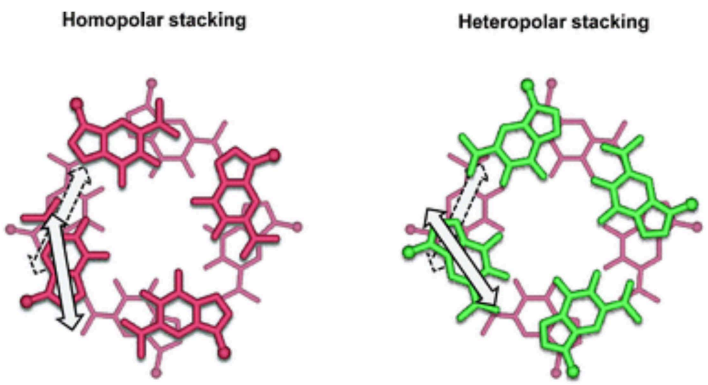


Figure 3. Depiction of homopolar vs. heteropolar stacking of two adjacent G-quartets. These different arrangements of the chromophore result in distinct CD spectra. (7)

Biological Significance

Background: The End-Replication Problem, Telomeres, and Telomerase

The end-replication problem is the inability of the DNA replication machinery to complete synthesis of the lagging strand of DNA. This is because DNA polymerase can only

add bases to a strand of DNA in a 5'—> 3' direction (10-12). Because of the antiparallel nature of DNA and the inability of the DNA to completely separate, only one strand of the template can be continuously replicated (termed the leading strand), while the other strand is synthesized in small 100-200 base pair increments, known as Okazaki fragments. A short RNA primer of ~10-20 nucleotides is needed to begin synthesis of each Okazaki fragment. As the Okazaki fragments are synthesized, the RNA primers are enzymatically replaced with the appropriate DNA bases, and the individual Okazaki fragments are then bonded together in a continuous complementary strand. Note that in order to replace the RNA primer with DNA, the polymerase must have DNA preceding the 5' end of the primer. In other words, the n^{th} RNA primer is only replaced once the $(n+1)^{th}$ Okazaki fragment has been synthesized. Obviously, a problem arises as we approach the very end of the telomere where we cannot synthesize an $(n+1)^{th}$ Okazaki fragment, and thus, the primer cannot be replaced (Figure 4). This problem is encountered with each replication event, and as a result, the chromosomes become progressively shorter until the cell can no longer divide, leading to replicative senescence and eventual cell death, or apoptosis. This limit to the number of times a eukaryotic cell can divide is known as the Hayflick limit (13).

The telomere (from the greek "telos," meaning end, and "mero," meaning part) is a repetitive sequence of oligonucleotides located at the ends of linear chromosomes that serves as a buffer to this shortening. The sequence of vertebrate telomeric DNA is $(TTAGGG)_n$. The telomeres do not encode for proteins, so the progressive loss of these DNA segments is not detrimental to the cell.

Cancer cells are not susceptible to this progressive shortening of the DNA because they adopt various methods of chromosome elongation, which allow them to replicate

continuously. In 80-90% of cancerous cells, an enzyme known as telomerase is expressed and is responsible for telomere maitenance (14). Telomerase is ribonucleoprotein complex that acts as a reverse transcriptase. Telomerase consists of two main components: a telomere reverse transcriptase (TERT) and a telomere RNA component (TERC). Part of TERC provides the RNA template for telomere extension and TERT provides the enzymatic machinery (15). The remaining 10-20% of cancers maintain their telomeres via the Alternative Lengthening of Telomeres (ALT) mechanism. This process is thought to be dependent on homologous recombination between the telomeric sequences of nearby sister chromatids (16).

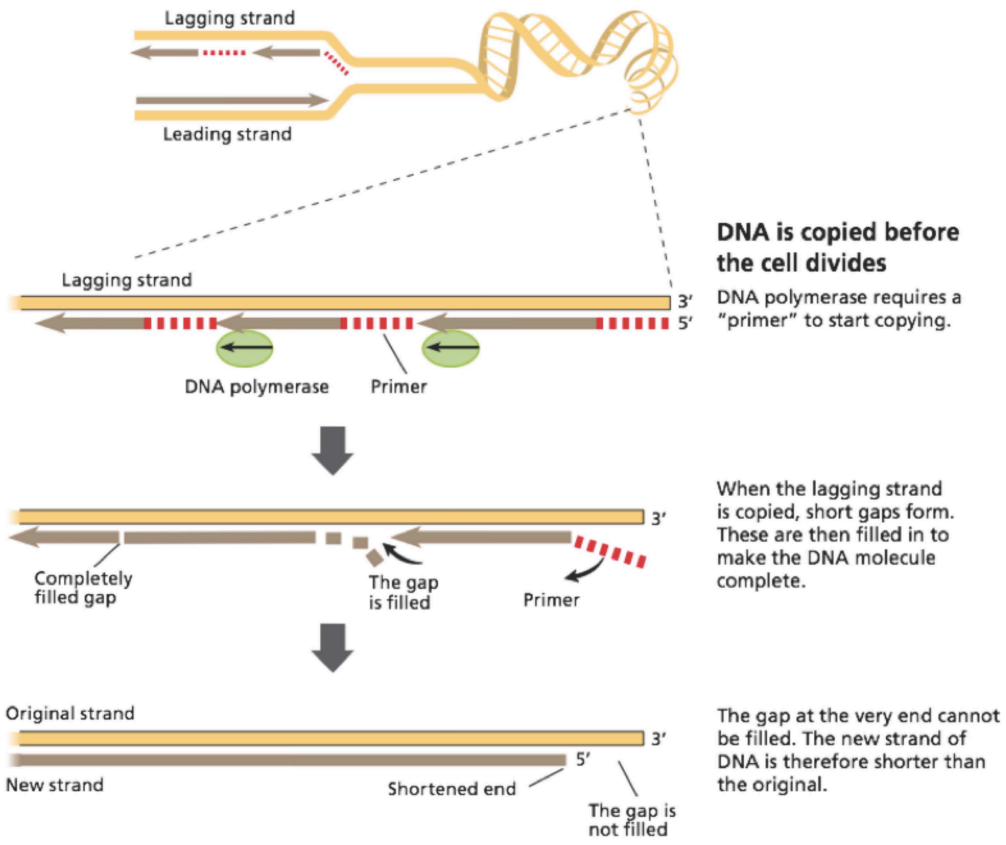


Figure 4. Schematic outlining the "end-replication problem." (12)

Due to the end-replication problem, the template strand for lagging strand synthesis is longer than its complement, and this leads to a 3' single-stranded G-rich region of the

telomeres, known as the G-overhang. In humans, the length of this overhang is 130-210 bases (17). Clearly, this overhang is longer than the 10-20 bases predicted by the end-replication problem, suggesting a mechanism in which the C-rich complementary strand undergoes some additional degradation (17).

In order to prevent the G-rich overhang from being recognized as a double stranded break, a group of proteins, known as the shelterin complex, caps the ends of telomeres. Without this complex, end-to-end fusions of the chromosomes would occur, resulting in widespread genomic instability and, eventually, cell death. The main components of the shelterin complex are telomeric repeat-binding factors 1 and 2 (TRF1 and TRF2), which bind to telomeric duplex DNA, and a protein named protection of telomeres protein 1 (POT1), which binds to the single-stranded g-overhang. With the aid of the shelterin complex, this g-overhang can fold back on itself and invade nearby duplex telomeric DNA, forming a telomere loop (t-loop) structure (Figure 5) (15).

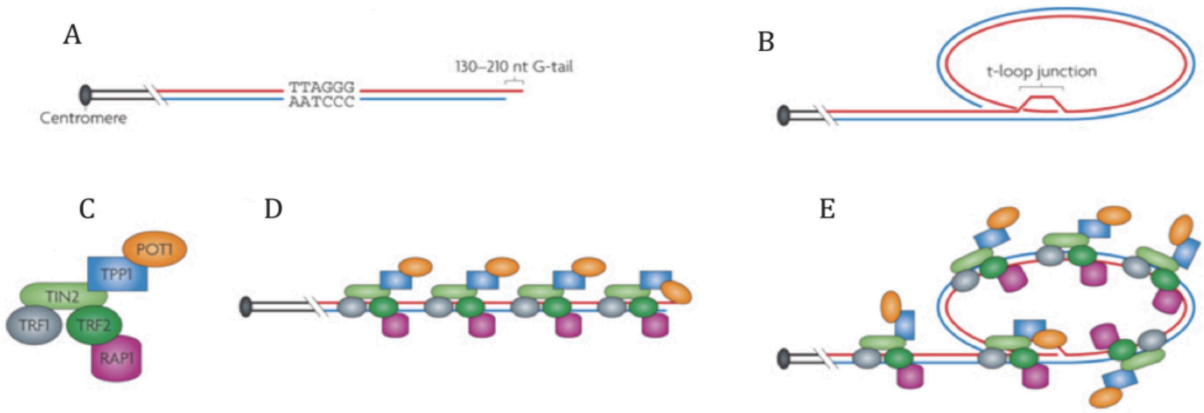


Figure 5. (A) Telomere (4-14 kb in normal human cells) **(B)** t-loop **(C)** Shelterin complex **(D)** Shelterin-bound linear telomere **(E)** Shelterin-bound t-loop (16)

***In vitro* structure of human telomeric G-quadruplex**

Due to the G-rich, single-stranded, and repetitive nature of the G-overhang, one would expect that G-quadruplexes should arise from this chromosomal locus. Indeed, in

isolation, the G-rich strand of many telomeres can form very stable GQ structures. Extensive research has been dedicated to the 22-mer sequence d[$AG_3(T_2AG_3)_3$] (known in this work as Tel22). Because it contains four G-tracts, the minimum required to form single intramolecular GQ, it is the smallest biologically relevant model.

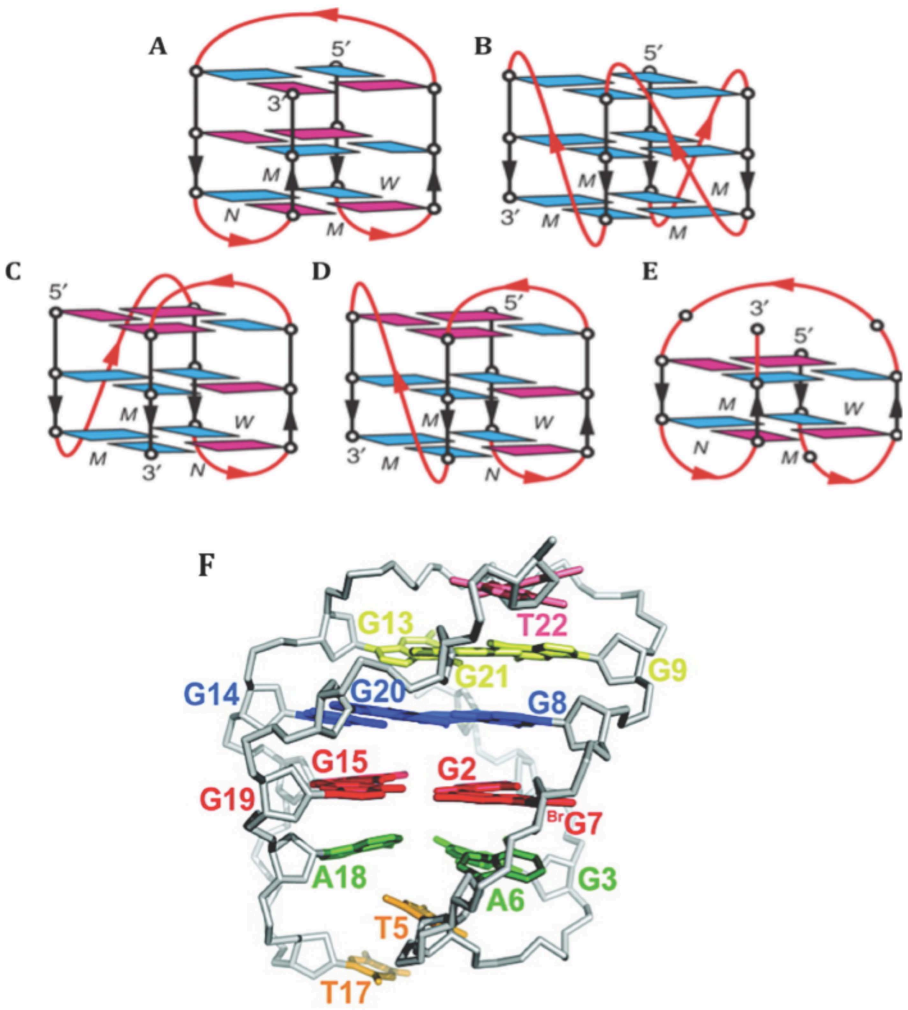


Figure 6. Structures formed by Tel22 under various conditions: **(A)** Antiparallel basket **(B)** Parallel propeller loop **(C)** Form 1, **(D)** Form 2, **(E)** Form 3. (6)
(F) NMR solution structure of Form 3 with stabilizing interactions highlighted by colored nucleotides: G-tetrads (blue, red); Base triples (green, yellow); Base pairs (pink, orange). (24)

Tel22 shows large variation in GQ topology depending on the conditions. In high Na^+ buffer, Tel22 adopts an antiparallel basket GQ (Figure 6A) (18); in crystalline conditions or

in the presence of polyethylene glycol (a crowding or dehydrating agent), Tel22 takes on a parallel-propeller loop structure (Figure 6B) (19); and in K^+, Tel22 likely takes on at least 3 conformations (Figure 6C-E). Due to the conformational heterogeneity of this sequence in K^+, there was heated controversy surrounding its structure. To address this problem, the sequences were altered by changing the nucleotides that flank the four G-runs or by substituting guanines with analogues such as 8-bromoguanine in order to favor only one of the predominant forms (20-24). Taken together, the results suggest a mixture of three structures named Form 1, 2, and 3 (Figure 6C-E). Note that Forms 1 and 2 differ only by the order in which their loops form, both containing two lateral loops and one propeller loop, resulting in (3+1) mixed hybrid structures. Form 3 is especially peculiar because the formation of only two G-tetrads should be thermodynamically disfavored (24). However, it appears to be quite stable due to extensive base-paring and stacking interactions between residues in the loops (Figure 6F) (24).

Function of and Evidence for G-quadruplexes *in vivo*

In biological systems, GQs are most likely to be formed by G-rich DNA sequences that are at least transiently single-stranded. It is unlikely, however, that the human telomeric G-overhang forms a GQ as a long-term, stable structure *in vivo* because POT1, a shelterin protein that binds to the single stranded G-overhang, has been shown to disrupt GQ formation (25). However, it is probable that a GQ forms transiently, as the POT1-TPP1 (another shelterin protein) complex has been shown, by single-molecule FRET-based assays, to slide back and forth along the telomeric DNA, inducing continuous folding and unfolding of the quadruplex structure *in vitro* (26). Also, it is hypothesized that GQs could serve as an

alternative mechanism of telomere capping in the absence of capping proteins. In temperature sensitive mutant *S. cerevisiae* strains lacking Cdc13, a POT1 homologue in yeast, capping defects were suppressed by using a variety of GQ-stabilizing treatments, which include the overexpression of GQ-binding proteins, loss of a GQ-unwinding helicase, and addition of a GQ-specific ligand, N-methyl mesoporphyrin IX (NMM) (27).

Although there are proteins, such as POT1, that disfavor G-quadruplex formation, there are some telomere binding proteins that actually induce GQ formation such as Telomere Binding Protein β (TEBPβ) in *S. lemnae* and Repressor Activator Protein 1 (RAP1) in *S. cerevisiae* (28-30). Furthermore, antibodies raised against GQs in both humans and ciliates have been shown to localize to the telomeres, among other sites in the chromosome (31,32).

In addition to their putative location at the telomeres, G-quadruplexes are also likely to form during DNA replication while the template strands of DNA are separated. GQs would be especially likely to form on the lagging strand, as it is transiently single-stranded due to its discontinuous synthesis. In order to be copied by DNA polymerase, the template strand must be single-stranded; therefore, if GQs do form, helicases must resolve these structures prior to replication. If left unresolved, genome instability results due to double stranded breaks and recombination events that take place once the replication machinery is stalled. A number of helicases have been shown to unwind GQs. However, these studies are mostly qualitative and as a result, it is difficult to asses whether a given helicase is particularly effective at unwinding GQs and whether GQs are the preferred substrate of the helicase (33).

Two particular GQ helicases, Pif1 and FANCJ, are well known for their association with disease and the compelling evidence for their GQ-unwinding capabilities *in vivo*. *S. cerevisiae* Pif1 has been shown to bind to and very actively unwind GQs *in vitro*. A study by Paeschke *et al.* in 2011, revealed, via ChIP-on-chip assay, that 25% of the GQ motifs were Pif1 associated. Furthermore, replication slows especially during s-phase of the cell cycle and on the lagging strand of the replicated DNA (where GQs are most likely to form) in Pif1 deficient cells. Recombination and DNA damage were also prevalent in these cells, with viable cells containing mutations in their G4 motifs (34). Similarly, in the case of FANCJ, a Fanconi Anemia associated helicase, cells from human patients with this disease carrying FANCJ mutations display deletions and mutations in regions of DNA with the potential to form GQs (35).

Another speculated role for GQs *in vivo* is gene regulation. Sequences bearing one or more GQ-motifs are found in the promoters of >40% of human genes, and there is enrichment for these motifs in the promoter regions of proto-oncogenes and various regulatory and developmental genes (36). One of the first well-characterized examples of a GQ-forming regulatory species is the nuclease hypersensitive element III_1 (NHE III_1), upstream of the P1 promoter of the c-MYC oncogene. Quadruplex formation is hypothesized to repress transcription of this gene, which encodes a transcription factor involved in cell proliferation. Indeed, mutation of this regulatory element such that GQ formation is no longer possible leads to a 3-fold increase in the basal transcriptional activity of the promoter, and stabilization of the GQ with a cationic porphyrin ligand, TMPyP4, suppresses c-MYC transcriptional activation (37).

A simple consideration of thermodynamics would dash all hopes of quadruplex formation at promoter sequences, as GQs are typically not nearly as stable as their corresponding duplex form. However, GQ formation is thought to be stimulated by negative supercoiling that is a result of transcription. Indeed, when placed in a supercoiled plasmid (to mimic the effect of transcriptionally induced superhelicity), GQ formation was observed in the c-MYC promoter (38). Moreover, even in the absence of this supercoiling, there are some exceptions. For example, the stability of the c-kit promoter quadruplex is great enough to favor its formation over the competing duplex *in vitro* (39). Additionally, the existence of GQ-binding agents and GQ chaperones such as Rap1 and TBPβ only further serve to support the idea that GQs could arise from duplex DNA *in vivo* under favorable conditions.

Duplex-derived interstrand quadruplex forming potential (ddiQFP)

Another interesting thermodynamic case to consider is whether a bimolecular GQ can form between a strand of DNA and its complement (e.g., 5'-GGGCCCGGGCCC-3' and 3'-CCCGGGCCCGGG-5'). Sequences of this type are neglected in bioinformatics studies examining GQ-forming potential of DNA sequences, as there is little to no evidence supporting (or refuting) the existence of GQs of this form. Recently, Cao *et al.* (2012) showed, using a bioinformatics approach, that sequences with, what they term, duplex-derived interstrand quadruplex forming potential (ddiQFP) may need to be more seriously considered in future studies (40).

A very loose thermodynamic justification may be used to argue that sequences with ddiQFP may in fact be able to form stable GQ structures. First, we must consider a duplex consisting of fourteen base-paired "segments" (Figure 7). If an intramolecular quadruplex

forms on one of these strands, four segments are bonded together, but in exchange, ten are lost (Figure 7A). Of course this argument is overly simplistic, as it assumes that no stabilizing interactions will form on the complementary strand and ignores any external stabilizing (or destabilizing) effects on the GQ (e.g., proteins). If instead we consider a sequence with ddiQFP, there are, again, fourteen base-paired segments with ten segments lost if a GQ is formed (Figure 7B). Note, however, this scheme regards loops and C-runs as separate entities. Alternatively, we could generate a more compact structural form in which the C-runs function as the loops (Figure 7C). In this case, there are only four base-paired segments lost upon quadruplex formation, suggesting that sequences with ddiQFP might be able to form especially stable structures. Furthermore, since the strands of a duplex are essentially tethered to one another in the context of genomic DNA, there is little entropic cost to forming a GQ from ddiQFP *in vivo* and the kinetic barriers should not differ greatly from those of intramolecular GQs. (40)

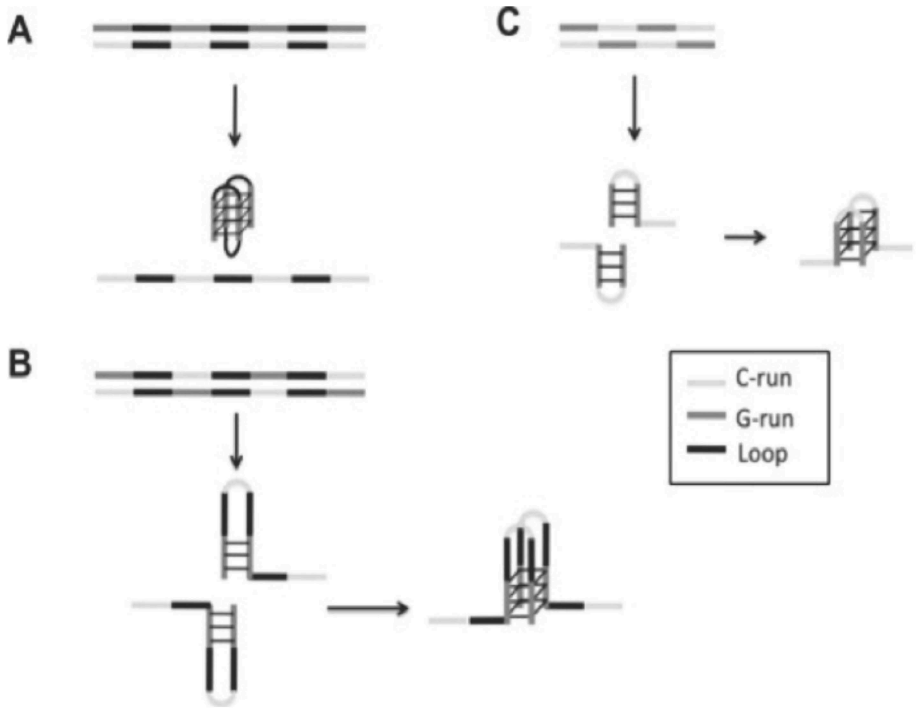


Figure 7. The interconversion between duplex DNA and **(A)** an intramolecular quadruplex, **(B)** a ddiQFP structure, and **(C)** a compact ddiQFP structure (40).

In addition to these arguments, one of the best pieces of evidence that these structures are actually found *in vivo* is their statistically significant association with sites targeted by the Pif1, a G-quadruplex-unwinding helicase (discussed previously) (40).

G-quadruplex Binding Ligands

Since GQs are implicated in the gene regulation of oncogenes and may also form at the telomeres, there is a search for GQ-binding small molecule ligands that could potentially act as cancer therapeutics. One hypothesized mechanism of action for these drugs is via telomerase inhibition because intramolecular GQs cannot serve as a substrate for this enzyme (41). As a result, treatment of cells with these ligands should result in normal rates of telomere attrition and, thus, hamper these cells' proliferative ability. This mechanism is not favored by most researchers in this field because the observed rates of cell death are much faster than those expected by this model (42,43).

Rather, the mechanism supported by the most *in vivo* evidence is one in which telomeres are uncapped and an early, massive DNA damage response results due to double-stranded break (DSB) formation at the telomeres, perhaps, from replication fork stalling. Telomestatin, a GQ-selective ligand, has been shown to provoke a strong dissociation of POT1 and TRF2 from their telomeric sites (44). It is hypothesized that due to the DSBs, the telomere length decreases rapidly, and once a critical length is reached, a cascade of signals activates p53, which, in turn, can trigger apoptosis or senescence (45). Although this effect would seem to influence all cells equally, making GQ-ligands an untargeted drug, it has been demonstrated by several groups that these drugs display enhanced selective uptake by malignant cells (46).

Research in the field is aimed at discovering ligands that both selectively bind and highly stabilize G-quadruplex DNA. Most ligands studied are cationic with a planar, aromatic core. These qualities are thought to endow the ligand with good GQ-stabilizing capabilities because of two main energetic contributions, namely electrostatic attraction to the DNA backbone and efficient π-π stacking between the aromatic core and the G-tetrads of the GQ (47).

There are four major structural motifs that initially gained recognition for their marked abilities to bind GQ DNA: 1) acridines such as BRACO-19 (Figure 8A); 2) perylenes such as PIPER (Figure 8B); 3) bisquinolinium compounds such as Phen-DC_3 (Figure 8C); and 4) porphyrins such as TMPyP4 (Figure 8D). Note that BRACO-19 and PIPER are not cationic as depicted in Figure 8. However, these ligands contain amine appendages that undergo *in situ* protonation at physiological pH (47).

Metallo-organic complexes are also widely used as G-quadruplex ligands because these positively charged metals can endow the ligands with increased cationic nature. Moreover, it is hypothesized that these metals will align with the cation channel, thus optimizing the stacking interactions with the nearby G-tetrad (47). The first reported examples of metallated GQ ligands were metallated derivatives of TMPyP4, which include CuTMPyP4, ZnTMPyP4, and MnTMPyP4 (47,48). Platinated ligands are especially interesting to consider, as both *cis*- and *trans*-$[Pt(NH_3)_2(H_2O)_2](NO_3)_2$ were shown to crosslink with guanines in the 5' G-tetrad of Tel22 (49). A drawback to using such a molecule as a GQ stabilizing agent is non-specificity, as *cis*-$[Pt(NH_3)_2(H_2O)_2](NO_3)_2$ (cisplatin) is a well-documented duplex binding compound and cancer treatment. Therefore, in order to be selective for quadruplexes, these ligands should have an aromatic stacking

component. Ligands of this type should, thus, have both covalent and non-covalent interactions with the GQ through crosslinking and π-π stacking, respectively. Indeed, a hybrid platinum-quinacridine ligand has been designed and was shown to interact with quadruplex-DNA via covalent and non-covalent interactions (50).

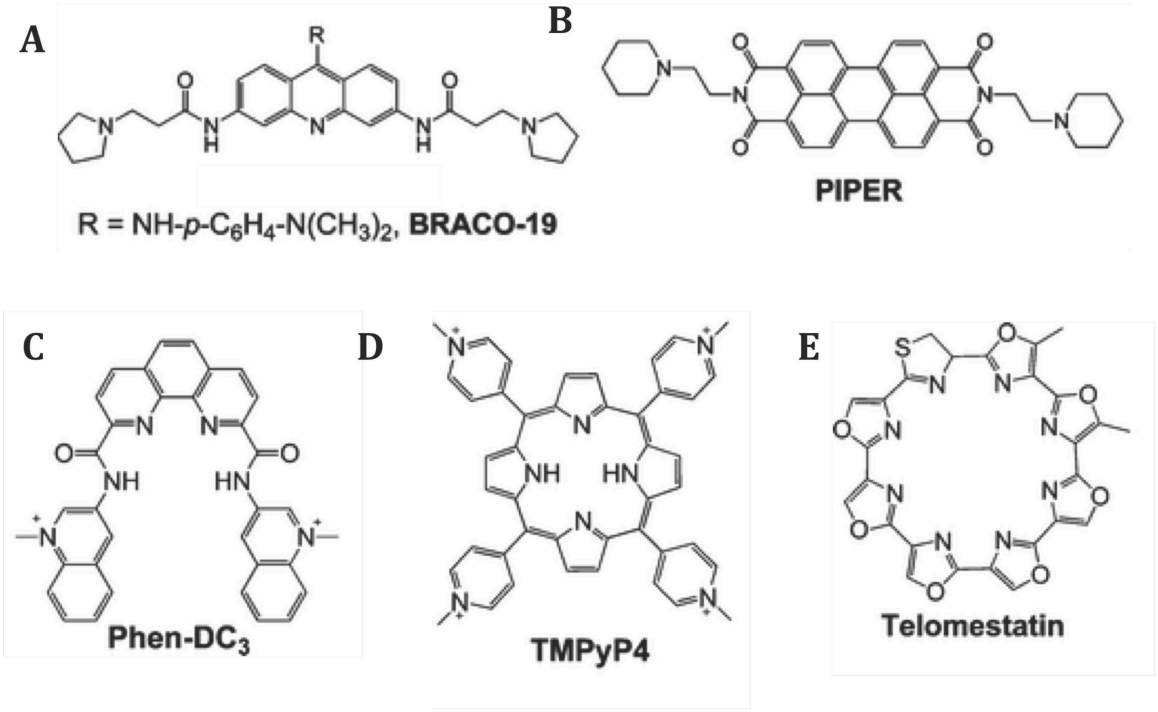


Figure 8. Examples of GQ binding ligands with structural motifs discussed in the text. (46)

It should also be mentioned that the gold standard for GQ-ligands is a neutral, polyheteroaromatic compound named telomestatin (Figure 8E). This molecule is a natural product isolated from *Streptomyces annulatus* with extremely high selectivity for GQ DNA (51). Interest in this ligand was strongly stimulated by its exceptional ability to inhibit telomerase (51). Nevertheless, one major drawback of this ligand is that it is difficult to obtain, and the total synthesis of the ligand is too complex to be feasible on a large scale (47,52,53).

N-methyl mesoporphyrin IX

Although cationic molecules are the primary focus of GQ-ligand research, N-methyl mesoporphyrin IX (NMM) (Figure 9), a negatively charged porphyrin at physiological pH, has gained prominence due to its ability to very selectively bind GQs.

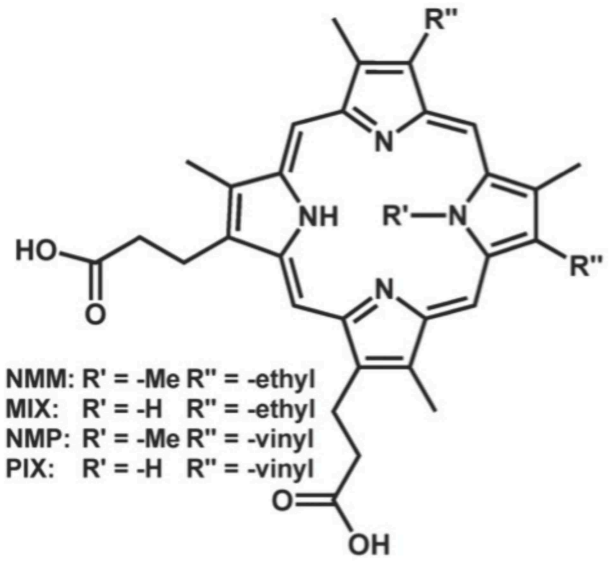


Figure 9. Protoporphyrin IX (PIX) and Mesoporphyrin IX (MIX) along with their core N-methylated derivatives, NMP and NMM, respectively.

NMM was first discovered in 1980 by F. De Matteis *et al.* as a potent ferrochelatase inhibitor, blocking the insertion of iron into protoporphyrin IX (PIX), the terminal step in heme biosynthesis (54,55). Furthermore, the crystal structure of the NMM in complex with ferrochelatase indicates that NMM acts as a transition state analogue due to the out-of-plane distortion caused by its N-methylated pyrrole group (56). Exploiting this property of NMM, Chocran and Shultz (1990) developed catalytic antibodies that could mimic ferrochelatase's metal insertion ability, catalyzing the insertion of Zn^{2+}, Co^{2+}, Mn^{2+}, and Cu^{2+} in mesoporphyrin IX (MIX) (57). In attempt to find a DNA-zyme with similar activity and aptamers that could bind to and distinguish negatively charged porphyrins, Li *et al.* subjected

NMM to *in vitro* selection. The result of this study was a variety of GQ-forming aptamers with submicromolar affinity for NMM (58).

Following this, two groups independently demonstrated that NMM is highly selective for GQ over duplex DNA (59,60). Furthermore, NMM can discriminate GQs from single-stranded DNA, Z-DNA, triplex DNA, and Holliday junctions (structures formed during recombination events) (60). Johnson and coworkers (2008) took advantage of NMM's GQ-selective nature to probe the influence of GQs on gene expression and found that NMM causes significant upregulation of genes containing promoters or open-reading frames (ORFs) with GQ-forming potential (61). In addition, NMM has been used in a variety of other assays such as affinity chromatography for the separation of GQ DNA from other nucleic acid folds (62) and fluorescence techniques for the detection of GQs *in vitro* (59).

Finally, it is interesting to note that commercially available NMM consists of four regioisomers, each differing by the pyrrole on which the core N-methyl group resides. Additionally, there is a pair of enantiomers for each regioisomer, due to the nonplanarity of the N-mehtylated pyrrole. Thus, in total, there are 8 NMM isomers present in a commercially synthesized sample. As of now, it is unknown whether a subset of these isomers interacts more strongly with GQ DNA than the others; it could even be the case that only a subset of these isomers interacts with GQ DNA at all.

Materials and Methods

Materials

Oligonucleotides

Fluorescently-labeled oligonucleotides, F21D and FGD, were purchased from IDT (Coralville, IA); Calf thymus (CT) DNA was purchased from Sigma-Aldrich; and all other sequences were purchased from Midland Certified Reagent Company (TX, USA) (**Table 1**). Fluorescent oligonucleotides were diluted in water to a concentration of ~100 μM and stored at -80°C until use. Prior to annealing, F21D and FGD were diluted 10 fold in the desired buffer. Other lyophilized DNAs were stored at -80 °C until use, at which point they were dissolved in water or the appropriate buffer to ~1 mM and stored at 4 °C. Stocks of Tel22 were further purfied by ethanol precipitation. CT was dissolved in 10 mM lithium cacodylate (LiCac), 1mM Na_2EDTA to a concentration of ~1 mM (in base pairs) and placed on a nutator for 1 week at 4 °C. Before use, the sample was centrifuged to remove any insoluble components, transferred to a clean eppendorf tube, and stored at 4 °C.

The ddiQFP constructs (general structure shown in **Scheme 1**) were engineered by F. Brad Johnson (University of Pennsylvania, Philadelphia, PA). This process was carried out via annealing of the self-complimentary sequence, sddiQFP, and ligation using T4 DNA ligase in the presence of 3 or 10 mM $MgCl_2$ (**Scheme 1**). The efficiency of the reaction was confirmed via gel electrophoresis. The construct was created with T-loops to mimic the tethering of these strands in the context of genomic duplex DNA. T_4 hairpins were used as these created the most stable sddiQFP hairpins prior to ligation.

It is important to note that GQ-forming oligonucleotides were annealed in the desired buffer by heating for ten minutes at 90 °C and slowly cooling to room temperature over

several hours prior to all experiments, unless otherwise specified. After reaching room temperature, DNA was stored at 4 °C until use.

Scheme 1.

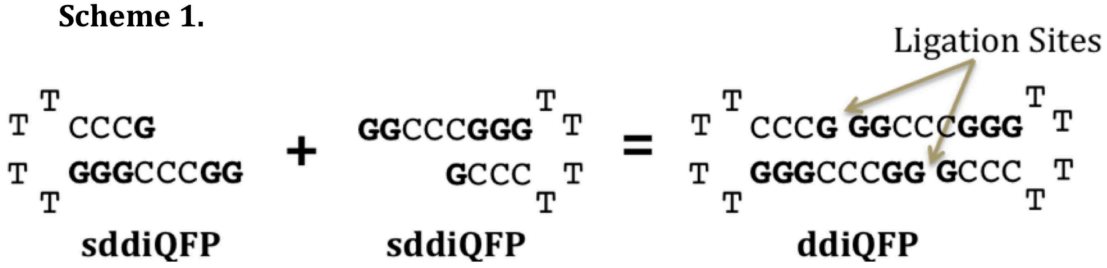


Table 1. Oligonucleotide names, sequences, and extinction coefficients.

Name	Sequence	$\Delta\varepsilon_{260\ nm}$ ($M^{-1}cm^{-1}$) per strand
Tel22	AGGGTTAGGGTTAGGGTTAGGG	228,500
F21D	5′-6-FAM-GGGTTAGGGTTAGGGTTAGGG-Dabcyl-3′	247,600
sddiQFP	GGCCCGGGTTTTCCCG	138,200
sddiQFP2	CCGGGCCCTTTTGGGC	134,600
sddiQFP3	CCTAGGGGTCCCCTTTTGGGGACC	214,000
sddiQFP4	CCTGGGGTCCCCTTTTGGGGACC	200,000
sddiQFP5	CCTAGGGGTTCCCCTTTTGGGGAACC	234,100
dscontrol1	TTTGGGCCCGGGCCCTTT	151,200
dscontrol2	AAAGGGCCCGGGCCCAAA	178,800
T-tailed dscontrol	TTCCCGGGCCCGGGTT	137,400
A	GGA TGG GGT GGG GAG G	165,800
Ac	AGA TGG AGT GGA GAG G	175,400
B	GGG GGA TGC GGG GG	140,400
Bc	AGG AGA TGC AGG AG	152,700
C’	GGA GGG TGG ATG G	135,500
C’c	AGA GGG TAG ATG G	142,800

GQ-Ligands

Porphyrins were purchased from Porphyrin Products, a division of Frontier Scientific (Logan, UT) and stored in solid form at -20 °C (**Table 2**). Stock solutions (~4 mM) aliquots of ligands were made in water (or DMSO, depending on solubility constraints) and stored at

-20 °C. While in use, aliquots were thawed and stored at 4 °C. Care was taken to shield the porphyrins from direct light.

Table 2. GQ ligand names and extinction coefficients.

Name	Abbrev.	Solvent	Δε ($M^{-1}cm^{-1}$)
***N*-methyl mesoporphyrin IX**	NMM	H_2O	1.45 x 10^5 (379 nm) (60)
***N*-methyl mesoporphyrin IX dimethyl ester**	NMME	DMSO	1.16 x 10^5 (410.5 nm) (This work)
***N*-methyl protoporphyrin IX**	NMP	H_2O	1.00 x 10^5 (383 nm) (This work)
Mesoporphyrin IX	MIX	DMSO	1.50 x 10^5 (401 nm)
Protoporphyrin IX	PIX	DMSO	1.70 x 10^5 (408 nm) (63)
5,10,15,20–tetrakis(*N*-methyl-4-pyridyl) porphyrin	TMPyP4	H_2O	2.26 x 10^5 (422 nm) (64)
Zn(II)-5,10,15,20–tetrakis(*N*-methyl-4- pyridyl) porphyrin	ZnT4	H_2O	204,000 (437 nm)

Buffers

With experiments involving GQ-ligands, cacodylate-based buffers enriched with potassium or sodium were used (**Table 3**). The ionic strength of these buffers was kept constant at ~110 mM salt (depending on the pH) by supplementing buffers with LiCl, which does not affect the formation of quadruplexes. This ionic strength was chosen to match that of the cell. The buffers are named for both the identity and concentrations (in mM) of the primary salt in the buffer. Unless otherwise noted, 5K buffer was used for the majority of experiments. Tris was also used in some experiments when a higher pH or alternative conditions were desired.

Table 3. Buffers

Name	Composition
5K	5 mM KCl, 95 mM LiCl, 10 mM LiCac (pH=7.2)
5K (pH=5.8)	5 mM KCl, 95 mM LiCl, 10 mM LiCac (pH=5.8)
5KT	5 mM KCl, 95 mM LiCl, 10 mM Tris (pH=8.6)
50K	50 mM KCl, 50 mM LiCl, 10 mM LiCac (pH=7.2)
100K	100 mM KCl, 10 mM LiCac (pH=7.2)
50Na	50 mM NaCl, 50 mM LiCl, 10 mM LiCac (pH=7.2)
100Li	100 mM LiCl, 10 mM LiCac (pH=7.2)
10Tris	10 mM Tris-HCl (pH=8.0)

Methods

Overview of Methods for the Study of GQ Structure

The secondary structure of any oligonucleotide can be determined using UV-vis ***thermal difference spectra (TDS)*** (65). This technique provides a spectral signature based on the structure of the nucleotide. If the GQ structure is indicated by the TDS, ***circular dichroism (CD) spectroscopy*** can be used to deduce the specific topology of the GQ, again, by comparing the obtained spectra to documented signatures (7). Finally, ***UV-vis and CD melts*** can be used to assess the stability of these GQs to begin to judge their potential significance *in vivo* (66).

Overview of Methods for the Study of GQ-Ligand Interactions

In order to determine the binding affinity of a ligand to a GQ, ***UV-vis or CD titrations*** are carried out, and the absorbance data taken from these spectra are used to extract a binding constant. Other values calculated from these titrations can suggest the binding mode and stoichiometry of the interaction. Normally, however, the stoichiometry of the interaction is more clearly obtained using ***Job's method, or continuous variation analysis*** (67).

The stabilizing abilities of these ligands can be assessed using a variety of techniques. Again, ***CD and UV-vis thermal melts*** are useful tools for the study of GQ stability in the presence of ligands (66). Another technique that allows for the quick and easy assessment of GQ stability with a ligand is ***fluorescence resonance energy transfer (FRET) melts*** in which the oligonucleotide is labeled with a fluorescent probe and quencher (68). Utilizing a 96-well plate format, this technique allows for the quick determination of melting temperatures at various ligand concentrations and in the presence of competitor sequences. A major downside of FRET is that the DNA used must be labeled, which is a much larger monetary expense than purchasing the unlabeled oligonucleotide.

UV-visible Spectrscopy Experiments

1. Titrations of NMM with Tel22

Because the ligands used in this work absorb in the visible range of the electromagnetic spectrum, UV-vis spectroscopy is a useful tool for examining the interactions between a ligand and DNA. The electronic structure of a ligand changes as it is complexed with DNA, leading to a change in the ligand's spectrum. All UV-vis data were taken on a Cary300 spectrophotometer (Varian) equipped with a Peltier-thermostated cuvette holder (temperature accuracy ± 0.3 °C). Spectra were baseline corrected with using the absorbance of the appropriate buffer.

UV-vis titrations are used to determine the binding constant of a ligand (in this case, NMM) to a G-quadruplex. In addition, information on the binding mode can also be inferred from the data. Titrations were carried out by the stepwise additions of ~0.3-0.7 mM Tel22 into a 1 mL solution of ~2-5 μM NMM in a 1 cm path-length methacrylate cuvette at room temperature. The Tel22 for these titrations was annealed with the same concentration (~2-5

μM) of NMM so that the effects of GQ-additions were solely due to the interactions between NMM and Tel22, not the dilution of NMM. The titration was terminated when at least two consecutive additions of DNA yielded identical spectra. Because the binding of NMM is weak, this standard could not always be achieved; so titrations were terminated when the GQ:NMM ratio exceeded 10.

To achieve an equilibrium state, titrations were also carried out using the batch method in which several samples with increasing, predetermined GQ:NMM ratios were prepared and allowed to equilibrate for 30 hours at 30 °C. However, complete isomerization of the Tel22 from mixed hybrid to a parallel fold cannot be induced by NMM at the GQ:NMM ratios needed for a complete titration profile.

Based on UV-vis data the following parameters were calculated: %H, Δλ, and K_a.

The percent hypochromicity was calculated as follows

$$\%H = \frac{\varepsilon_{Free} - \varepsilon_{Bound}}{\varepsilon_{Free}} x100\% \qquad \text{(Eq. 1)}$$

Where ε_{Free} is the extinction coefficient of the free porphyrin, and $\varepsilon_{Bound}=Abs_{bound}/[NMM]_t$ is the maximum absorbance of the final spectrum divided by the total concentration of porphyrin (Note, this assumes that the titration is complete and that all porphyrin is in the complexed form). The red shift (Δλ) is calculated as

$$\Delta\lambda = \lambda_{\max,complex} - \lambda_{\max,NMM} \qquad \text{(Eq. 2)}$$

These two parameters (red shift and percent hypochromicity) are thought to contain information about the nature of the binding of the ligand to the GQ. However, this information cannot be used as an absolute identification of the binding mode because these parameters were only shown to be correlated between ligands with duplex DNA (64).

The binding constant (K_a) was determined via a direct fit to the following equation

$$Y = Y_0 - \frac{x}{C_t}(Y_0 - Y_f) \qquad \text{(Eq. 3)}$$

where Y_0 and Y_f are the initial and final absorbances, respectively, at a given wavelength (normally, $\lambda_{max,complex}$ or $\lambda_{max,NMM}$), x is the concentration of NMM-GQ complex, and C_t is the total NMM concentration.

The concentration of the complex (x) is derived from the equilibrium expression for the chemical equation NMM (C_t) + Tel22 (C_s)→Complex (x)

$$K_a = \frac{[x]}{[C_s - x] \bullet [C_t - x]} \qquad \text{(Eq. 4)}$$

where C_s is the concentration the total concentration of GQ. From this equation, x can be written in terms of C_t (a known constant value), C_s (known variable values), and K_a (an unknown constant). By substituting back into Eq. 3, we can now plot Y vs. C_s and fit the data to obtain K_a using GraphPad Prism software.

2. Continuous Variation Analysis

Continuous variation experiments (Job's Plots) are used in order to determine the binding ratio of two interacting species (67). In this method, the total concentration of the two interacting species (e.g., GQ and ligand) is held constant, but their mole fractions are altered over the course of the experiment.

This method requires two sets of titrations (69). In one set of titrations, Tel22 is titrated into a solution of ligand (in our case, NMM). In the second set of titrations, NMM is titrated into a solution of Tel22. For each of these titrations, the reference cell is titrated in the same way with NMM (with buffer replacing Tel22). This is done to correct for the

absorbance of free NMM. As a result, using this method we obtain absorbance differences and only observe the effects due to the Tel22-NMM interaction. As an alternative method, each sample can be prepared at once at the desired ratios of NMM to Tel22; this is known as the batch method, and it is useful if equilibrating samples for an extended period of time is necessary.

In either method, by determining the mole fraction (using the equation below) at which the greatest absorbance difference occurs, we can calculate the relative ratios of the components of the system in the complex, or the stoichiometry of the interaction.

$$X_{Ligand} = \frac{mol_{Ligand}}{mol_{Ligand} + mol_{GQ}} \quad \text{(Eq. 5)}$$

Note that, often, the peak will not be well-defined due to weak binding between the GQ and ligand; to account for this, tangents are fit to the linear portions of the curve before and after the peak. These lines effectively tell us what the curve would look like if the binding were very tight. Thus, the stoichiometry can be obtained from the mole fraction at which these lines intersect.

Job Plots were executed in both the batch format and titration format. In the titration format, ~18 μM Tel22 was annealed alone (i.e., without NMM), and a stock NMM solution of the same concentration (within 5%) was prepared in 5K buffer. For all experiments with NMM and Tel22, two titrations were performed as described above. In the batch method, 12 samples were made with the same total concentration of both NMM and Tel22 (~17 μM), but with varying molar ratios. These were incubated for 30 hours at 30°C before spectra were recorded in order to ensure complete thermodynamic equilibrium, as NMM isomerizes GQs.

The data were analyzed by plotting the absorbance at 402 nm vs. mole fraction of NMM (X_{NMM}) to produce Job plots (67). 402 nm was chosen because it exhibited the greatest changes in absorbance difference over the course of the titrations.

3. Beer's Law Aggregation Studies

The extinction coefficients and aggregation states of NMP and NMME were determined by adherence to the linear relationship proposed by Beer's Law. Stock solutions of NMP and NMME at 0.867 mM and 0.575 mM, respectively, were prepared using analytical techniques in H_2O and DMSO, respectively. From this stock, 1 mL dilutions of the porphyrins, ranging from 1-10 μM, were prepared in methacrylate cuvettes. Ten samples ranging from 13-50 μM were prepared in a 2 mM quartz cuvette. For these samples, a 300 μL solution of 13 μM porphyrin was made, and the concentration of this solution was increased gradually by stepwise addition of the stock solution. The absorbance vs. concentration of porphyrin was plotted, and from the slope of this line, the extinction coefficients could be determined if there is no aggregation. If the dependence deviates from linear, aggregation is expected. In some cases, even where there is no deviation from a straight line, dimers (or any other oligomers) can be expected to form if their stability is not affected under the given conditions (i.e., the oligomers still form under dilute conditions).

4. Thermal Difference Spectra (TDS)

TDS is used to confirm the secondary structure of nucleic acids in conjunction with its CD signature (65). By subtracting the UV-vis spectra of the DNA in its single stranded state (at high temperatures, above the melting point) from the spectra of the DNA in its folded state (at low temperatures), a signature difference-spectrum results from which information about the secondary structure can be determined.

To carry out this experiment, DNA in the appropriate buffer was heated in a 1 cm or 1 mm cuvette to 90 °C and allowed to melt for 10 minutes (assuming we achieve complete unflolding) after which time the UV-vis spectrum of the sample was taken from 330 to 220 nm. This sample was cooled slowly to ambient temperature over several hours to allow for complete thermodynamic folding, and the UV-vis spectrum was taken again. The high

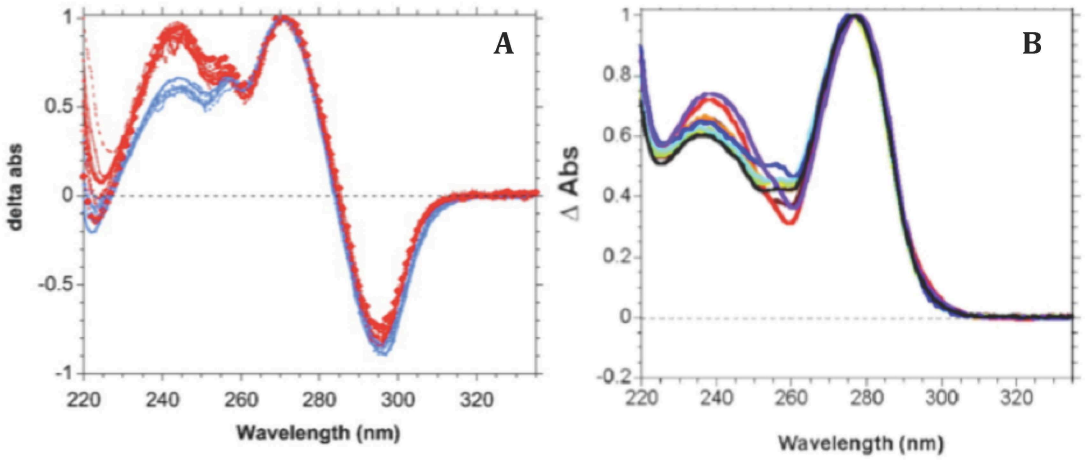


Figure 10. Thermal difference spectra (TDS) of various **(A)** GQ-forming sequences and **(B)** G-C rich duplex DNA sequences. (65)

temperature spectrum was subtracted from the low temperature, and the resultant spectrum was normalized and examined for peaks of interest. Quadruplexes are identified by a negative peak at 295 nm and positive peaks at 243 nm and 273 nm (Figure 10) (65).

5. UV-vis melting studies

The melting temperature of a GQ reports on the stability of the structure under the conditions tested. In order to understand the melting of a G-quadruplex we can monitor the absorbance of the DNA at 295 nm, 273 nm, or 243 nm, where the absorbance spectra exhibits the greatest changes with increasing temperature (see TDS). Often a reference wavelength that exhibits no change in absorbance with temperature (e.g., 350 nm) is also monitored to ensure that the instrument is performing properly.

In these experiments, the temperature change rate was 0.2 °C min^{-1} with an averaging time of 2.00 s, temperature interval step of 1.0 °C, and spectral bandwidth of 2.0 nm. The value of signal at 350 nm was subtracted from each data set before the data were plotted. To extract the melting temperature, the data were fit with a two-state Van't Hoff equation (full equation with derivation is shown in the Appendix) with constant C_p (66). If the data could not be reliably fit, the first derivative curve of the melting data was calculated where the minimum of this curve is the $T_{1/2}$ for the sequence.

Circular Dichroism (CD) Experiments

As mentioned previously, circular dichroism is a useful probe of GQ structure, as GQs are chiral structures with significantly varying structural topologies. Each GQ topology gives a unique CD signature: Parallel GQs have a maximum at ~260 nm and a minimum at ~240 nm, while antiparallel quadruplexes have a maxima at ~290 and 240 nm and a minimum at ~260 nm (7).

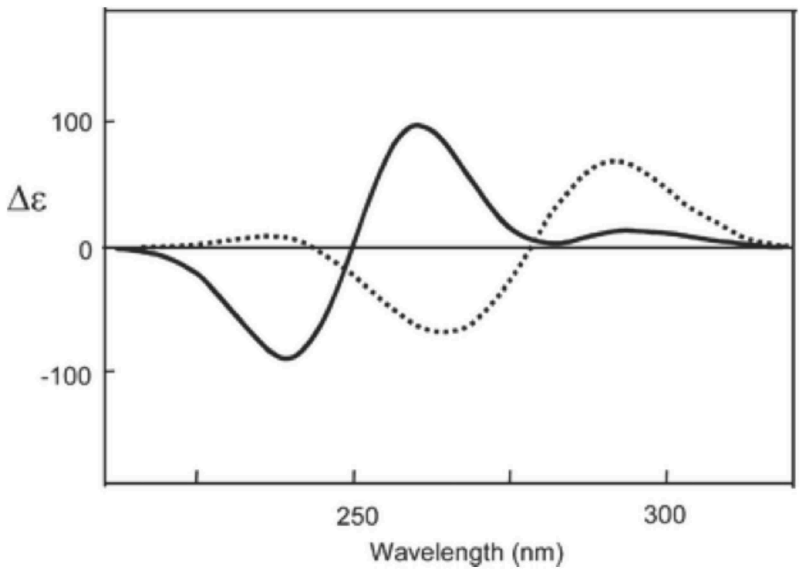


Figure 11. CD spectra of a parallel GQ (solid line) and an antiparallel GQ (dashed line). (7)

CD experiments were performed at 25 °C on an AVIV 410 spectropolarimeter equipped with Peltier thermostated heating unit (temperature accuracy = ±0.3 K). Most spectra were collected from 330 to 220 nm with a 1 nm bandwith and an averaging time of 1 s. For each experiment, two or more scans were taken and then averaged. All data shown here are baseline corrected by subtracting the signal of the cuvette with buffer, zeroed by subtracting the average signal from the first ten data points from the entire spectrum, and converted to molar ellipticity, using the following formula:

$$\Delta\varepsilon = \frac{\theta}{\ell \times c \times 3.298x10^{4}} \qquad \text{(Eq. 6)}$$

where θ is the CD signal (mdeg), c is the concentration of the DNA (M), and ℓ is the pathlength of the cuvette (cm). All CD data are smoothed using the Savitzky-Golay smoothing filter with a 13-point quadratic function in Origin 8.1.

1. CD Annealing Studies of Tel22 with NMM and NMME

CD annealing studies are performed in order to determine if there is an effect of the annealing process on the interaction between Tel22 and the ligand in question. We assume that the annealing process would yield a thermodynamically favored product that may not be achievable by simply adding the ligand to pre-annealed DNA. Samples of Tel22 were diluted in the appropriate buffer to a concentration of ~5 μM. Two samples of Tel22 without ligand were annealed alone, and a third sample was annealed in the presence of the ligand in a 2:1 ligand:DNA ratio. After annealing, the ligand was added to one of the native Tel22 samples, achieving a ligand:DNA ratio of 2:1. CD spectra were recorded for all three samples.

In the case of NMM or NMME these samples were allowed to equilibrate for 30 hours at 30 °C, and CD spectra were recorded again.

2. CD Melting Studies of Tel22 with NMM and NMME

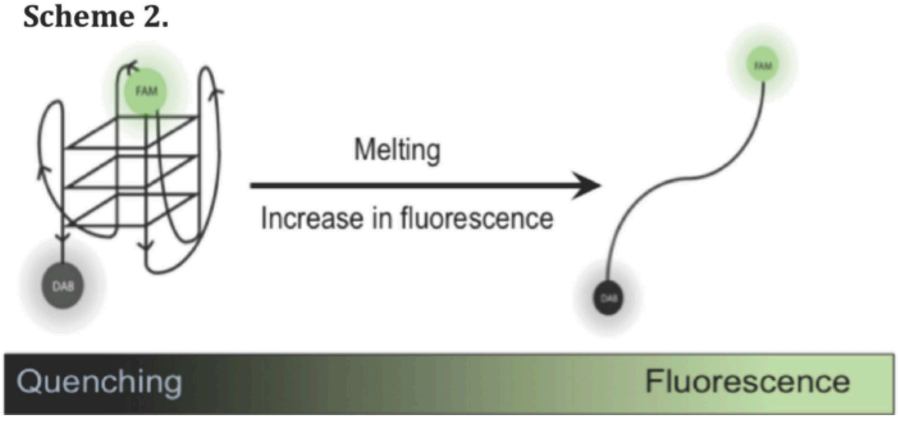
As an alternative method to UV-vis melting, CD melts are also used to examine the stability of GQs with various ligands or under different conditions. The melting of GQs was monitored at 264 and 295 nm (see Figure 11). Samples were heated from 10 to 95 °C with 1 °C step, an overall temperature change rate of 0.27 °C min^{-1}, maintained at 95 °C for less than 5 min, and then the temperature was decreased to 10 °C at the same rate with 0.5 min equilibration time, 4 second averaging time, and 2 nm bandwidth. Again, either the first derivative or a two state Van't Hoff equation with constant ΔC_p were used to fit the data and obtain the melting temperature.

Fluorescence Resonance Energy Transfer (FRET) melting studies

Like CD melts, FRET melts allow us to probe the stability of quadruplexes with varying ligand concentrations under varying conditions. But it provides an additional advantage; the selectivity of the ligands for GQ vs. ds DNA or selectivity for different quadruplex folds can be easily assessed using this technique (68). In these experiments, the GQ-forming DNA is tagged with a fluorphore on one end and a quencher at the other. If a GQ is formed, the fluorophore and quencher will be in close proximity, leading to high FRET and low fluorescence (**Scheme 2**). Conversely, if a GQ is not formed, then the probes will be distant from one another, leading to low FRET and high fluorescence. Since FRET, can be used as a molecular ruler (i.e., it is very sensitive to distances) we can monitor processes in which the GQ progresses from a folded to unfolded state, such as melting.

In our experiments, annealed fluorescent oligonucleotides (F21D, Table 1) were diluted with 5K buffer to a concentration of 0.25 μM. DNA and ligands were added to PCR tubes such that the final concentration of the oligo was 0.2 μM with a total volume of 50 μL. For stabilization assays, the final concentration of ligand ranged from 0.4 μM to 4 μM. Ligands in DMSO were diluted from their 1 mM stock concentration down to 20 μM using the appropriate buffer. For competition studies, an 800 μM stock solution of ds or parallel GQ competitors were added to wells with 0.2 μM F21D and 0.4 μM ligand for the desired competitor:GQ ratio, which ranged from 24 to 480 equivalents.

Scheme 2.



This scheme demonstrates the theory underlying the data obtained from FRET melting studes. Figure prepared by J. Nicoludis.

The typical setup for a stabilization assay is shown below:

Well	Description	V of 0.25 μM F21D (μL)	V of 5K Buffer (μL)	V of 20 μM NMM (μL)
1	Blank	0	50	0
2	F21D	40	10	0
3	2eq NMM	40	9	1
4	5eq NMM	40	7.5	2.5
5	8eq NMM	40	6	4
6	12eq NMM	40	4	6
7	16eq NMM	40	2	8
8	20eq NMM	40	0	10

The typical setup for a competition assay is shown below:

Well	Description	V of 0.25 μM F21D (μL)	V of 5K Buffer (μL)	V of 20 μM NMM (μL)	V of 800 μM CT (μL)	V of 80 μM CT (μL)
1	Blank	0	50	0	0	0
2	F21D	40	10	0	0	0
3	8eqNMM	40	6	4	0	0
4	8eqNMM, 24eqds	40	3	4	0	3
5	8eqNMM, 80eqds	40	5	4	1	0
6	8eqNMM, 160eqds	40	4	4	2	0
7	8eqNMM, 480eqds	40	0	4	6	0
8	F21D, 80eqds	40	9	0	1	0

When NMM was used, samples were allowed to equilibrate at 30 °C for 30 hours, and mixtures with all other ligands were allowed to equilibrate at 4 °C for 1 hour. The fluorescence at 519 nm of 6-FAM dye was measured by an MJ research DNA Engine Real-Time PCR machine when excited at 496 nm. This instrument is available in Nick Kaplinsky's biology research lab at Swarthmore College. Each well was made in duplicate each time an experiment was performed, and all experiments were preformed three times.

In all scenarios (stabilization and competition with both ligands), the following protocol was used for the melting:

1. Incubate at 15 °C for 5 minutes.
2. Melt from 15.0 °C to 95.0°C, reading every 1.0 °C, 1 °C min^{-1}.
3. Cool from 95.0 °C to 15.0°C, reading every 1.0 °C, 1 °C min^{-1}.

All blanks and baselines were automatically subtracted from the data for the wells after the melt using software provided with the instrument. The fluorescence was normalized between 0 and 1. Duplicate wells were then averaged, and the resulting melting curves were smoothed. First derivatives of these averages were used to determine the melting points (T_m) by manually picking the maxima of the resultant curves (with an associated error of ±0.5 °C).

High Performance Liquid Chromatography (HPLC) Separation of NMME Isomers

Isomers of N-methyl mesoporphyrin IX dimethyl ester (NMME) were separated using a silica column (put dimensions here) with the following mobile phase composition: 48.5% THF, 48.5% hexanes, and 3% MeOH. This allowed separation into two groups—species with an N-Me group on an ethyl-substituted pyrrole and species with an N-Me group on a pyrrole substituted with a methyl propionate group. NMR was be used to assess the successfulness of the separation. Note, this experiment was performed by Professor Alison Holliday of Swarthmore College (Swarthmore, PA).

^{1}H-Nuclear Magnetic Resonance (NMR) Spectroscopy

In order to assess the success of the separation of the NMME regioisomers, NMR spectroscopy was carried out. All NMR were taken on a Bruker Avance DRX-400 MHz spectrometer. All spectra of porphyrins were taken in deuterated DMSO from -5.5-14 ppm. Spectra were Fourier transformed and automatically phase corrected.

Mass Spectrometry (MS)

To determine the purity of the separated NMME samples, electrospray ionization mass spectrometry (ESI-MS) experiments were carried out. Experiments were done via direct injection using an Agilent 1100 Series LC-MS. Prior to injection, 1-mL samples in DMSO were diluted 100-fold in isopropanol and spiked with glacial acetic acid.

Results and Discussion

N-methyl mesoporphyrin IX (NMM) as a G-quadruplex Ligand

N-methyl mesoporphyrin IX (NMM) is an exceptionally selective GQ ligand with the ability to discriminate GQ structures from almost any other DNA conformation (59,60,70). Until recently, however, this molecule was largely ignored by scientists because it was presumed to have a low binding affinity to GQs. This argument is based primarily on electrostatic considerations: NMM has two propionic acid side chains that should be deprotonated at physiological pH; therefore, Coulombic repulsion between the DNA backbone and the side chains should cause a weakened interaction.

Although a large GQ binding affinity would be favorable for a small molecule that could serve as a potential pharmaceutical agent or biotechnological tool, the ability to stabilize the GQ is, in fact, a more important factor, as a stabilizing effect would prevent unwinding of GQs by helicases or destabilization by normal cellular processes. It is important to note that there is no direct correlation between binding affinity and stabilization of GQs. Additionally, selectivity is, in actuality, a more important factor to examine, as duplex DNA is far more ubiquitous in the cell than GQ DNA, and the off-target effects of duplex binding could be detrimental. In reality, NMM would likely never be used as a human anticancer drug, as it is a potent inhibitor of ferrochelatase, the enzyme that inserts iron into protoporphyrin IX (54,55). However, the molecular basis of NMM's selectivity certainly warrants investigation, as a deep understanding of the origin of this effect could inform future work on rational structure-based GQ ligand drug design.

Here, I summarize the major findings of this work, which culminated in two publications entitled "Interaction of human telomeric DNA with N-methyl mesoporphyrin

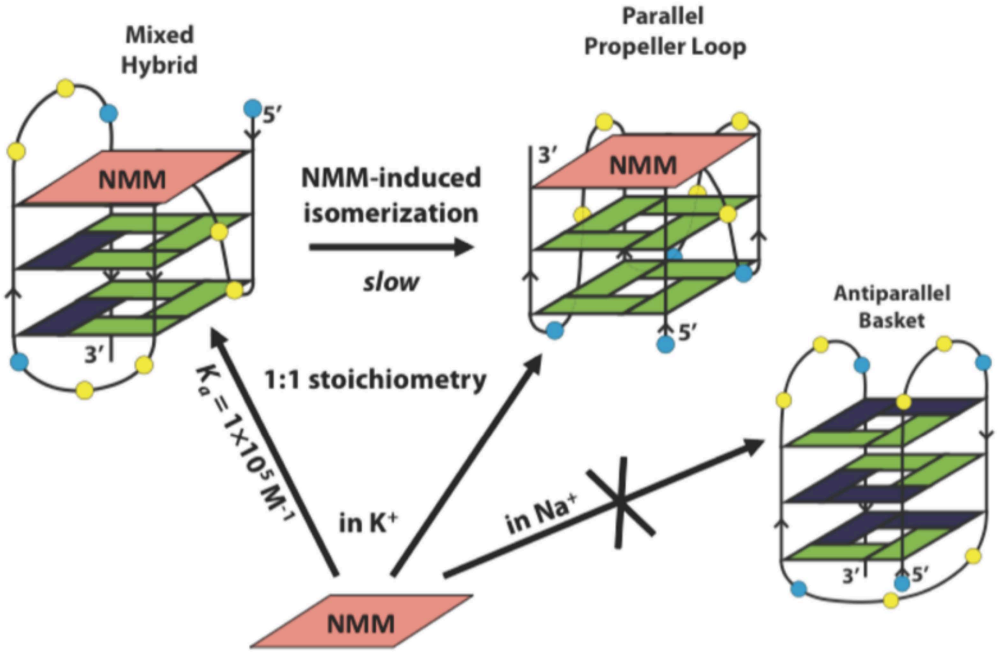
IX" and "Optimized end-stacking provides specificity of N-methyl mesoporphyrin IX for human telomeric G-quadruplex DNA." Both publications are appended for convenience.

Interaction of human telomeric DNA with N-methyl mesoporphyrin IX (70)

This work is a detailed spectroscopic examination of the Tel22-NMM interaction that is summarized in **Scheme 3**. On the whole, the experimental work performed here was a joint effort with John Nicoludis (class of 2012) who conceived the work and started its investigation prior to my joining the lab. My contributions led to the data in Figures 2, 4, and 6 along with Supplementary Figures S1, S2, S6, S12, and S14.

One of our primary goals in this work was to investigate the validity of the previously untested claim that the binding of NMM to Tel22 should be quite weak. Using UV-vis titrations, we determined a moderate binding constant (K_a) of 1.0×10^5 M^{-1}, and using Job's method of continuous variation, the binding stoichiometry was revealed to be 1:1. Furthermore, in FRET melting assays, NMM demonstrated strong stabilization of F21D, and in the presence of various duplex competitors, this stabilization did not decay, confirming previous findings on NMM's selectivity (60). Interestingly, mesoporphyrin IX (MIX), the unmethylated form of NMM, showed no thermal stabilization of Tel22, highlighting the essential nature of the N-methyl group in this interaction. Moreover, in Na^+ buffer, in which Tel22 adopts an antiparallel conformation, NMM was unable to stabilize the GQ, suggesting that NMM can discriminate between GQ topologies. Indeed, circular dichroism (CD) melts with various biologically relevant GQ-forming sequences revealed that this selectivity applies broadly (i.e., not only to Tel22). Thus, NMM is the second example in the literature, after TOxaPy (71), of a ligand that exhibits a strong preference for a specific GQ structure.

Finally, NMM was able to induce a conformational switch in Tel22 from mixed hybrid to the all-parallel propeller loop structure in low K^+ buffer. This isomerization did not occur in Na^+ buffer, as NMM is unable to bind to this structure and there is no population of parallel quadruplex in solution.



Scheme 3. Summary of "Interaction of human telomeric DNA with N-methyl mesoporphyrin IX." Figure prepared by Jack Nicoludis.

Note that NMM likely does not play an active role in the isomerization process, as a protein with a similar function might. Rather, NMM's effect is probably indirect, shifting the equilibrium from a mixed hybrid Tel22 topology to the parallel structure. It is interesting to consider that, in order to shift this equilibrium, a parallel structure must pre-exist in solution. If this is true, this is one of the only indications that such a structure is formed in K^+ buffer, adding to the family of structures speculated to form in this buffer (as depicted in Figure 6C).

Alternatively, while the structures in solution equilibrate, it is possible that NMM could bind a partially unfolded structure and template a parallel GQ by more favorable π-π stacking interactions with an all *anti* G-quartet. In essence, NMM would function as a catalyst, stabilizing an important transition state *en route* to the parallel topology, which must be both kinetically and thermodynamically unfavorable without NMM (by this scheme).

Optimized end-stacking provides specificity of N-methyl mesoporphyrin IX for human telomeric G-quadruplex DNA (72)

This work examines the X-ray crystallographic structure of the Tel22-NMM complex (72) (Figure 12), addressing some unanswered questions from our previous publication (70). For example, how does NMM bind to Tel22? Why is NMM selective for G-quadruplexes over duplex DNA? What role do the N-methyl group and side chains play in this interaction?

The crystals of the Tel22-NMM complex were grown by John Nicoludis (Class of 2012, however the crystals were grown from the condition that I designed and prepared), and diffraction data was collected by Phillip Jeffery (Princeton). The structure was solved by Stephen Miller (Swarthmore), and theoretical calculations were carried out by Paul Rablen (Swarthmore). I am responsible for all of the spectroscopic experiments, which appear in Figure 7, along with supplementary Figures S3, S8, and S9 and Table S3.

The structure of the Tel22-NMM complex was solved in two space groups, P6 and $P2_12_12$, at 2.15 and 1.65 Å resolution, respectively. The latter is the highest resolution structure of the human telomeric GQ to date. The structure of the GQ itself does not show much variation from the native, unbound Tel22 (19). It is NMM's positioning and conformation that were the most remarkable aspects of this structure (Figure 12); the

geometry of NMM's macrocycle becomes distorted to optimize its interaction with the 3' G-tetrad where it binds via end-stacking, and in addition, NMM's core methyl group is positioned directly over the potassium ion channel, avoiding steric clashes with the tetrad. Due to the absence of such a channel in duplex DNA, it is hypothesized that NMM is sterically hindered from interacting with this structure, leading to its exceptional selectivity. In the absence of the N-methyl group, the distortion of porphyrin ring becomes thermodynamically costly, revealing the reason MIX is unable to stabilize GQs as we showed previously with FRET (70). Spectroscopic data, including UV-vis titrations at various pHs, FRET melting studies with the dimethyl ester derivative of NMM (NMME), and CD melts at high ionic strengths, reveal that the negative charge of the propionic acid side chains likely does not have an antagonistic effect on the interaction.

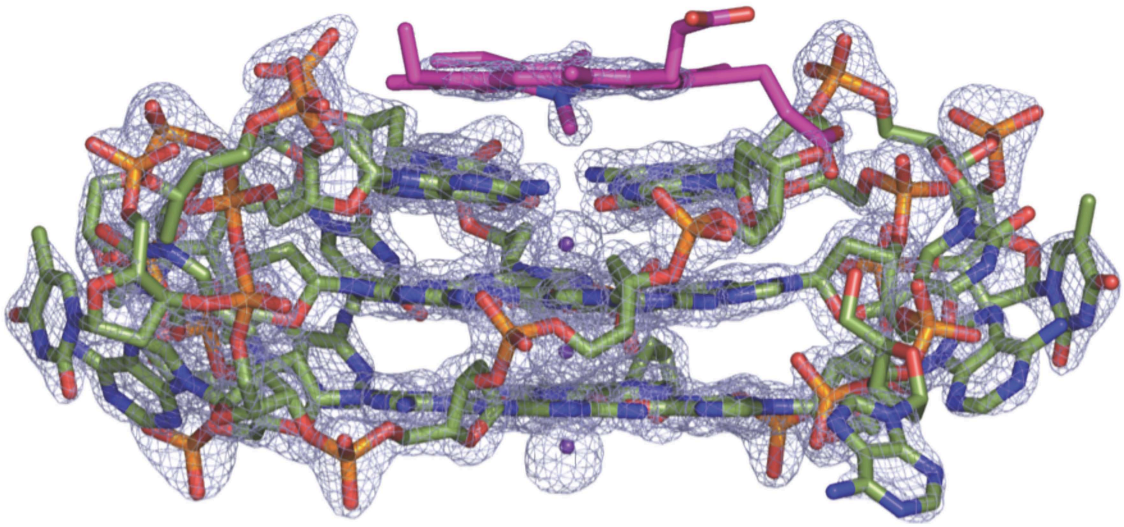


Figure 12. Side-on view of the Tel22-NMM electron density map (at 2σ) with the cartoon model for the structure. (72)

This interpretation of the effect of negative charge presented in this work is likely an overly simplistic (though accurate) view of the effect of the propionic acid side chains. As

previously noted, there is no direct relationship between binding affinity and thermal stabilizing ability of a GQ ligand. Thus, we must treat these effects separately when considering the effect of charge on the interaction.

If we first consider binding affinity, it appears as though the negative charge of the propionates decreases the affinity of NMM for Tel22, as expected. First, UV-vis titrations reveal decreasing binding constants as pH increases, or as the side chains become more deprotonated: $(1.2 \pm 0.1)\times10^5$ M^{-1} (pH=5.8), $(1.0 \pm 0.3)\times10^5$ M^{-1} (pH=7.2), and $(0.3 \pm 0.1)\times10^5$ M^{-1} (pH=8.6). Similarly, with increasing ionic strengths (shielding of negative charge), NMM exhibits a tighter binding to Tel22, as evidenced by the decrease in absorbance of free, unbound NMM at a constant NMM:GQ ratio (Figure 13).

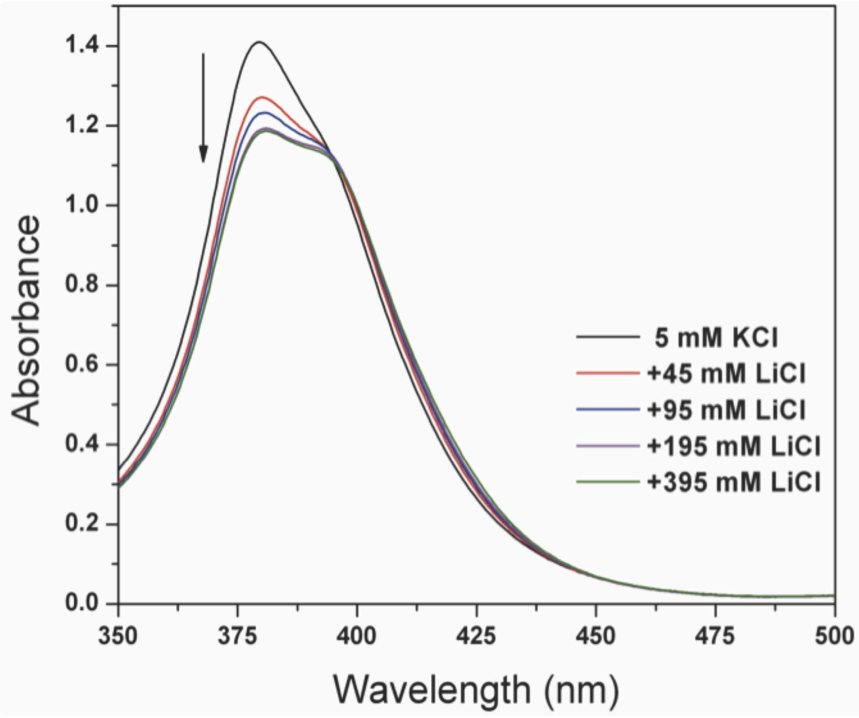


Figure 13. UV-vis spectra of 10.0 μM NMM with 5.0 μM Tel22 at varying ionic strengths indicated in the figure legend. As the ionic strength increases, the amount of free porphyrin decreases, as indicated by the decreasing absorbance at 379 nm.

If we instead consider the melting temperature of the GQ to be a measure of the strength of the interaction, then we can ascertain no clear answer as to the effect of charge on the interaction; there is no obvious trend in melting temperature of the Tel22-NMM complex with increasing ionic strength (shown in Ref. 72, Figure S9), and the use of NMME does not lead to increased thermal stabilization relative to those produced by NMM (shown in Ref. 72, Figure 7). On the contrary, the relative stability of the NMME-Tel22 complex appears to be slightly lower than that of the NMM-Tel22 complex. Although this result is completely counter to our electrostatic predictions, it is possible that the potential formation of a hydrogen bond between the GQ backbone and the propionate group (shown in Ref. 72, Figure S6) is prevented by the inability of NMME to become protonated, leading to a decrease in stability. If we assume that a hydrogen bond does indeed form, and if we assume that this hydrogen bond is well-formed at pH=5.8 and completely disrupted at pH=8.6, we can estimate the energy of this hydrogen bond by calculating $\Delta\Delta G^0$ via the following equation:

$$\Delta\Delta G^0 = -RT\ln\frac{K_{pH=5.8}}{K_{pH=8.6}}$$

Using the K_a values cited above, we obtain $\Delta\Delta G^0$ of approximately -3.2 kJ/mol, an appropriate value for hydrogen bond formation (73).

It is important to note that the conclusions based on the affinity constants and thermal stabilities are not irreconcilable. It is very likely that the small increase in binding affinity (e.g., at high ionic strengths) is too weak to give rise to noticeable changes in melting temperature.

NMM Regioisomers

A close examination of the electron density for the structure shown in Figure 12 reveals that NMM's side chains are not well resolved. Additionally, the temperature B-factors for NMM show that the porphyrin core has a minimal amount of movement, while its peripheral substituents are more disordered (Shown in Ref. 72, Figure 5). Remarkably, the N-methyl group of the porphyrin is very well defined, indicating that it docks at a unique position with respect to the quadruplex core.

As previously mentioned, NMM exists as a racemic mixture of eight isomers: four regioisomers, which differ by the position of the N-methyl group, each of which has an enantiomer pair due to the out-of-plane protrusion of the N-methyl group. If each of these isomers binds to Tel22 (with a localized methyl group), it is clear that the porphyrin side chains will not be resolved in the crystal structure. And indeed clear electron density is only observed for the first peripheral carbon atom that is present in each isomer. As an alternative explanation, it is possible that one isomer of NMM binds to Tel22, but due to thermal motion, the side chains remain unresolved in the crystal structure. This stereospecificity was observed in the NMM-ferrochelatase structure (56). In order to address these uncertainties, our lab has begun the separation and characterization of these isomers.

Following the protocol provided by Ortiz de Montellano *et al.* for the separation of the N-methyl protoporphyrin IX dimethyl ester (NMPE) isomers (74), the isomers of NMME were separated using normal phase high-performance liquid chromatography (HPLC) using a polar amino cyano column and 48.5/48.5/3 hexane/tetrahydrofuran/methanol mobile phase composition (Figure 14). Significant overlap between peaks 1 and 2 and between peaks 3 and 4 is visible on the chromatogram. Therefore, "pure" compounds were obtained by selective

collection of the leading or trailing edge of the peak. This work was performed by Professor Alison Holliday (Swarthmore). By analogy with the previously documented separation of the NMPE isomers, peaks 1 and 2 likely correspond to isomers with the N-methyl group on the

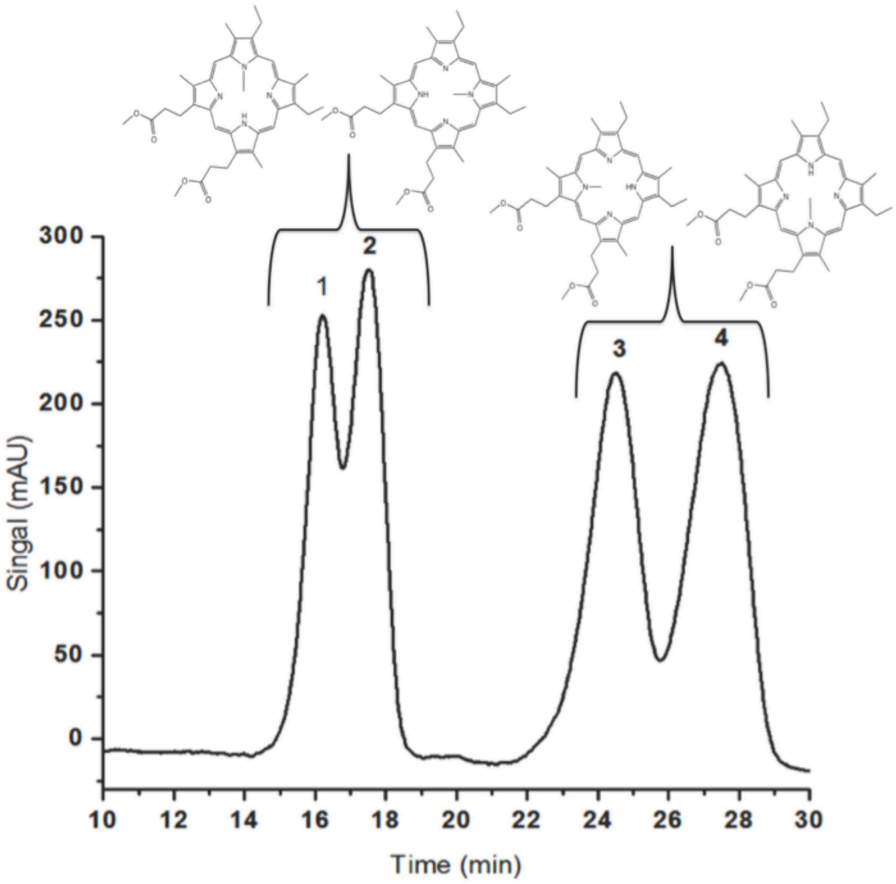


Figure 14. HPLC chromatogram for the separation of the four isomers of NMME. Speculated elution order of isomers is shown above the peaks of the chromatogram.

ethyl-substituted pyrroles of the porphyrin, while peaks 3 and 4 correspond to isomers with the N-methyl group on the ester-substituted rings (Figure 14) (74). Due to the difficulties hydrolyzing the separated NMME isomers (to obtain the respective NMM isomers), we instead focused on characterizing the NMME isomers with the assumption that these isomers should behave similarly to those of NMM, as demonstrated in Ref. 72, Figure 7.

As a first step in characterization, UV-vis spectra of the four separated isomers and the commercial mixture were recorded in DMSO. Absorption spectra of the isolated isomers 1, 3, and 4 and the commercial mixture in DMSO exhibit a Soret maximum at 410 nm (Figure 15). Oddly, the Soret (located at 402 nm) and Q-bands (Figure 4, inset) of isomer 2 differ quite significantly from the other isomers. And notably, isomer 1 exhibits a shoulder at this wavelength, indicating that there could be cross-contamination of these species, a distinct possibility given the overlap in the chromatograms. It may also be the case that

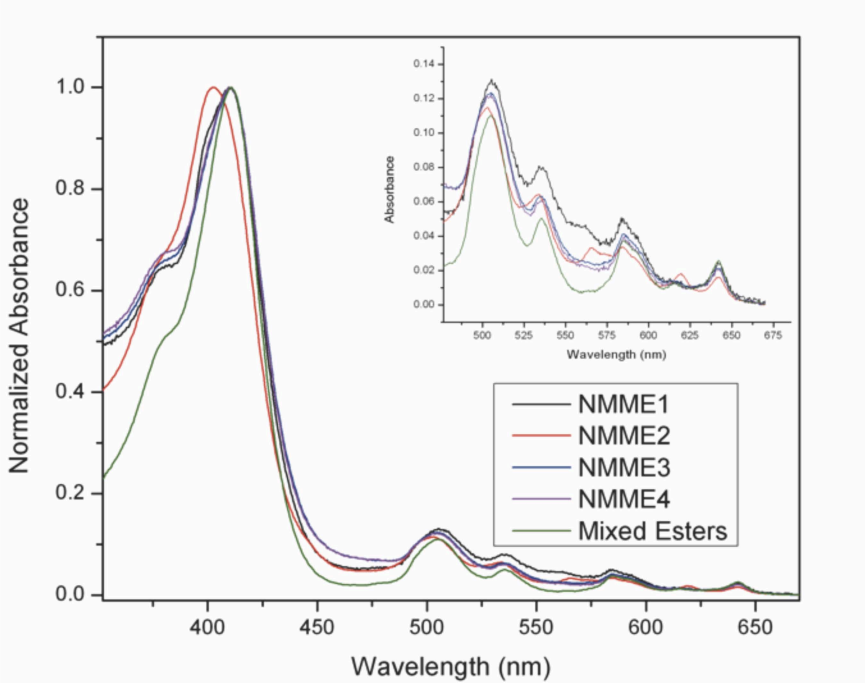


Figure 15. Electronic absorbance spectra of the four separated NMME isomers along with the spectrum of the commercial mixture of isomers. The Soret of isomers 1, 3, 4 and the mixture of isomers is located at 410 nm, while the Soret of isomer 2 is located at 402 nm. The inset is a magnification of the Q-band region.

isomers 2 simply has a different absorbance spectra than isomers 3 and 4. However, this seems less likely, as the substituents on the periphery should not have a very large impact on the electronic structure of the porphyrin ring.

To investigate isomer 2's anomalous absorption spectrum and to provide more concrete evidence for the separation of the NMME isomers, ^{1}H-NMR and electrospray

ionization mass spectrometry (ESI-MS) experiments were undertaken. For the commercial mixture and each isomer, mass spectra appeared identical with one major peak at m/z 609.2 (See Appendix for MS of Isomers), which corresponds to the exact mass of protonated NMME (Figure 16). Other large peaks located at m/z 101.0 and 179.0 correspond to DMSO+Na^+ and 2DMSO+Na^+, respectively; these are simply solvent derived peaks.

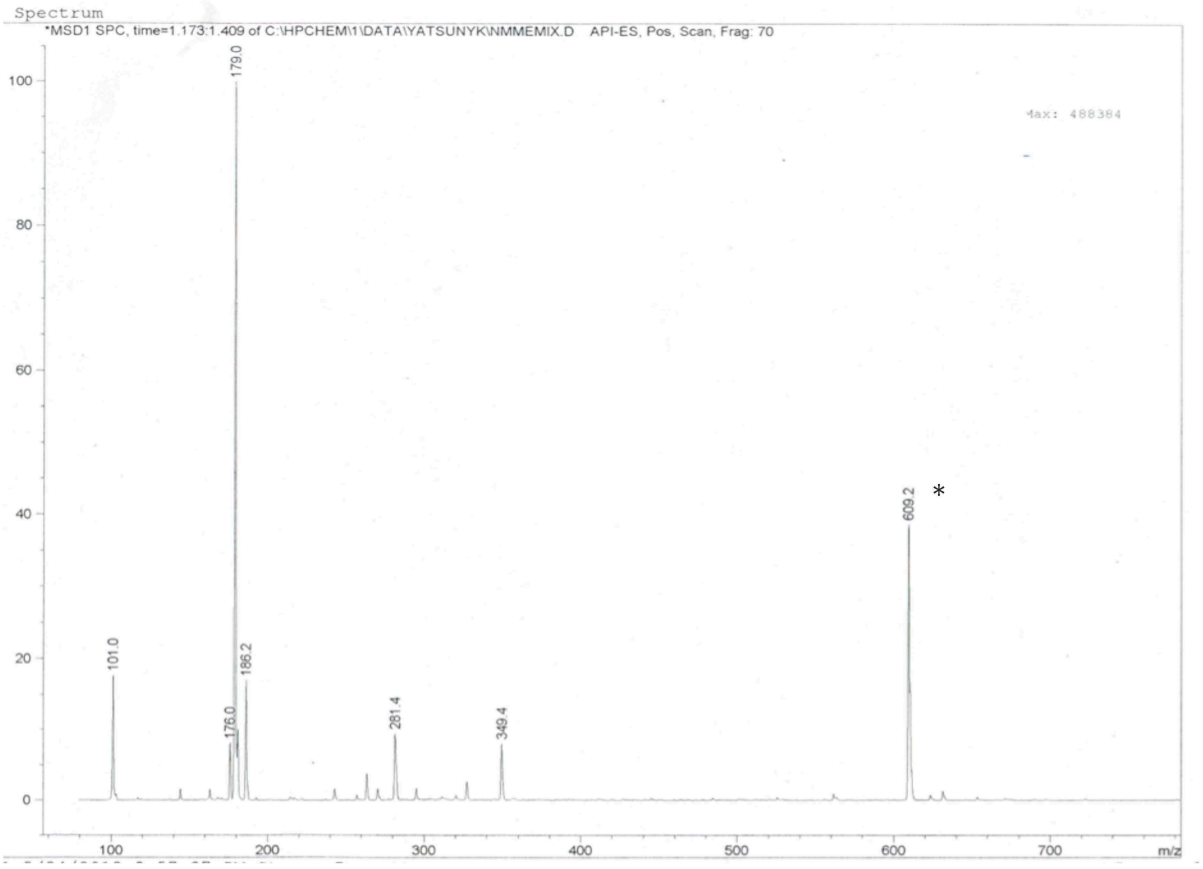


Figure 16. Mass spectrum of the commercial mixture of NMME isomers. Peak at m/z=609.2 corresponds to the exact mass of NMME. The mass spectra of the separated isomers is identical. This spectrum was collected at Haverford with the help of R. Scarrow and his research group members.

Results from proton NMR can be used to examine the separation of the isomers, as tilting of the N-methylated pyrrole out of the plane of the porphyrin ring alters the position of its substituents relative to the ring current and causes them to shift upfield in the NMR spectrum relative to the same substituents on the unmethylated rings; this leads to greater

separation of the peaks corresponding to these substituents on N-methylated rings (74). Furthermore, the *meso* protons (the protons located on the carbons that bridge the porphyrin pyrroles) are highly influenced by the ring current and, therefore, are sensitive to the site of N-methylation. Each of the four *meso* proton gives rise to a downfield singlet resonance (~10-11 ppm), leading to a total of sixteen singlets in the commercial mixture (four for each isomer). Not all of these singlets will be well-resolved due to overlap; as a result, in our experimental NMR spectrum of the commercial mixture of NMME, we observe only twelve peaks (Figure 17).

Given the range of ~10.8 to 10.4 ppm for *meso* proton resonances, it is safe to assume, based on symmetry, that the four *meso* resonances for a given isomer will be distributed throughout this range (i.e., all four peaks will not be clustered together). Therefore, we can simply examine the well-resolved resonances from 10.45 to 10.6 ppm, as these peaks integrate to yield eight out of the sixteen *meso* protons, and based on our assumption, these resonances should correspond to two *meso* protons from each isomer. A summary of the observed *meso* proton peak positions is given in Table 4. The ^{1}H-NMR spectrum for NMME isomer 1 is shown in Figure 18 and the spectra for isomers 2, 3, and 4 can be found in the Appendix.

Table 4. NMME *meso* proton resonances.

Sample	Observed *meso* proton resonances						
NMME commercial mixture	10.541	10.504	10.497	10.483	10.476	10.462	10.455
NMME Isomer 1	X	X	10.496	X	X	X	10.455
NMME Isomer 2	X	X	X	X	10.476	10.462	X
NMME Isomer 3	10.539	10.504	X	X	X	X	X
NMME Isomer 4	10.54	X	X	10.483	X	X	X

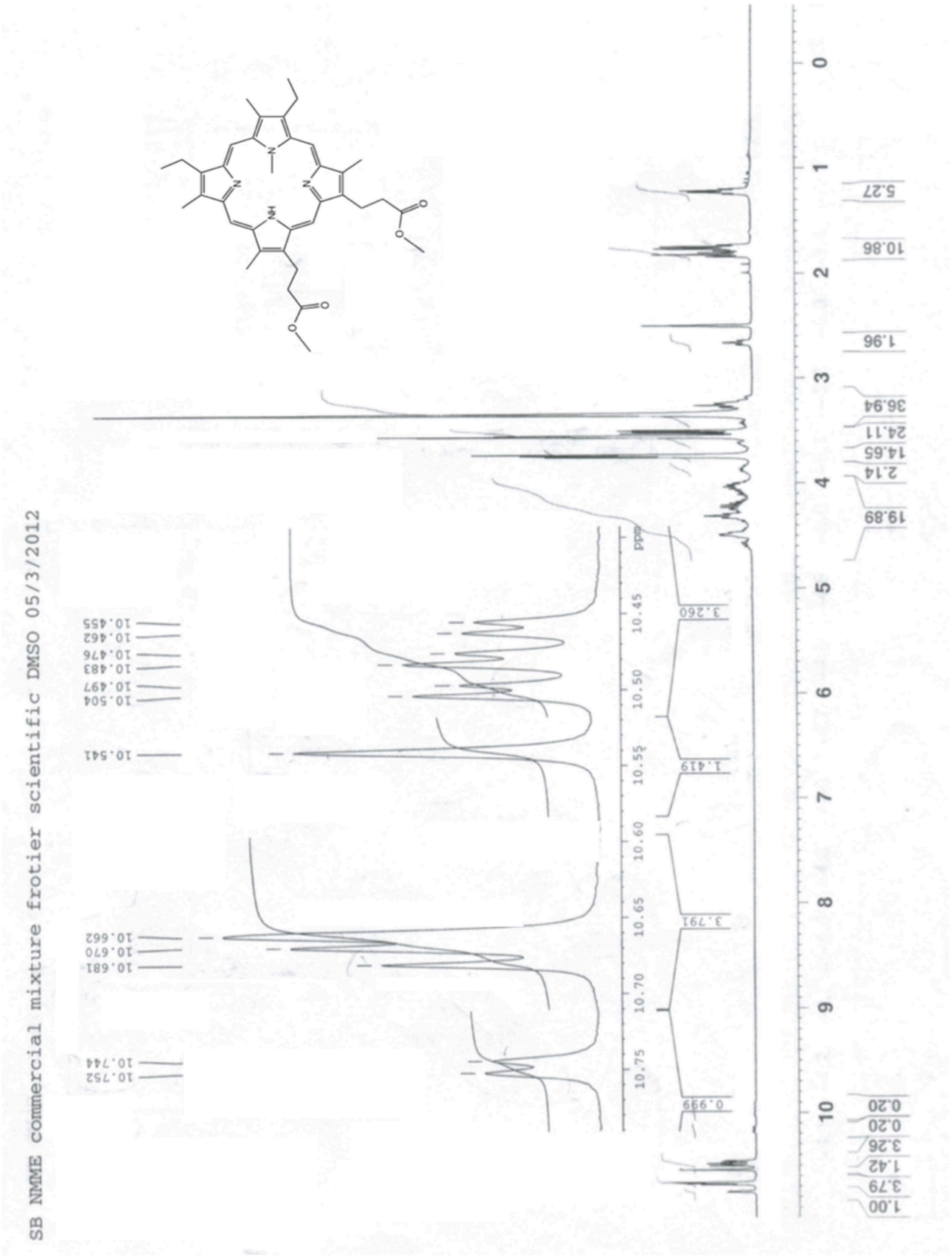


Figure 17. ^{1}H-NMR spectrum of the commercial mixture of NMME in d6-DMSO. The chemical shifts of the *meso* protons are expanded in the inset.

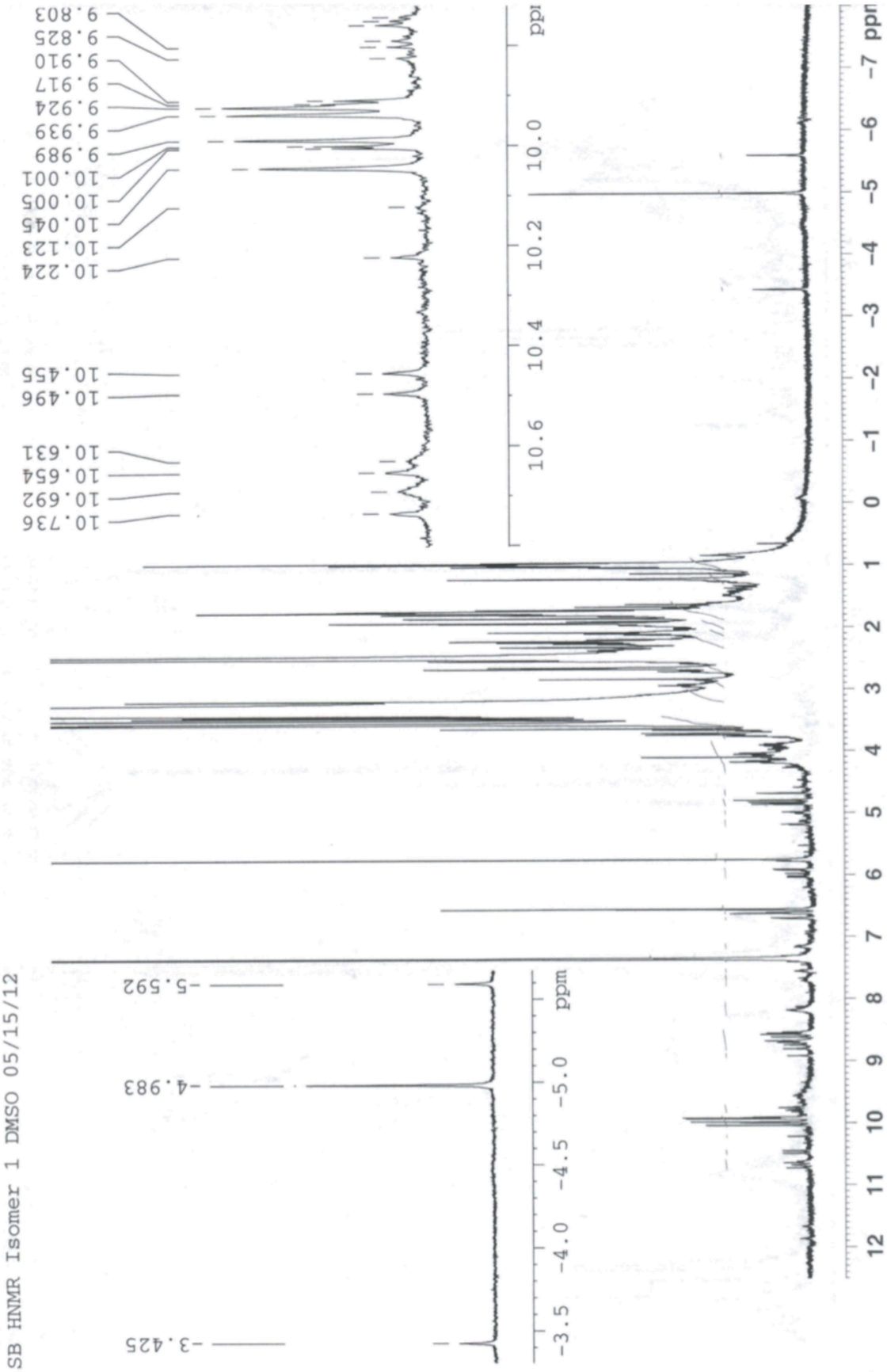


Figure 18. ^{1}H-NMR spectrum of NMME isomer 1 in d6-DMSO. The chemical shifts of the *meso* protons and core protons are expanded in the insets.

Significantly, in the narrow range we are analyzing (10.6 to 10.45 ppm), we only observe two proton resonances per isomer, as expected. And perhaps even more significantly, taken together, the observed two resonances from each isomer yield all of the resonances of the commercial mixture in the correct proportions. In other words, the sum of the NMR spectra of the four separate isomers in this region would produce the NMR spectrum of the commercial mixture.

Of course, we cannot simply limit our analysis to this small region of the NMR spectra. Notably, there are countless peaks present in the NMR spectra of the separated isomers that are not present in the commercial mixture. The most concerning peaks appear in the far upfield (negative) region of the NMR spectra. In this region, the commercial mixture of NMME exhibits two major resonances at -4.545 and -5.581 ppm, corresponding to the core N-H and N-CH_3 protons, respectively. Disturbingly, in the NMR spectra of the separated isomers, there is a large peak present at around -4.98 ppm that is not present in the commercial mixture spectrum, suggesting that there may be a porphyrin based contaminant in the separated isomers. Moreover, a set of chemical shifts ~10.0 ppm (in the range of porphyrin *meso* protons) in the separated isomers grows in proportion to the height of this peak, providing more evidence that this species is indeed porphyrin derived. Strangely, one would expect that such a species would have been discovered in our mass spectrometry experiments. However, it is possible that this species is uncharged due to a lack of protonatable side chains, and thus did not appear in the spectra.

We proceeded to investigate the GQ-stabilizing abilities of these separated isomers via FRET melting assays; although we bear in mind that there may be significant contamination of these samples with unknown (perhaps, porphyrin) species. The results of

this preliminary assay suggest that all NMME isomers can stabilize F21D, but not equally well—there is a slight preference for isomers 3 and 4, which putatively correspond to the isomers with their N-methyl group on an ester substituted pyrrole (Figure 19).

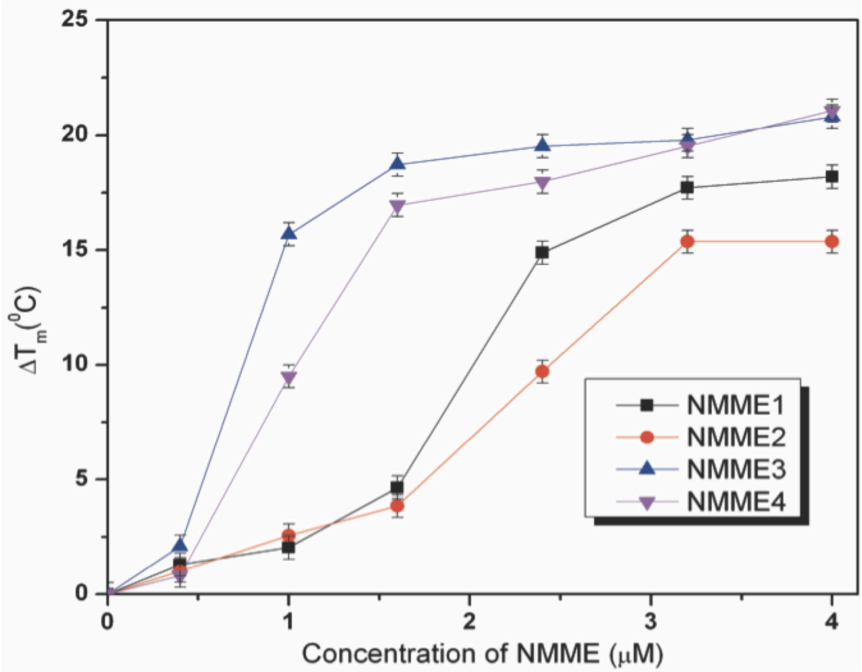


Figure 19. Increase in the stabilization temperature of F21D with increasing concentration of NMME isomers as measured by FRET melting assay.

Thus, in the crystal structure, it is less likely that Tel22 forms a complex with just one isomer of NMM (though it is still possible that only one of these isomers crystallized). Crystallization with separated NMM isomers might more clearly reveal the interactions of the propionic acid side chains with the GQ backbone.

Conclusions and Future Work

Through our biochemical, biophysical, and spectroscopic studies with NMM and Tel22, we have discovered the following: 1) that NMM is able to isomerize Tel22 from a mixed hybrid to parallel topology in dilute K^+ conditions and 2) that NMM cannot bind to the antiparallel form of Tel22 in Na^+, making it one of the few ligands with selectivity for a specific quadruplex fold (70). The X-ray crystal structure of NMM in complex with Tel22 reveals that the origin of both NMM's selectivity for GQ DNA and for specific GQ topologies is its methyl group, which distorts the shape of the porphryrin tetrapyrrole core from planarity and optimizes its interactions with the 3' G-tetrad (72). Contrary to electrostatic predictions, NMM's negatively charged propionic acid side chains seem to have a negligible effect on its ability to stabilize Tel22 and a small, but noticeable impact on the binding affinity. However, the nature of these interactions (if any) is not fully uncovered (72).

To address this query and others, we have attempted to separate and characterize the isomers of N-methyl mesoporphyrin IX dimethyl ester (NMME). The four regioisomers of NMME were separated via HPLC with a polar-amino cyano column. And although the HPLC chromatogram depicts an adequate separation, proton NMR data reveals the presence of several contaminants possibly derived from porphyrins. Perplexingly, these contaminants are not present in the mass spectra of the isomers, possibly because they are uncharged. FRET melts of the separated isomers reveal that all of the isomers are able to bind to and stabilize GQs with a slight preference for isomers 3 and 4 (those that are methylated on pyrroles with propionic acid substituents).

Although significant progress has been made in our study of NMM, there are many questions that must still be addressed. First and foremost, the binding constant of NMM to parallel Tel22 has yet to be ascertained. The main obstacle for the determination of this value is experimental design. Tel22 does not adopt a parallel conformation under commonly used conditions and the two conditions that drive the formation of this topology, high concentrations of NMM and polyethylene glycol (PEG) (75), preclude our study. Beginning our titration with a high NMM:Tel22 ratio would make obtaining a complete binding curve impossible, as we would never be able to obtain NMM:Tel22 ratios lower than our initial conditions. Utilizing PEG to induce a parallel GQ conformation would highly complicate the experiment, as there is data that suggests that PEG can bind to GQs, and furthermore, we have collected preliminary UV-vis data that suggests that PEG can bind to NMM. In order to extract the binding constant from such experiments, it would first be necessary to deconvolute the Tel22-NMM binding data from the effects due to PEG.

Another interesting facet of the NMM-Tel22 interaction is the potential stereochemical selectivity. Although the difference in stabilization between the four regioisomers does not seem to vary greatly (based on our preliminary FRET data), it is possible that the interaction is enantio-specific. However, this idea can be easily discarded if we examine any UV-vis titration of NMM with Tel22. If the interaction were enantio-specific, there would be a significant population of unbound NMM even at high Tel22:NMM ratios, which is not what we observe.

In order to increase the binding constant of NMM for Tel22, attempts are underway to synthesize NMM derivatives with amide side chains. We are curious to know if these side chains would affect NMM's selectivity for GQs over duplex DNA, as an electrostatic

component would be introduced to the interaction in some cases, depending on the specific derivative that is created. It might also be interesting to metallate NMM as an alternative method to probe the electrostatic component of the interaction. Although metallation with M^{2+} (such as Zn^{2+}) would not result in a species with overall positive charge, the molecule would be zwitterionic in actuality and positive charge density would be clustered at the metal site, still allowing for an electrostatic contribution.

Duplex-derived Interstrand Quadruplex Forming Potential (ddiQFP)

There is abundant *in vitro* evidence that GQ structures can form from single-stranded DNA or RNA oligonucleotides that contain the proper sequence motifs (76). Evidence for the stable formation of unimolecular GQs in the presence of their complementary Watson-Crick duplex sequence is not as ubiquitous, and in most cases, quadruplex formation will not occur spontaneously; rather, specific conditions are necessary to drive the formation of the GQ from the competing duplex (38,77,78).

Another potentially biologically significant quadruplex structure could be formed between two complementary strands of DNA (intermolecular GQ). Sequences with the speculated propensity to form these structures are said to have duplex-derived quadruplex forming potential (ddiQFP) (40). If such structures exist *in vivo*, they could have widespread impacts on cellular function; specifically, sequences with ddiQFP are highly associated with upstream promoter regions and genes involved in mitochondrial functions (40). The mere demonstration that these sequences form GQs *in vitro* would likely largely affect bioinformatics, as these sequences would need to be considered in search algorithms for potential GQs *in vivo* (40).

Here, we spectroscopically characterize a sequence with ddiQFP and attempt to induce the structural conversion of this sequence from duplex to GQ.

ddiQFP Sequence Design

The sequences used in this study were engineered by Professor Brad Johnson at the University of Pennsylvania. The engineered sequence (hereafter simply referred to as ddiQFP) is a G-C rich duplex with two T_4 hairpin loops (Figure 20).

Figure 20. ddiQFP construct.

The structure of the expected GQ would contain three G-tetrads, two seven nucleotide loops (sequence T_4C_3), and two three nucleotide loops (sequence C_3).

The rationale for this construct essentially can be decomposed into three arguments. First, the structure is in the compact form (i.e., the cytosines will function as the loops of the GQ), as the addition of spacer nucleotides would only increase the number of hydrogen bonds lost from the duplex structure upon GQ formation. Therefore, GQ formation from this ddiQFP construct should be more thermodynamically favorable than any other ddiQFP design. Secondly, the complementary strands are tethered by the T_4 loops in order to mimic the constraints the sequence might experience in the context of genomic DNA. Note, these constraints are not entirely geometric and should not restrict the formation of GQ folds. On the contrary, the tethering should merely restrict how far the two duplex strands can separate; this makes "bimolecular" GQ formation more kinetically favorable, as the "two" strands will be kept in close proximity. Finally, this alternating G tract-C tract sequence likely bears some biological significance, as over a third of mitochondrial QFP motifs are of this form (40); these sequences appear related to the mitochondrial "GC clusters," which might be involved in transcriptional regulation (79).

ddiQFP Spectroscopic Results

The main methods used to investigate the structural form of ddiQFP were circular dichroism (CD) spectroscopy and UV-vis thermal difference spectroscopy (TDS), as both

methods can produce spectra that are indicative of either duplex or quadruplex formation (65,80). It is important to note that this project is at its initial stages, and the approach we took was to examine a wide range of conditions and DNA sequences that could potentially lead to GQ formation.

To begin our studies, we investigated the spectroscopic characteristics of two duplex control sequences: 5'-$A_3(G_3C_3)_2A_3$-3' annealed to its complement (referred to as dscontrol) and two strands of 5'-$T_2(C_3G_3)_2T_2$-3' annealed together (referred to as T-tailed dscontrol). Essentially, these duplex controls should have a very similar structure to the ddiQFP (besides the T_4 loops), and thus they should give very similar spectroscopic signals. Note that T-tailed dscontrol is not fully complementary, as each strand will have two flanking Ts on each end. This design was used to mimic the single stranded T_4 loops in our ddiQFP construct to gain an understanding of how these unpaired nucleotides might affect the observed spectra.

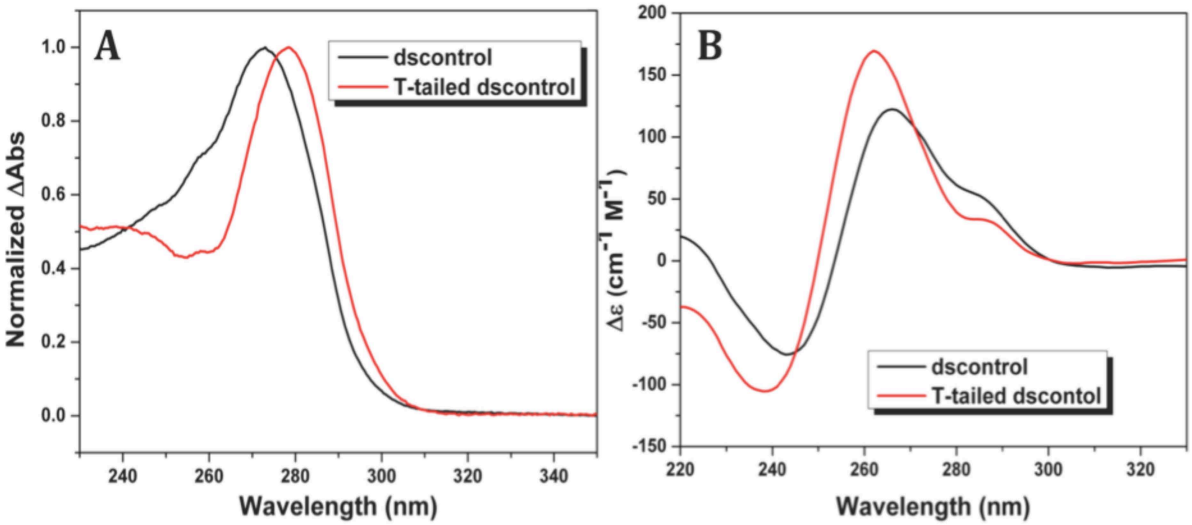


Figure 21. (A) TDS and **(B)** CD spectra of duplex control sequences, dscontrol and T-tailed dscontrol.

The TDS obtained for both of the control sequences are both characteristic of a GC rich duplex, with a peak at around 275 nm (Figure 21A). Since dscontrol has both AT and

GC base pairs, there is a slight shoulder at ~260 nm and a slight shift of the major peak toward this value, which are indicative of partial AT duplex formation (65). The CD spectra of these control sequences also reveal duplex formation (Figure 21B). Both spectra exhibit a peak at ~265 nm and a shoulder ~285 nm, features both indicative of GC-rich duplexes (65). Note that the CD spectrum of a parallel GQ would also give rise to a peak at 264 nm (without the presence of a shoulder). Thus, it is difficult to discern duplex or quadruplex formation by CD spectroscopy alone, and other methods must be used to complement the data obtained. These data indicate the that both controls adopt duplex structure as expected.

In an attempt to drive the structural equilibrium of the ddiQFP sequence from duplex to quadruplex, we subjected the construct to various treatments (listed in Table 5). The annealing temperature of 105 °C reflects the predictably high melting temperature of starting duplex structures in high salt conditions (~95 °C in 100 mM K^+ (81)). The ligands used were ZnT4, NMM, and PEG. NMM was used because of its extreme GQ selectivity (59,60,70); ZnT4 was used because it has been shown to induce bimolecular quadruplex formation (48); and PEG was used because it functions as a molecular crowding or dehydrating agent, which has been shown to support intramolecular GQ formation in the context of duplex DNA (77). We hoped that melting and slowly cooling the sequence in the presence of agents that favored GQ formation would shift the equilibrium from duplex the GQ. However, as of now, the results of these studies do not provide clear evidence of GQ formation.

Table 5. Conditions for ddiQFP annealing

Annealing Temp. (°C)	Buffers	Ligands	Cations
90	LiCac	NMM	Li^+
105	(pH=7.2, 5.8)	ZnT4	K^+
	Tris	PEG	Na^+
	(pH=8.0)	NMM+PEG	Mg^{2+}

To begin, the TDS of the ddiQFP sequence in LiCac buffer (pH=7.2) exhibits a negative peak at 258 nm and a positive peak at 283 nm with a shoulder at 273 nm (Figure 22A). These features are uniformly observed under all conditions tested here. Given that hypochromicity is the result of stacked and specifically oriented DNA bases, the negative change in absorbance is quite strange since absorbance normally increases upon melting or denaturing a secondary structure of DNA as a result of less stacking (82). TDS relies on the idea that the DNA is fully denatured and single-stranded at high temperatures and has a well-formed secondary structure at low temperatures. In actuality, the former cannot be completely achieved by our construct, since the "two strands" are effectively bound together by the T_4 loops. This likely impacts the shape of the observed curves and the way we can interpret the TDS with respect to previously obtained spectral fingerprints (65).

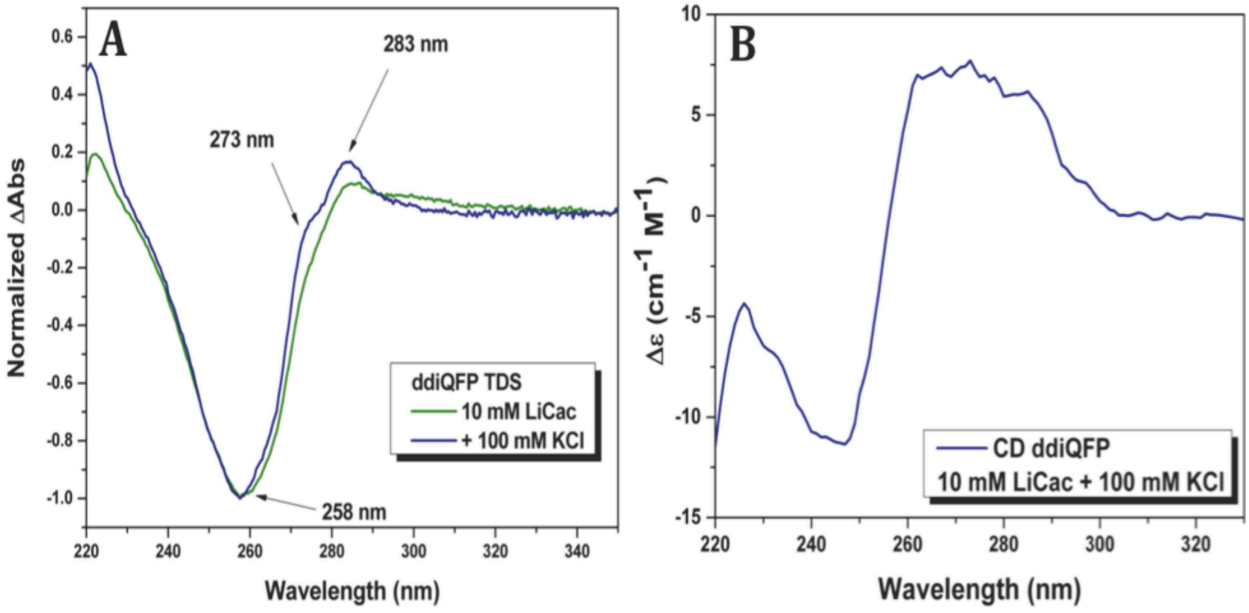


Figure 22. (A) TDS and **(B)** CD spectrum of ddiQFP in 10 mM LiCac buffer (pH=7.2). Data depicted are representative of the spectra collected under the conditions tested. The complete set of spectra can be found in the appendix.

The CD spectra are likewise uninformative in this instance. In each case, the signal intensity is too low to be reliably interpreted, suggesting that the sequence is largely unstructured in this buffer, although the reason remains unclear (Figure 22B). Some consistently observed features of the spectra include a very broad peak spanning from 260-280 nm and a trough at 245 nm; the spectra vaguely resemble those for the duplex controls, but the signal is too low to make any reliable conclusions. The CD spectra of ddiQFP annealed at 105 °C in LiCac (pH=5.8) were also collected, but similarly yield results that are indeterminate.

Given that the results from our very first experiments in cacodylate-based buffer were quite uncertain, we tried using Tris as an alternative, as the buffer condition might favor a certain structural form. The TDS and CD spectra of the ddiQFP sequence in Tris buffer (pH=8.0) yielded much clearer results indicative of duplex formation: the TDS exhibit a major positive signal at ~280 nm (Figure 23A), and the CD spectra demonstrated a peak around 265 nm and a shoulder at 285 nm (Figure 23B). Note that the unusual trough at ~260

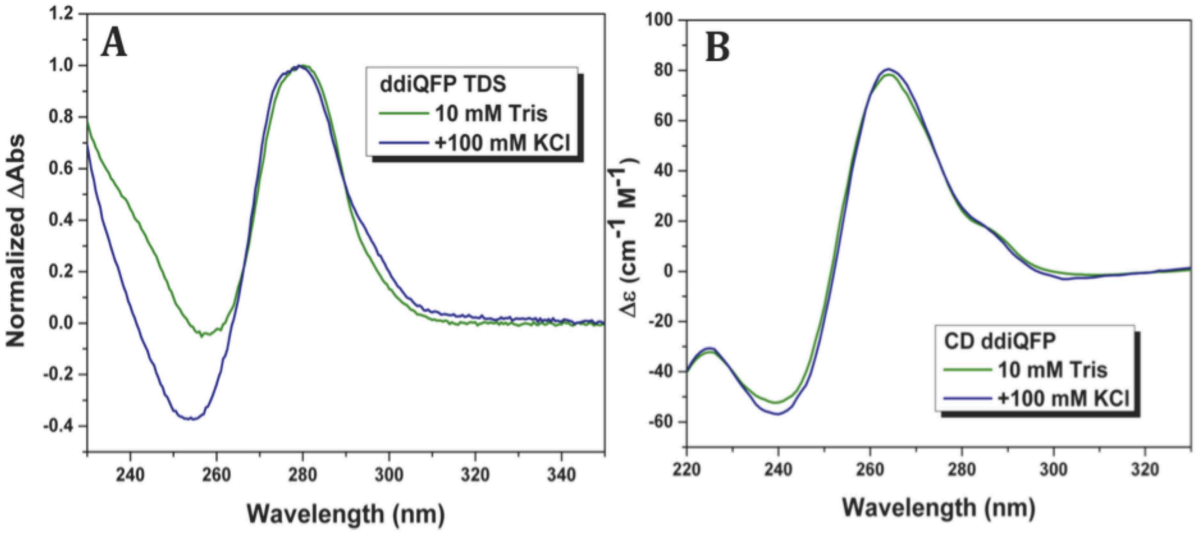


Figure 23. (A) TDS and **(B)** CD spectra of ddiQFP in Tris (pH=8.0) in the presence and absence of 100 mM KCl. The results are indicative of duplex formation. Data are representative of all the conditions tested.

nm is still observed in the TDS (though less prominently), and thus, is presumed to be a spectroscopic feature of the loops. We are currently still experimenting with this sequence in Tris buffer and are hopeful for quadruplex formation under other, untested conditions.

The discrepancy between the data obtained in cacodylate buffer and the data obtained in Tris is perplexing, as neither condition should particularly favor duplex formation. One distinct aspect of Tris is its large pH dependence on temperature (approximately -0.03 units/°C) (83). At room temperature (~22 °C), the buffer's pH is 8.0, but in TDS experiments the temperature increases to 95 °C, where the buffer would be at a drastically lower pH of 5.8. This drastic change in pH during the annealing process may have had an impact on the duplex's ability to form, although the mechanism remains unclear.

Secondly, the enzymatic ligation of the ddiQFP construct used in these experiments (using 10 mM Tris) was performed at only 3 mM Mg^{2+} concentration, whereas in the previous experiments (using 10 mM cacodylate) 10 mM Mg^{2+} was used for ligation. This decrease in magnesium concentration led to significantly higher yields for the ligation, perhaps, because the magnesium stabilizes an alternative structure of the construct. Therefore, under 10 mM Mg^{2+} some population of these non-duplex structures could have influenced the CD and TDS. Note, construct preparation was again by B. Johnson.

ddiQFP Sequence Variants

In addition to the previously discussed conditions, we also attempted to induce GQ structure with ddiQFP sequence variants (named ddiQFP2-5, Table 5). All sequences (besides ddiQFP2) have GGGG (vs. GGG) tracts and AT base pairs located just before or after the C-tracts in the stretch of DNA that would become a quadruplex loop. The logic for

these designs is based on two thermodynamic arguments: 1. Increasing the number of Gs in each tract would increase the stability of the GQ, as this increases the number of G-tetrads formed. 2. Although at first glance the addition of AT basepairs would seem to stabilize the duplex form of the ddiQFP structure (since the total number of hydrogen bonds increases because nucleotides were added), this argument neglects the nearest-neighbor effect on nucleic acid thermal stability; that is, the stability of a certain base pair of DNA depends on the identities of its neighboring base pairs (84). According to this model neighboring G-C base pairs are very stable and interrupting these GC base pairs actually destabilizes the duplex (84) (See T_m values in Table 6). Finally note that ddiQFP4 has a T-T base mismatch in the center of the sequence; this should have a destabilizing effect on the duplex with minimal impact on GQ-formation.

Table 6. ddiQFP sequence variants.

Sequence Name	Sequence Structure	Estimated T_m (°C)*
ddiQFP	T T CCC**GGG**CCC**GGG** T T GGG**CCC**GGG**CCC** T T T	96.3
ddiQFP2	T T **GGG**CCC**GGG**CCC T T CCC**GGG**CCC**GGG** T T T	96.3
ddiQFP3	T T **GGGG**ACCCCTA**GGGG**TCCCC T T CCCCT**GGGG**ATCCCCA**GGGG** T T T	91.6
ddiQFP4	T T **GGGG**ACCCCT**GGGG**TCCCC T T CCCCT**GGGG**TCCCCA**GGGG** T T T	89.5
ddiQFP5	T T **GGGG**AACCCCTA**GGGG**TTCCCC T T CCCCTT**GGGG**ATCCCCAA**GGGG** T T T	89.4

*Est. with IDT's OligoAnalyzer using the melting temperature of the relevant duplex (81).

Although these thermodynamic arguments are sound, the effects of these structural alterations are not so simple—we must consider how each change affects the stability of the quadruplex structure *relative to* the duplex. Thus, we cannot consider each change with respect to each structure in isolation. For instance, it is immediately evident that increasing the number of Gs in each G-tract would not only increase GQ stability but also greatly increase the stability of the duplex by increasing the number of GC base pairs. It is also true that adding AT base pairs to destabilize the duplex will likewise have a destabilizing effect on the GQ structure, as increasing loop length negatively impacts the stability of GQs (85). Additionally, the binding affinity of intramolecular GQs towards their complementary strands increases with GQ loop length, and likewise, the kinetics of GQ unfolding and duplex formation increases upon increasing loop lengths (86). In designing these sequences, we hoped that any antagonistic effects on the stability of the GQ would be outweighed by factors favoring its formation.

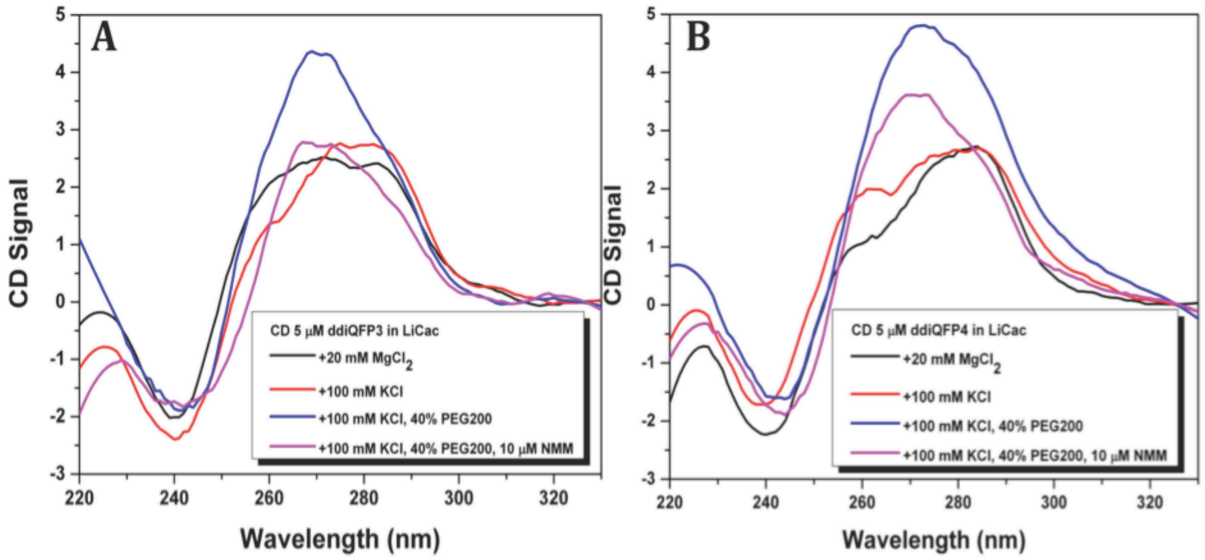


Figure 24. CD spectra of 5 μM **(A)** ddiQFP3 and **(B)** ddiQFP4 in LiCac (pH=7.2) with various additional components (described in the legend). In all cases, the signal is low, but in the presence of PEG200 there is a notable increase, indicating the presence of some secondary structure. Spectra for ddiQFP2 and ddiQFP5 can be found in the Appendix.

The CD spectra were collected in cacodylate buffer (pH=7.2) and yield spectra that have generally low signal intensity. ddiQFP3 and ddiQFP4 show a slight induction in signal around 270 nm upon the addition of PEG that may indicate the presence of some secondary structure. As we had better success in Tris, these studies should be repeated in Tris buffer to ensure the effects we see are not due to low signal to noise.

Conclusions and Future Work

The goal of this work is to provide experimental support to the theory that quadruplexes can form between two strands of complimentary DNA discussed in depth with additional bioinformatics evidence by Cao *et al.* (40). The results depicted here outline the effects of a variety of different conditions on ddiQFP secondary structure formation. In all instances in which cacoldylate based buffer was utilized, the TDS produced an uncharacteristic large negative peak at ~260 nm and the observed CD spectra demonstrated a low signal that could not reliably be interpreted but that suggests the absence of significant secondary structure. In Tris buffer, the ddiQFP sequence reliably exhibited duplex character in both the TDS and CD spectra. However, this result could simply be attributed to a different ligation procedure used for these samples and not necessarily to a difference in buffer. GQ formation was not clearly observed in any samples tested, which is not completely surprising given the extremely high stability of the duplex.

Although it is important to contemplate the results presented here, it is perhaps more important to consider potential future experiments. As we have done here, we can tackle this problem from two perspectives: we can either alter the ddiQFP construct or we can alter the conditions to which we subject these sequences.

In addition to all that we have tried, another potential way in which the ddiQFP construct could be altered to favor GQ formation is by reducing the number of thymines in the loops (making a TT-loop structure, see Figure 25). Shortening the loops causes strain on the duplex and, thus, lowers its melting temperature. Moreover, two of the potential GQ loops would be shortened (from 7 nt to 5 nt in the case of the TT-loop structure) which should have a stabilizing effect on the quadruplex. A major hurdle in this scheme is the

generation of such constructs, which requires the formation of a structurally unstable intermediate prior to ligation which could lead to homodimer formation in place of the TT loop structure. One way to counteract homodimer formation is by altering the sequence such that more nucleotides base pair in the intermediate loop structure, leaving a shorter overhang for ligation. The most drastic alteration would result in a blunt-end intermediate (Figure 25), which would have a very inefficient ligation, as there would be no way to orient the intermediates correctly.

Figure 25. **(A)** ddiQFP hairpin loop intermediate with T_2 loop (TT-loop structure). **(B)** Undesired homodimer formation. **(C)** "Blunt-end" intermediate.

Another future direction for this project is to test additional conditions which could drive the structural equilibrium toward quadruplex formation. The use of quadruplex selective ligands will likely be an important feature in future experiments. NMM, although a very GQ selective ligand, may not be an adequate choice for these experiments. First, NMM has a strong preference for parallel-stranded GQs and is insensitive to the antiparallel GQ structures. If ddiQFP sequences cannot take on a parallel GQ conformation, this ligand will not perform well in this context. Second, even at very high concentrations of NMM, the stabilization of GQs is only ~20 °C (at NMM:GQ~20:1) (70). Given the extremely high melting points of the competing duplex sequences, NMM's effect on the GQ stability is likely not enough to have a significant impact unless the GQ formed is already quite stable (T_m>75 °C). For these reasons, it would be best to use a ligand, such as telomestatin, with broad GQ selectivity and the ability to highly stabilize these structures (51). In addition,

using low K^+ concentrations (e.g., 5 mM; so far we have tried only 100 mM) may actually assist GQ-formation; low salt concentrations will ensure that the duplex is completely melted at 95 °C, thereby allowing the ddiQFP to experience alternative folds at these temperatures.

Another external factor that might influence the ability of the ddiQFP constructs to take on GQ folds is the cooling rate. For normal single-stranded intermolecular sequences that form GQ structures *in vitro* our cooling method involves removing the heat source while keeping the sample well-insulated. This method usually works well because these GQ structures do not start forming until relatively low temperatures are reached and there are not many other competing products. Newton's law of cooling states that, when cooling an object, the temperature will decrease rapidly at first and slow as the object approaches the ambient temeprature. Therefore, at the beginning of the cooling process, when the competition between duplex and quadruplex formation will be greatest (i.e., the first 10-15 degrees), any structures that form are likely the products under kinetic control. In our case, we suspect the duplex to be the product of kinetic control, as GQ formation would probably have a relatively high-energy transition state as the backbone of the DNA contorts to create this structure. Thus, it would be advisable to have a slow, controlled cooling rate at the very start of our cooling procedure to avoid duplex formation. This control cooling could be easily achieved using either our UV-vis or CD instrument.

Quadruplex Formation in Mitochondrial DNA (mtDNA) Sequences

Mitochondrial DNA (mtDNA) is a non-nuclear form of DNA that is exclusively maternally inherited in humans. Mutations in mtDNA, especially deletions, can cause a wide range of disorders (87,88). In the past, attempts were made to determine the relationship between the site of the deletion and the DNA sequences surrounding these sites. The first identified features were short direct repeats that overlapped the deletion junction (89-91). An alternative mechanism proposed by Damas *et al.* states that sequences with the ability to form cruciforms and stem-loop structures are the most highly associated with deletion (92). However, these structures have only been implicated in a small subset of deletions.

Since G-qadruplexes are associated with nuclear genomic instability (33), it is hypothesized that these structures may also be involved in mtDNA deletions. In a recent bioinformatics investigation, Brett Kaufman's lab at the University of Pennsylvania showed that, indeed, sequences with GQ-forming potential (QFP) were significantly associated with mtDNA breakpoints (unpublished results). Here, we demonstrate that three sequences overlapping with 3' mtDNA deletion breakpoints are able to form quadruplex structures *in vitro* under physiological conditions of 150 mM K^+, and under low salt conditions (10 mM K^+), these GQ structures are induced by Pb^{2+}, providing a potential mechanism for the widely recognized effects of lead-induced toxicity on mitochondrial form and function.

Spectroscopic analysis of quadruplex formation and stability in mtDNA sequences

The sequences used for analysis were designed by Brad Johnson at the University of Pennsylvania and are listed in Table 7. Control sequences (noted with 'c') were also

constructed in which guanines were replaced with adenines in order to impede quadruplex formation.

Table 7. mtDNA sequences with QFP and their corresponding control sequences along with melting temperatures extracted from the melting (T_m) and cooling (T_c) curves.

Sequence Name	Sequence	T_m (°C)	T_c (°C)
A	GGA TGG GGT GGG GAG G	80. ± 2	69.9 ± 0.5
Ac	**A**GA TGG **A**GT GG**A** GAG G	12 ± 1	11.6 ± 1.0
B	GGG GGA TGC GGG GG	85.2 ± 0.9	74 ± 1.3
Bc	**A**GG **A**GA TGC **A**GG **A**G	14 ± 2	14 ± 2
C'	GGA GGG TGG ATG G	54 ± 1	42 ± 0.2
C'c	**A**GA GGG T**A**G ATG G	26.9 ± 0.5	26.9 ± 0.5

Spectroscopic results indicate that sequences A, B, and C' are all able to form GQs in 10 mM lithium cacodylate (pH=7.2) with 150 mM KCl. TDS spectra indicate GQ formation with positive peaks 243 and 273 nm and a negative peak at 295 nm (Figure 26A, C, E), and CD spectra suggest parallel character in sequences A and B with a peak at 264 nm and mixed-hybrid character in sequence C' with peaks at both 264 and 295 nm (Figure 26B, D, F). Note that the corresponding control sequences do not demonstrate GQ formation as demonstrated by the TDS spectrum whose positive peak at 260 nm is simply indicative of AT base pair formation, likely the result of homodimers (Figure 26A, C, E). These homodimers apparently have significant secondary structure as evidenced by their strong CD signatures (Figure 26B, D, F).

The melting temperature of these structures was determined by monitoring the CD signal at 264 nm over a range of temperatures from 5 to 95 °C. The melting temperatures extracted for sequences A, B, and C' (Table 7) indicate very stable secondary structure formation ($T_m \geq 54$ °C), whereas the corresponding controls are very unstable ($T_m \leq 26.9$ °C).

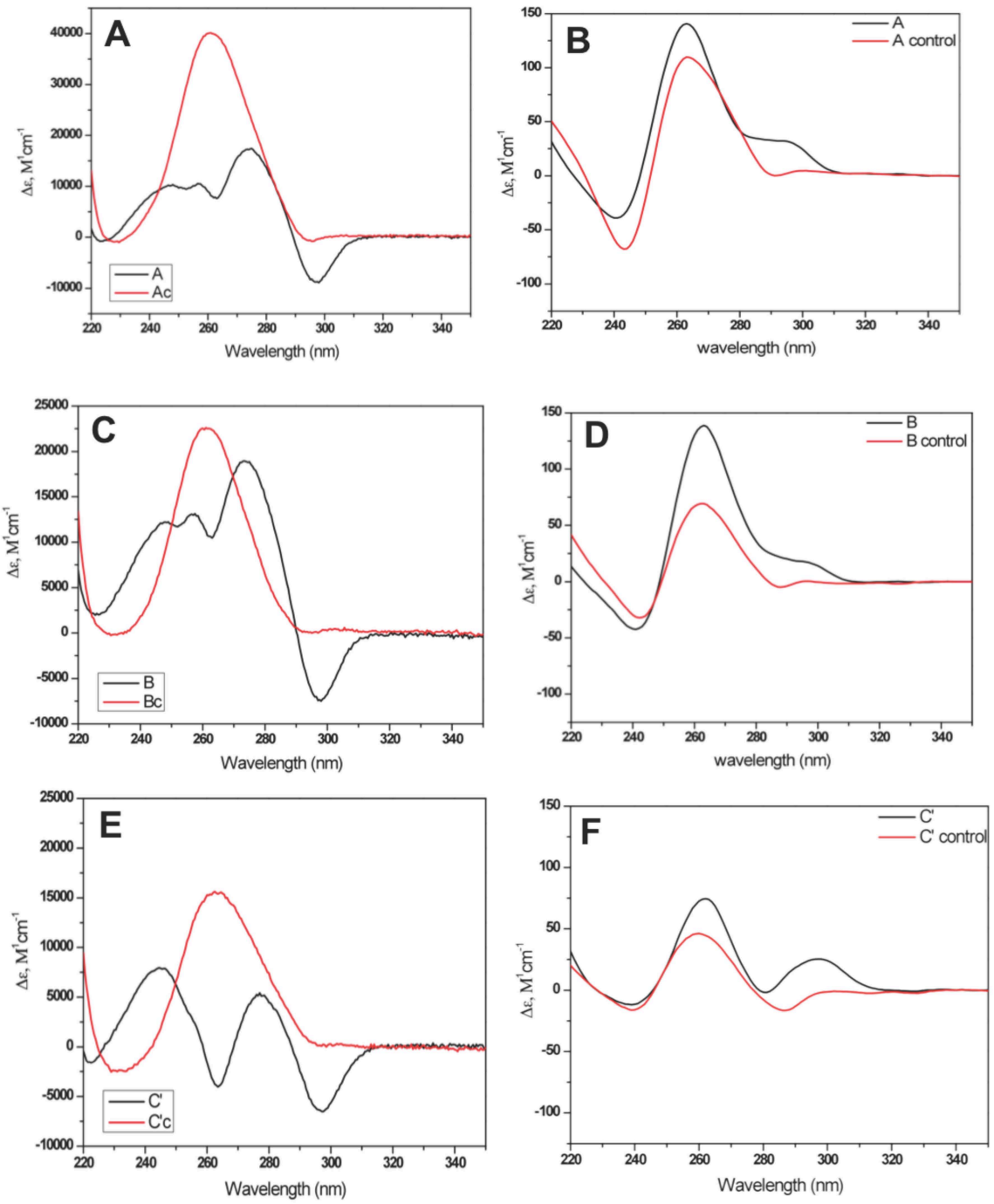


Figure 26. (A, C, E) TDS and **(B, D, F)** CD spectra of sequences A, B, and C' and their respective controls in 10 mM LiCac (pH=7.2) with 150 mM KCl. Taken together, the spectra indicate GQ formation in A, B, anc C' but not in the controls.

Effects of lead on mtDNA quadruplex formation

As a toxicant, lead (Pb^{2+}) can operate through a variety of different mechanisms. As a soft acid, lead binds tightly to the sulfhydryl groups of proteins and causes structural distortion in enzymes and structural peptides (93). Additional mechanisms of toxicity include the cleavage of the ribophosphate backbone of tRNAs (94) and competition with calcium for binding sites in proteins (93). One major target of lead is the mitochondria where it produces swelling and distortion of the cristae along with inhibited cellular respiration (95). Based on our previous evidence indicating the formation of GQ at these mtDNA breakpoints, we speculated that one mechanism of lead's effects on mitochondrial dysfunction might be through the induction or increased stability of quadruplex structures.

Using two of the sequences investigated previously, B and C', we examined the effects of lead on the CD spectra of the sequences. Because the GQs produced by these sequences were well-formed and quite stable at 150 mM K^+, we reduced the potassium concentration to 10 mM to make any effects due to lead more prominent. In the absence of lead, the CD spectrum of sequence B demonstrated the previously observed peak at 264 nm, representing a parallel-stranded GQ structure (Figure 27A). Upon the addition of lead to this sequence, there was a noticeable decrease in signal at 295 nm, representing the formation of a more parallel GQ topology. In contrast, sequence C' yielded a CD spectrum with very low signal intensity, indicating that this sequence is largely unstructured at 10 mM K^+ (Figure 27B). However, as lead is added, the DNA adopts a parallel GQ formation, as evidenced by the marked increase in signal at 264 nm. Thus, it appears that Pb^{2+} induces GQ formation in C' at low concentrations of K^+.

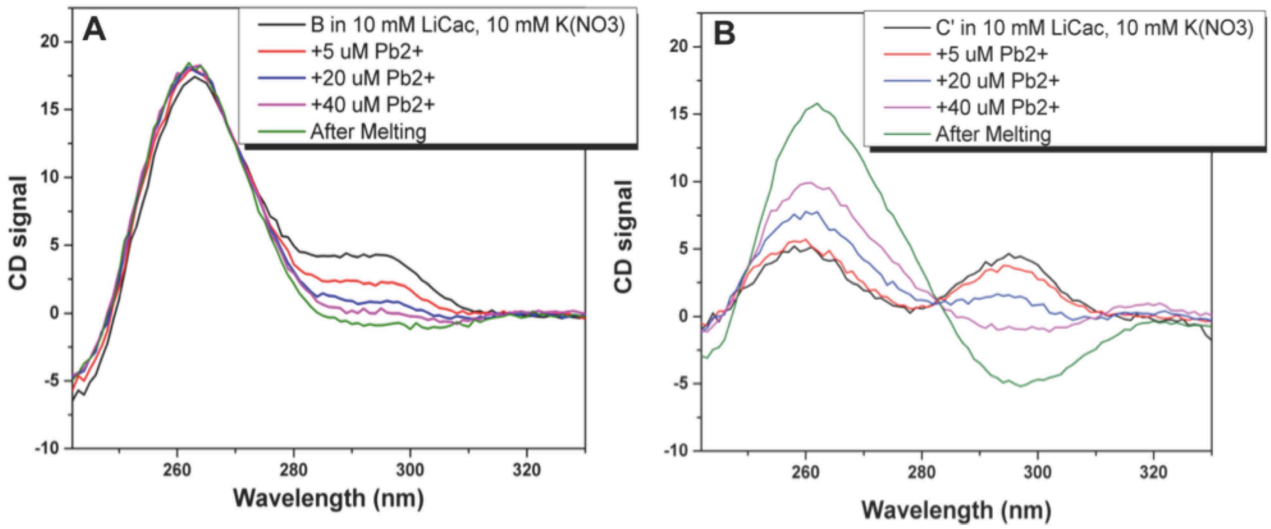


Figure 27. CD wavelength scans of 5 μM **(A)** sequence B and **(B)** sequence C' in 10 mM LiCac (pH=7.2) with 10 mM KNO_3 and increasing concentrations of $Pb(OAc)_2$ (indicated in the figure legends).

The thermal stability of these structures was assessed in 10 mM LiCac buffer with 10 mM KNO_3 in the presence and absence of 40 μM $Pb(OAc)_2$. Note that these samples were not annealed in the presence of lead; rather, Pb^{2+} was added to the samples after they were annealed with 10 mM K^+. Melting and cooling temperatures are indicated in Table 8. In the absence of lead, the structures appear to be generally unstable upon cooling, and the curves do not demonstrate the expected sigmoidal shape (Figure 28A). There is also significant hysteresis present in the melting of these samples (indicated by the values for T_m-T_c). Although the exact reason for this hysteresis is not completely clear, we speculate that it indicates the presence of multiple species, some of which may be kinetically favored and others that are thermodynamically favored. The melting curves for these samples likely depict the melting of the thermodynamically favored species, while upon cooling, the kinetically favored species is formed, leading to the significantly lower melting temperatures observed. Alternatively, the thermodynamic species may reform, but the barrier to its

formation may be very large such that it is not well formed until the end of the melting experiment. Possibly, decreasing the rate of cooling might yield curves with less hysteresis.

In contrast, in the presence of lead, the melting curves take on a much more well-defined shape, and the hysteresis observed in these samples decreases dramatically, suggesting that lead may act as a catalyst, speeding up the formation of a thermodynamically favored product (Figure 29). The stabilities of these species are only slightly lower than the thermodynamically favored products in the absence of lead.

Table 8. Melting temperatures obtained from the melting curves (T_m) and cooling curves (T_c) for mtDNA sequences B and C'. Values of T_m-T_c are also calculated and are indicative of hysteresis.

	10 mM LiCac, 10 mM KNO_3*			10 mM LiCac, 10 mM KNO_3, 40 μM $Pb(OAc)_2$**		
Sequence	T_m	T_c	T_m-T_c	T_m	T_c	T_m-T_c
B	75.7 ± 0.5	27.3 ± 0.8	48.4	67.8 ± 0.5	53.9 ± 0.2	11.9
C'	46.7 ± 0.2	15 ± 5	31.7	41.9 ± 0.4	28.9 ± 0.5	13.0

*T_m and T_c determined by examination or fitting of first derivative curves, except for T_c for C', which was obtained by inspection.
** T_m and T_c determined by fitting the data to a two-state Van't Hoff equation with constant heat capacity.

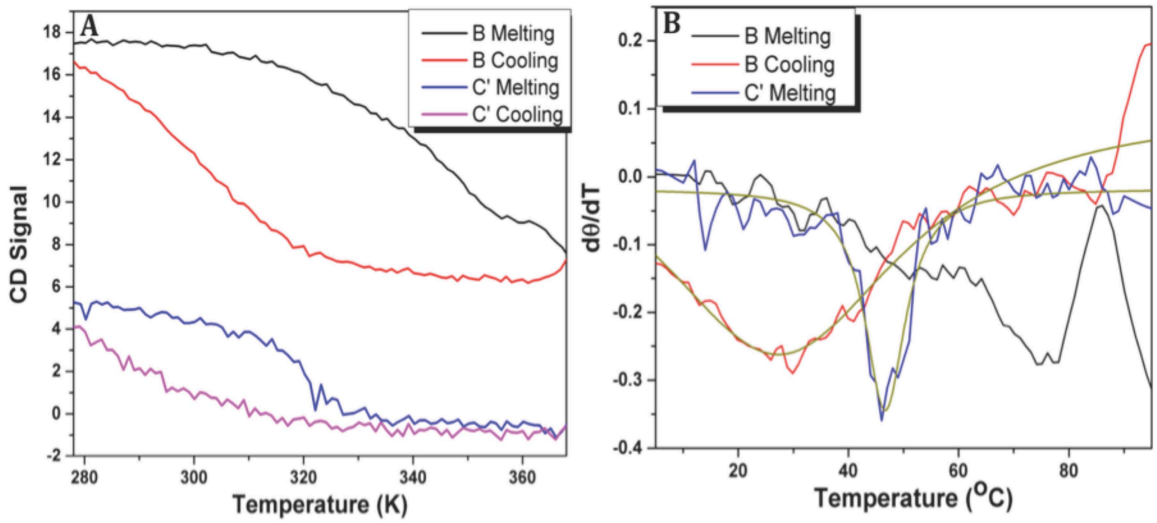


Figure 28. (A) Melting and cooling curves and **(B)** their corresponding derivatives for the melting of sequences B and C' in 10 mM LiCac (pH=7.2) with 10 mM KNO_3. The derivatives were fit with a Lorentzian function (gold curves) to accurately determine the location of the peak, which indicates the melting point.

It is interesting to note that upon completing the melting and cooling process with lead, sequence C' demonstrates a dramatic an increase in CD signal at 264 nm (Figure 29 and Figure 27B, 'After Melting'). This could either be caused by the formation of a distinct species (one that can only be formed through the annealing process) or could simply be caused by the formation of a larger amount of the previously observed species. The latter explanation is in good agreement with the presence of the isoelliptic point observed at 280 nm (Figure 27B), as this strongly suggests the presence of an equilibrium between only two species.

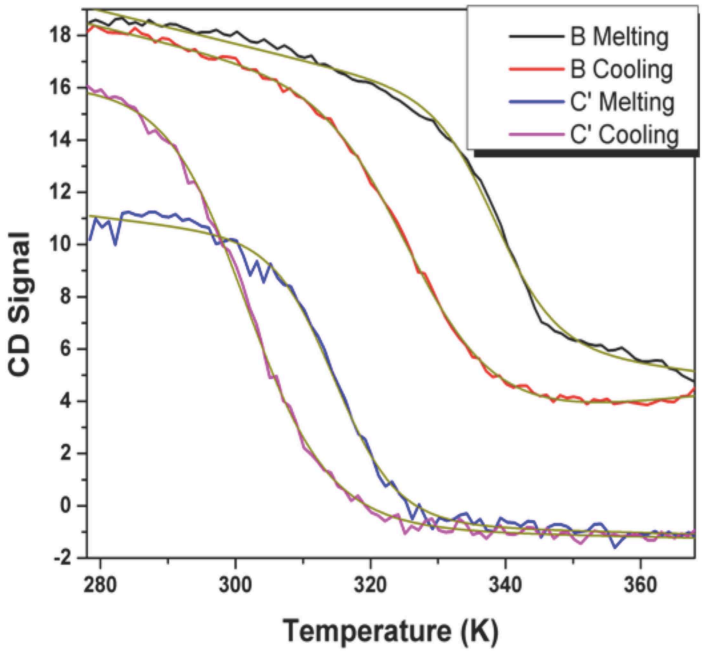


Figure 29. Melting and cooling curves and for the melting of sequences B and C' in 10 mM LiCac (pH=7.2) with 10 mM KNO_3 and 40 μM $Pb(OAc)_2$. The curves were fit to a two-state Van't Hoff function (gold curves) with constant Cp to extract the melting temperatures.

Conclusions and Future Work

Here, we demonstrate quadruplex formation by sequences found at mtDNA breakpoints, implicating GQs in the mechanism of the formation of these deletions under physiological conditions. At low concentrations of K^+, Pb^{2+} appears to induce GQ formation in one of these sequences (C') and, based on the decreased hysteresis between the melting and cooling curves, lead likely catalyzes the formation of a small population of thermodynamically stable species. These results bear strong medicinal implications, as they provide a direct link from lead toxicity to mitochondrial deletions and, thus, to mitochondrial dysfunction.

It is important to note that, in actuality, we have not explicitly shown the induction of GQ structure in C' or even the formation of GQs under these conditions. Thus, as an immediate project for future work, we must to conduct TDS on these samples to ascertain the validity of our assertions. It would also be very interesting to directly test for the effects of lead on the frequency of mtDNA deletions *in vivo*, and furthermore, whether a statistically significant ratio of these deletions occur at sites with GQ forming potential.

References

1. Bacolla, A. and Wells, R.D. (2004) Non-B DNA conformations, genomic rearrangements, and human disease. *J Biol Chem*, **279**, 47411-47414.
2. Bang, I. (1910) Untershungen uber die Guanylsaure. *Biochem. Z.*, 293-231.
3. Gellert, M., Lipsett, M. and Davies, D. (1962) Helix Formation by Guanylic Acid. *Proceedings of the National Academy of Sciences of the United States of America*, **48**, 2013-&.
4. Risitano, A. and Fox, K.R. (2004) Influence of loop size on the stability of intramolecular DNA quadruplexes. *Nucleic Acids Res*, **32**, 2598-2606.
5. Davis, J.T. (2004) G-quartets 40 years later: from 5'-GMP to molecular biology and supramolecular chemistry. *Angew Chem Int Ed Engl*, **43**, 668-698.
6. Phan, A. (2010) Human telomeric G-quadruplex: structures of DNA and RNA sequences. *Febs Journal*, **277**, 1107-1117.
7. Masiero, S., Trotta, R., Pieraccini, S., De Tito, S., Perone, R., Randazzo, A. and Spada, G.P. (2010) A non-empirical chromophoric interpretation of CD spectra of DNA G-quadruplex structures. *Org Biomol Chem*, **8**, 2683-2692.
8. Kelley, S., Boroda, S., Musier-Forsyth, K. and Kankia, B.I. (2011) HIV-integrase aptamer folds into a parallel quadruplex: a thermodynamic study. *Biophys Chem*, **155**, 82-88.
9. Jing, N. and Hogan, M.E. (1998) Structure-activity of tetrad-forming oligonucleotides as a potent anti-HIV therapeutic drug. *J Biol Chem*, **273**, 34992-34999.
10. Watson, J. (1972) Origin of concatemeric T7 DNA. *Nature-New Biology*, **239**, 197-&.
11. Olovnikov, A.M. (1973) A theory of marginotomy. The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon. *J Theor Biol*, **41**, 181-190.
12. Danielsson, O. (2009). Karolinska Institutet, Stockholm.
13. Hayflick, L. (1965) The limited *in vitro* lifetime of human diploid cell strains. *Exp Cell Res*, **37**, 614-636.
14. Shay, J.W. and Bacchetti, S. (1997) A survey of telomerase activity in human cancer. *Eur J Cancer*, **33**, 787-791.
15. Wai, L.K. (2004) Telomeres, telomerase, and tumorigenesis--a review. *MedGenMed*, **6**, 19.
16. Cesare, A.J. and Reddel, R.R. (2010) Alternative lengthening of telomeres: models, mechanisms and implications. *Nat Rev Genet*, **11**, 319-330.
17. Makarov, V., Hirose, Y. and Langmore, J. (1997) Long G tails at both ends of human chromosomes suggest a C strand degradation mechanism for telomere shortening. *Cell*, **88**, 657-666.
18. Wang, Y. and Patel, D. (1993) Solution Structure of the human telomeric repeat $D[AG_3(T_2AG_3)_3]$ G-tetraplex. *Structure*, **1**, 263-282.
19. Parkinson, G.N., Lee, M.P. and Neidle, S. (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*, **417**, 876-880.
20. Xu, Y., Noguchi, Y. and Sugiyama, H. (2006) The new models of the human telomere d[AGGG(TTAGGG)(3)] in K+ solution. *Bioorganic & Medicinal Chemistry*, **14**, 5584-5591.

21. Phan, A.T., Luu, K.N. and Patel, D.J. (2006) Different loop arrangements of intramolecular human telomeric (3+1) G-quadruplexes in K+ solution. *Nucleic Acids Res*, **34**, 5715-5719.
22. Dai, J., Carver, M., Punchihewa, C., Jones, R. and Yang, D. (2007) Structure of the Hybrid-2 type intramolecular human telomeric G-quadruplex in K+ solution: insights into structure polymorphism of the human telomeric sequence. *Nucleic Acids Research*, **35**, 4927-4940.
23. Phan, A., Kuryavyi, V., Luu, K. and Patel, D. (2007) Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K+ solution. *Nucleic Acids Research*, **35**, 6517-6525.
24. Lim, K., Amrane, S., Bouaziz, S., Xu, W., Mu, Y., Patel, D., Luu, K. and Phan, A. (2009) Structure of the Human Telomere in K+ Solution: A Stable Basket-Type G-Quadruplex with Only Two G-Tetrad Layers. *Journal of the American Chemical Society*, **131**, 4301-4309.
25. Zaug, A., Podell, E. and Cech, T. (2005) Human POT1 disrupts telomeric G-quadruplexes allowing telomerase extension in vitro. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 10864-10869.
26. Hwang, H., Buncher, N., Opresko, P.L. and Myong, S. (2012) POT1-TPP1 regulates telomeric overhang structural dynamics. *Structure*, **20**, 1872-1880.
27. Smith, J., Chen, Q., Yatsunyk, L., Nicoludis, J., Garcia, M., Kranaster, R., Balasubramanian, S., Monchaud, D., Teulade-Fichou, M., Abramowitz, L. *et al.* (2011) Rudimentary G-quadruplex-based telomere capping in Saccharomyces cerevisiae. *Nature Structural & Molecular Biology*, **18**, 478-U119.
28. Paeschke, K., Simonsson, T., Postberg, J., Rhodes, D. and Lipps, H. (2005) Telomere end-binding proteins control the formation of G-quadruplex DNA structures in vivo. *Nature Structural & Molecular Biology*, **12**, 847-854.
29. Giraldo, R. and Rhodes, D. (1994) The yeast telomere-binding protein RAP1 binds to and promotes the formation of DNA quadruplexes in telomeric DNA. *Embo Journal*, **13**, 2411-2420.
30. Fang, G. and Cech, T.R. (1993) The beta subunit of Oxytricha telomere-binding protein promotes G-quartet formation by telomeric DNA. *Cell*, **74**, 875-885.
31. Biffi, G., Tannahill, D., McCafferty, J. and Balasubramanian, S. (2013) Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat Chem*, **5**, 182-186.
32. Schaffitzel, C., Berger, I., Postberg, J., Hanes, J., Lipps, H.J. and Plückthun, A. (2001) In vitro generated antibodies specific for telomeric guanine-quadruplex DNA react with Stylonychia lemnae macronuclei. *Proc Natl Acad Sci U S A*, **98**, 8572-8577.
33. Bochman, M.L., Paeschke, K. and Zakian, V.A. (2012) DNA secondary structures: stability and function of G-quadruplex structures. *Nat Rev Genet*, **13**, 770-780.
34. Paeschke, K., Capra, J.A. and Zakian, V.A. (2011) DNA replication through G-quadruplex motifs is promoted by the Saccharomyces cerevisiae Pif1 DNA helicase. *Cell*, **145**, 678-691.
35. London, T.B., Barber, L.J., Mosedale, G., Kelly, G.P., Balasubramanian, S., Hickson, I.D., Boulton, S.J. and Hiom, K. (2008) FANCJ is a structure-specific DNA helicase associated with the maintenance of genomic G/C tracts. *J Biol Chem*, **283**, 36132-36139.

36. Huppert, J.L. and Balasubramanian, S. (2007) G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res*, **35**, 406-413.
37. Siddiqui-Jain, A., Grand, C.L., Bearss, D.J. and Hurley, L.H. (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc Natl Acad Sci U S A*, **99**, 11593-11598.
38. Sun, D. and Hurley, L.H. (2009) The importance of negative superhelicity in inducing the formation of G-quadruplex and i-motif structures in the c-Myc promoter: implications for drug targeting and control of gene expression. *J Med Chem*, **52**, 2863-2874.
39. Shirude, P.S., Okumus, B., Ying, L., Ha, T. and Balasubramanian, S. (2007) Single-molecule conformational analysis of G-quadruplex formation in the promoter DNA duplex of the proto-oncogene c-kit. *J Am Chem Soc*, **129**, 7484-7485.
40. Cao, K., Ryvkin, P. and Johnson, F.B. (2012) Computational detection and analysis of sequences with duplex-derived interstrand G-quadruplex forming potential. *Methods*, **57**, 3-10.
41. Zahler, A.M., Williamson, J.R., Cech, T.R. and Prescott, D.M. (1991) Inhibition of telomerase by G-quartet DNA structures. *Nature*, **350**, 718-720.
42. Gunaratnam, M., Greciano, O., Martins, C., Reszka, A., Schultes, C., Morjani, H., Riou, J. and Neidle, S. (2007) Mechanism of acridine-based telomerase inhibition and telomere shortening. *Biochemical Pharmacology*, **74**, 679-689.
43. Phatak, P., Cookson, J., Dai, F., Smith, V., Gartenhaus, R., Stevens, M. and Burger, A. (2007) Telomere uncapping by the G-quadruplex ligand RHPS4 inhibits clonogenic tumour cell growth in vitro and in vivo consistent with a cancer stem cell targeting mechanism. *British Journal of Cancer*, **96**, 1223-1233.
44. Gomez, D., Wenner, T., Brassart, B., Douarre, C., O'Donohue, M.F., El Khoury, V., Shin-Ya, K., Morjani, H., Trentesaux, C. and Riou, J.F. (2006) Telomestatin-induced telomere uncapping is modulated by POT1 through G-overhang extension in HT1080 human tumor cells. *J Biol Chem*, **281**, 38721-38729.
45. d'Adda di Fagagna, F., Reaper, P.M., Clay-Farrace, L., Fiegler, H., Carr, P., Von Zglinicki, T., Saretzki, G., Carter, N.P. and Jackson, S.P. (2003) A DNA damage checkpoint response in telomere-initiated senescence. *Nature*, **426**, 194-198.
46. Cogoi, S., Shchekotikhin, A.E., Membrino, A., Sinkevich, Y.B. and Xodo, L.E. (2013) Guanidino anthrathiophenediones as G-quadruplex binders: uptake, intracellular localization and anti Harvey-ras gene activity in bladder cancer cells. *J Med Chem*.
47. Monchaud, D. and Teulade-Fichou, M.P. (2008) A hitchhiker's guide to G-quadruplex ligands. *Org Biomol Chem*, **6**, 627-636.
48. Bhattacharjee, A.J., Ahluwalia, K., Taylor, S., Jin, O., Nicoludis, J.M., Buscaglia, R., Brad Chaires, J., Kornfilt, D.J., Marquardt, D.G. and Yatsunyk, L.A. (2011) Induction of G-quadruplex DNA structure by Zn(II) 5,10,15,20-tetrakis(N-methyl-4-pyridyl)porphyrin. *Biochimie*, **93**, 1297-1309.
49. Redon, S., Bombard, S., Elizondo-Riojas, M.A. and Chottard, J.C. (2003) Platinum cross-linking of adenines and guanines on the quadruplex structures of the AG3(T2AG3)3 and (T2AG3)4 human telomere sequences in Na+ and K+ solutions. *Nucleic Acids Res*, **31**, 1605-1613.

50. Bertrand, H., Bombard, S., Monchaud, D. and Teulade-Fichou, M. (2007) A platinum-quinacridine hybrid as a G-quadruplex ligand. *Journal of Biological Inorganic Chemistry*, **12**, 1003-1014.
51. Shin-ya, K., Wierzba, K., Matsuo, K., Ohtani, T., Yamada, Y., Furihata, K., Hayakawa, Y. and Seto, H. (2001) Telomestatin, a novel telomerase inhibitor from Streptomyces anulatus. *J Am Chem Soc*, **123**, 1262-1263.
52. Doi, T., Yoshida, M., Shin-ya, K. and Takahashi, T. (2006) Total synthesis of (R)-telomestatin. *Org Lett*, **8**, 4165-4167.
53. Doi, T., Shibata, K., Yoshida, M., Takagi, M., Tera, M., Nagasawa, K., Shin-ya, K. and Takahashi, T. (2011) (S)-stereoisomer of telomestatin as a potent G-quadruplex binder and telomerase inhibitor. *Org Biomol Chem*, **9**, 387-393.
54. de Matteis, F., Gibbs, A.H., Jackson, A.H. and Weerasinghe, S. (1980) Conversion of liver haem into N-substituted porphyrins or green pigments. Nature of the substituent at the pyrrole nitrogen atom. *FEBS Lett*, **119**, 109-112.
55. De Matteis, F., Gibbs, A.H. and Smith, A.G. (1980) Inhibition of protohaem ferro-lyase by N-substituted porphyrins. Structural requirements for the inhibitory effect. *Biochem J*, **189**, 645-648.
56. Lecerof, D., Fodje, M., Hansson, A., Hansson, M. and Al-Karadaghi, S. (2000) Structural and mechanistic basis of porphyrin metallation by ferrochelatase. *J Mol Biol*, **297**, 221-232.
57. Cochran, A.G. and Schultz, P.G. (1990) Antibody-catalyzed porphyrin metallation. *Science*, **249**, 781-783.
58. Li, Y., Geyer, C.R. and Sen, D. (1996) Recognition of anionic porphyrins by DNA aptamers. *Biochemistry*, **35**, 6911-6922.
59. Arthanari, H., Basu, S., Kawano, T.L. and Bolton, P.H. (1998) Fluorescent dyes specific for quadruplex DNA. *Nucleic Acids Res*, **26**, 3724-3728.
60. Ren, J. and Chaires, J. (1999) Sequence and structural selectivity of nucleic acid binding ligands. *Biochemistry*, **38**, 16067-16075.
61. Hershman, S.G., Chen, Q., Lee, J.Y., Kozak, M.L., Yue, P., Wang, L.S. and Johnson, F.B. (2008) Genomic distribution and functional analyses of potential G-quadruplex-forming sequences in Saccharomyces cerevisiae. *Nucleic Acids Res*, **36**, 144-156.
62. Smith, J. and Johnson, F. (2010) *G-quadruplex DNA*.
63. Scolaro, L., Castriciano, M., Romeo, A., Patane, S., Cefali, E. and Allegrini, M. (2002) Aggregation behavior of protoporphyrin IX in aqueous solutions: Clear evidence of vesicle formation. *Journal of Physical Chemistry B*, **106**, 2453-2459.
64. Pasternack, R.F., Gibbs, E.J. and Villafranca, J.J. (1983) Interactions of porphyrins with nucleic acids. *Biochemistry*, **22**, 5409-5417.
65. Mergny, J., Li, J., Lacroix, L., Amrane, S. and Chaires, J. (2005) Thermal difference spectra: a specific signature for nucleic acid structures. *Nucleic Acids Research*, **33**.
66. Ramsay, G.D. and Eftink, M.R. (1994) Analysis of multidimensional spectroscopic data to monitor unfolding of proteins. *Methods Enzymol*, **240**, 615-645.
67. Job, P. (1928) Studies on the formation of complex minerals in solution and on their stability. *Annales De Chimie France*, **9**, 113-203.
68. De Cian, A., Guittat, L., Kaiser, M., Saccà, B., Amrane, S., Bourdoncle, A., Alberti, P., Teulade-Fichou, M.P., Lacroix, L. and Mergny, J.L. (2007) Fluorescence-based melting assays for studying quadruplex ligands. *Methods*, **42**, 183-195.

69. Hill, Z. and Maccarthy, P. (1986) Novel approach to Job method. *Journal of Chemical Education*, **63**, 162-167.
70. Nicoludis, J.M., Barrett, S.P., Mergny, J.L. and Yatsunyk, L.A. (2012) Interaction of human telomeric DNA with N-methyl mesoporphyrin IX. *Nucleic Acids Res*, **40**, 5432-5447.
71. Hamon, F., Largy, E., Guédin-Beaurepaire, A., Rouchon-Dagois, M., Sidibe, A., Monchaud, D., Mergny, J.L., Riou, J.F., Nguyen, C.H. and Teulade-Fichou, M.P. (2011) An acyclic oligoheteroaryle that discriminates strongly between diverse G-quadruplex topologies. *Angew Chem Int Ed Engl*, **50**, 8745-8749.
72. Nicoludis, J.M., Miller, S.T., Jeffrey, P.D., Barrett, S.P., Rablen, P.R., Lawton, T.J. and Yatsunyk, L.A. (2012) Optimized end-stacking provides specificity of N-methyl mesoporphyrin IX for human telomeric G-quadruplex DNA. *J Am Chem Soc*, **134**, 20446-20456.
73. Larson, J. and Mcmahon, T. (1984) Gas-Phase bihalide and pseudobihalide ions. *Inorganic Chemistry*, **23**, 2029-2033.
74. Ortiz de Montellano, P.R., Beilan, H.S. and Kunze, K.L. (1981) N-Methylprotoporphyrin IX: chemical synthesis and identification as the green pigment produced by 3,5-diethoxycarbonyl-1,4-dihydrocollidine treatment. *Proc Natl Acad Sci U S A*, **78**, 1490-1494.
75. Miyoshi, D., Nakao, A. and Sugimoto, N. (2002) Molecular crowding regulates the structural switch of the DNA G-quadruplex. *Biochemistry*, **41**, 15017-15024.
76. Huppert, J.L. (2010) Structure, location and interactions of G-quadruplexes. *FEBS J*, **277**, 3452-3458.
77. Kumar, N. and Maiti, S. (2008) Role of molecular crowding in perturbing quadruplex-Watson Crick duplex equilibrium. *Nucleic Acids Symp Ser (Oxf)*, 157-158.
78. Bucek, P., Jaumot, J., Aviñó, A., Eritja, R. and Gargallo, R. (2009) pH-Modulated Watson-Crick duplex-quadruplex equilibria of guanine-rich and cytosine-rich DNA sequences 140 base pairs upstream of the c-kit transcription initiation site. *Chemistry*, **15**, 12663-12671.
79. Santangelo, G.M. and Tornow, J. (1997) A Saccharomyces cerevisiae mitochondrial DNA fragment activates Reg1p-dependent glucose-repressible transcription in the nucleus. *Curr Genet*, **32**, 389-398.
80. Cantor, C.R., Warshaw, M.M. and Shapiro, H. (1970) Oligonucleotide interactions. III. Circular dichroism studies of the conformation of deoxyoligonucleolides. *Biopolymers*, **9**, 1059-1077.
81. (2012) In program, P. (ed.), *Scitools*. IDT, Vol. 2013.
82. Tinoco, I. (1960) Hypochromism in polynucleotides. *Journal of the American Chemical Society*, **82**, 4785-4790.
83. Elharakany, A., Halim, F. and Barakat, A. (1984) Dissociation-constants and related thermodynamic quantities of the protonated acid form of tris-(hydroxymethyl)-aminomethane in mixtures of 2-methoxyethanol and water at different temperatures. *Journal of Electroanalytical Chemistry*, **162**, 285-305.
84. SantaLucia, J. (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc Natl Acad Sci U S A*, **95**, 1460-1465.

85. Guedin, A., Gros, J., Alberti, P. and Mergny, J. (2010) How long is too long? Effects of loop size on G-quadruplex stability. *Nucleic Acids Research*, **38**, 7858-7868.
86. Kumar, N., Sahoo, B., Varun, K.A. and Maiti, S. (2008) Effect of loop length variation on quadruplex-Watson Crick duplex competition. *Nucleic Acids Res*, **36**, 4433-4442.
87. Schaefer, A., Taylor, R., Turnbull, D. and Chinnery, P. (2004) The epidemiology of mitochondrial disorders - past, present and future. *Biochimica Et Biophysica Acta-Bioenergetics*, **1659**, 115-120.
88. Schaefer, A., McFarland, R., Blakely, E., He, L., Whittaker, R., Taylor, R., Chinnery, P. and Turnbull, D. (2008) Prevalence of mitochondrial DNA disease in adults. *Annals of Neurology*, **63**, 35-39.
89. Schon, E.A., Rizzuto, R., Moraes, C.T., Nakase, H., Zeviani, M. and DiMauro, S. (1989) A direct repeat is a hotspot for large-scale deletion of human mitochondrial DNA. *Science*, **244**, 346-349.
90. Mita, S., Rizzuto, R., Moraes, C.T., Shanske, S., Arnaudo, E., Fabrizi, G.M., Koga, Y., DiMauro, S. and Schon, E.A. (1990) Recombination via flanking direct repeats is a major cause of large-scale deletions of human mitochondrial DNA. *Nucleic Acids Res*, **18**, 561-567.
91. Samuels, D.C., Schon, E.A. and Chinnery, P.F. (2004) Two direct repeats cause most human mtDNA deletions. *Trends Genet*, **20**, 393-398.
92. Damas, J., Carneiro, J., Goncalves, J., Stewart, J., Samuels, D., Amorim, A. and Pereira, F. (2012) Mitochondrial DNA deletions are associated with non-B DNA conformations. *Nucleic Acids Research*, **40**, 7606-7621.
93. Needleman, H. (2004) Lead poisoning. *Annual Review of Medicine*, **55**, 209-222.
94. Brown, R.S., Hingerty, B.E., Dewan, J.C. and Klug, A. (1983) Pb(II)-catalysed cleavage of the sugar-phosphate backbone of yeast tRNAPhe--implications for lead toxicity and self-splicing RNA. *Nature*, **303**, 543-546.
95. Holtzman, D., Devries, C., Nguyen, H., Olson, J. and Bensch, K. (1982) Maturation of resistance to lead encephalopathy - cellular and subcellular mechanisms. *Neurotoxicology*, **3**, 115-116.

Appendix

Contents

Fitting CD melting curves using Two-State Model

In a two-state model, there is equilibrium between the species in folded (native, N) and unfolded (U) state with the equilibrium constant K. No other states exist.

N⇔ U, $K = \frac{[U]}{[N]}$ (1)

This results in the reversibility of the melting process. The system is a true two-state system if a) the forward and reverse melting curves are identical (no hysteresis) b) if two or more techniques (f. ex. UV/vis and CD) give consistent results for the unfolding of a macromolecule.
In a two state model both folded and unfolded macrmolecules contribute to the observe parameter (Y, in the case of CD Y = mdeg):

$Y = F_N Y_N + F_U Y_U$ (2)
Where F_N and F_U are fractions of folded and unfolded macromolecule, such that $F_N+F_U = 1$ (3); Y_N and Y_U are absorbances of folded and unfolded macromolecule. These absorbances usually have linear temperature dependence. From equation (1) and (3)

$F_N = \frac{1}{1+K}$ and $F_U = \frac{K}{1+K}$ (4)

Now putting (2) and (4) together,

$$Y = \frac{1}{1+K} Y_N + \frac{K}{1+K} Y_U \qquad (5)$$

Finally, from the expressions, $\Delta G = -RT\ln K$, $\Delta G = \Delta H - T\Delta S$, $\Delta S = \frac{\Delta H}{T_m}$ (6),

(where T_m is a temperature at which half molecules are in the unfolded state) we can derive the expression for K,

$$K = e^{-\frac{\Delta H(1-\frac{T}{T_m})}{RT}} \qquad (7)$$

Combining (5) and (7) gives us the final equation for the fitting, taking into account that ΔH is constant.

$$Y = \frac{1}{1+e^{-\frac{\Delta H(1-\frac{T}{T_m})}{RT}}} \cdot Y_N + \frac{e^{-\frac{\Delta H(1-\frac{T}{T_m})}{RT}}}{1+e^{-\frac{\Delta H(1-\frac{T}{T_m})}{RT}}} \cdot Y_U \qquad (8)$$

User 1 function in Origin: ΔH is temperature independent
y=(P3+P4*x)*(1/(1+ K))+(P5+P6*x)*(K/(1+K))
now taking into account that
K = exp((-P1*(1-x/P2))/(8.31451*x)) we have:

y=(P3+P4*x)*(1/(1+exp((-P1*(1-x/P2))/(8.31451*x))))+(P5+P6*x)*(exp((-P1*(1-x/P2))/(8.31451*x))/(1+exp((-P1*(1-x/P2))/(8.31451*x))))

Definition of parameters:

P1 = ΔH, J/mol
P2 = T_m, K
P3, P4 – intercept and slope for an absorbance of a folded G-quadruplex (or any other macromolecule); Y_F = P3+P4*T(K)
P5, P6 – intercept and slope for an absorbance of an unfolded G-quadruplex
Y_U = P5+P6*T(K)

$$F_U = \frac{Y_F - Abs}{Y_F - Y_U}$$

However, if ΔH depends upon temperature (usually it is true when T range is quite wide), then this dependence have to be taken into account.

$\Delta H = \Delta H_m + \Delta C_P(T-T_m)$ (9)

$$\Delta S = \frac{\Delta H_m}{T_m} + \Delta C_P \ln\frac{T}{T_m} \qquad (10)$$

Combining (9) and (10) gives us new expression for ΔG:

$$\Delta G = \Delta H - T\Delta S = \Delta H_m + \Delta C_P(T-T_m) - T(\frac{\Delta H_m}{T_m} + \Delta C_P \ln\frac{T}{T_m}) = \Delta H_m(1-\frac{T}{T_m}) + \Delta C_P(T - T_m - T\ln\frac{T}{T_m})$$

(11)

Expressions (11) and ΔG = -RTlnK are used to derive K:

$$K = e^{-\frac{\Delta H_m(1-\frac{T}{T_m})+\Delta C_P(T-T_m-T\ln\frac{T}{T_m})}{RT}} \qquad (12)$$

Combining (12) and (5) results in the desired final equation for the fit.

User 2: ΔH is a function of temperature

y=(P3+P4*x)*(1/(1+exp(-(P1*(1-x/P2)+P7*(x-P2-x*ln(x/P2)))/(8.31451*x))))+(P5+P6*x)*(1-(1/(1+exp(-(P1*(1-x/P2)+P7*(x-P2-x*ln(x/P2)))/(8.31451*x)))))

Definition of parameters:

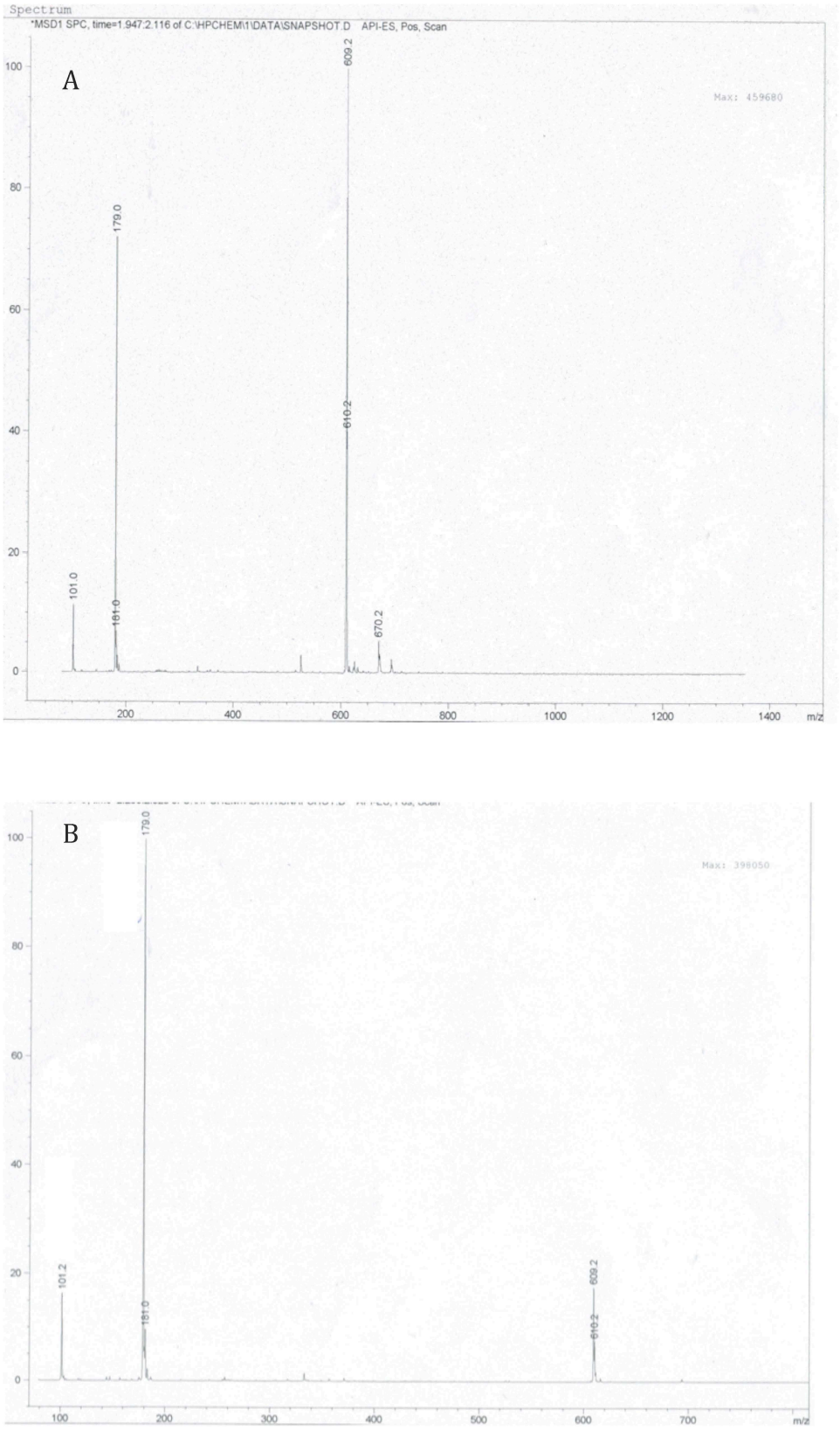
P1 = ΔH_m, J/mol, enthalpy at T = T_m
P7 - ΔCp – heat capacity, J/molK

Finally, to find free energy of activation combine the following formulas:

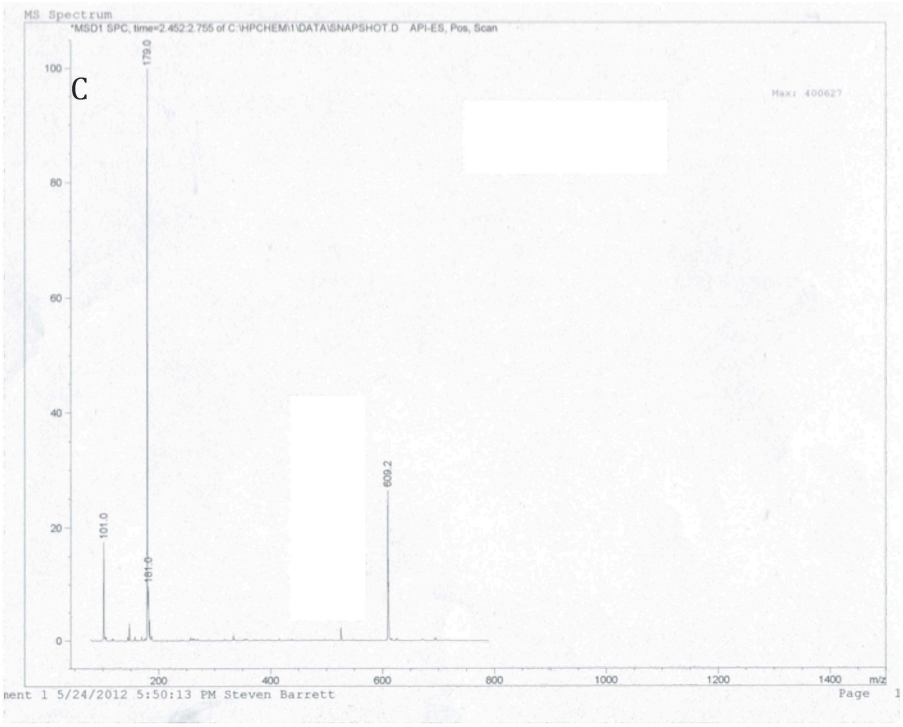
ΔG = ΔH-TΔS, $\Delta S = \frac{\Delta H}{T_m}$

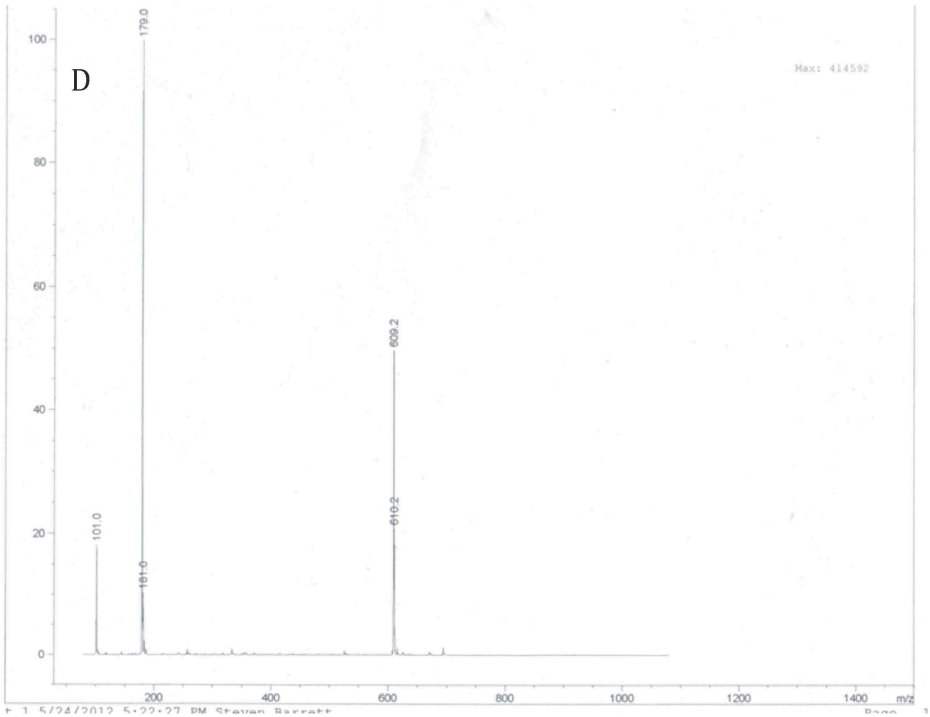
You will get:

$$\Delta G_{unf\ olding}(25^o C) = \Delta H_m(1 - \frac{T}{T_m})$$

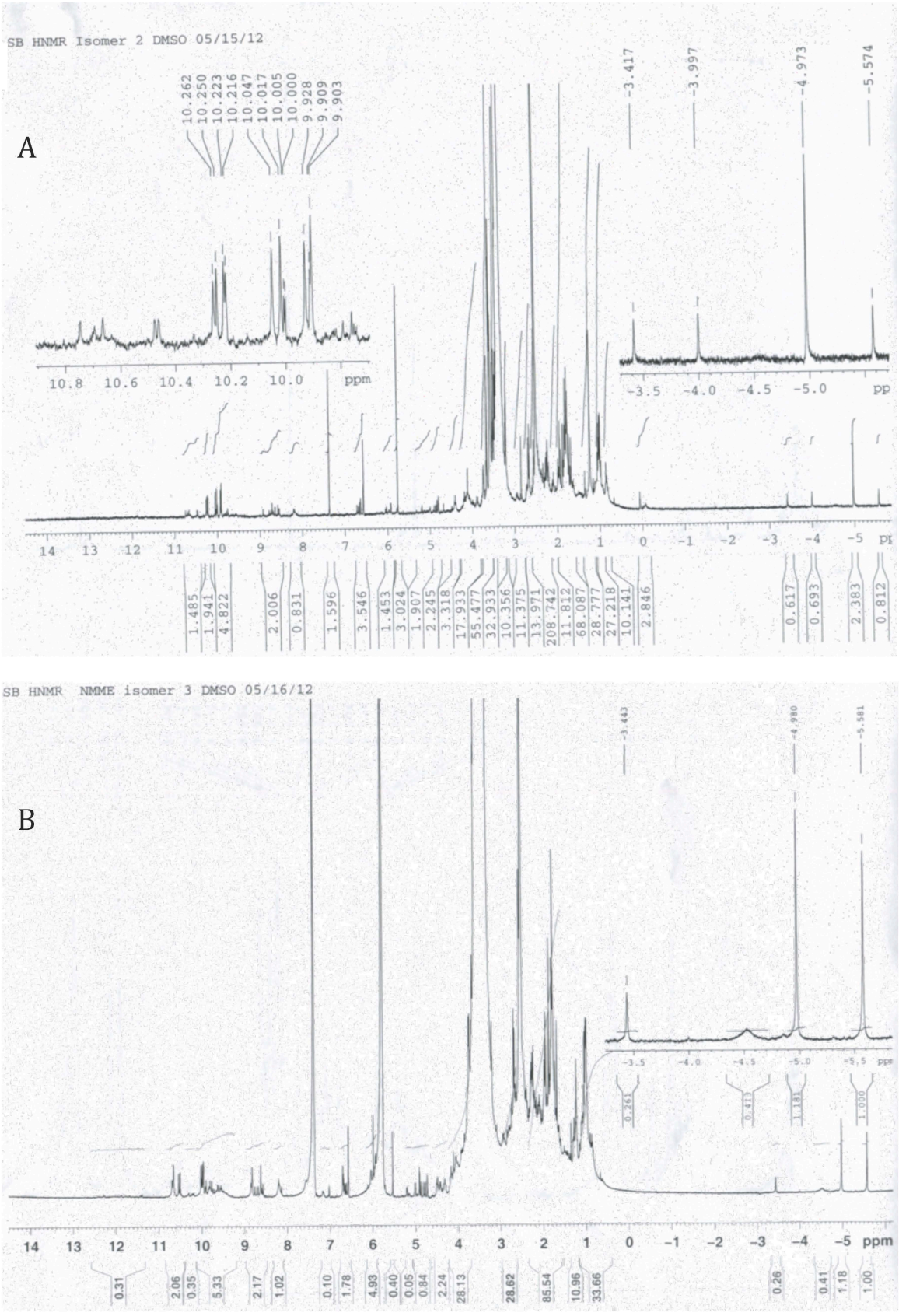


ESI MS data for NMME **(A)** Isomer 1 and **(B)** Isomer 2. Mass spectrum of the commercial mixture of NMME isomers. Peak at m/z=609.2 corresponds to the exact mass of NMME.

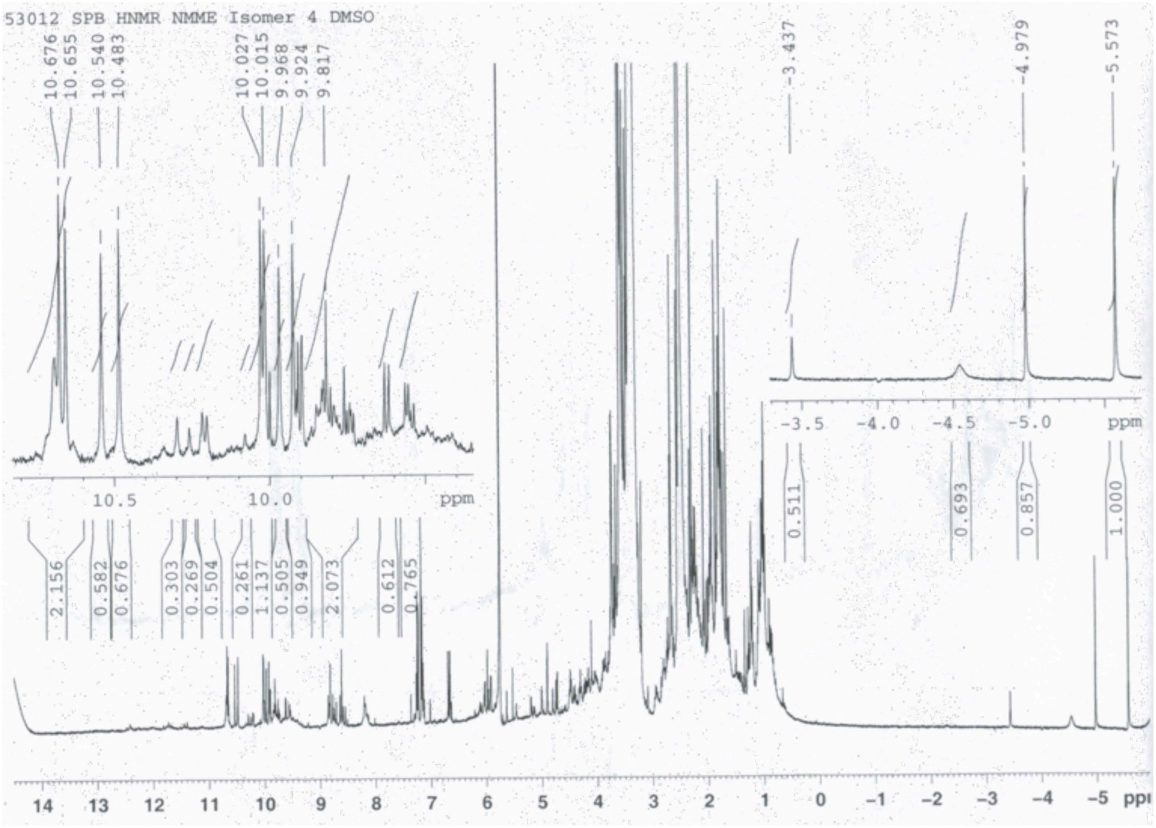




ESI MS data for NMME **(C)** Isomer 3 and **(D)** Isomer 4. Mass spectrum of the commercial mixture of NMME isomers. Peak at m/z=609.2 corresponds to the exact mass of NMME.

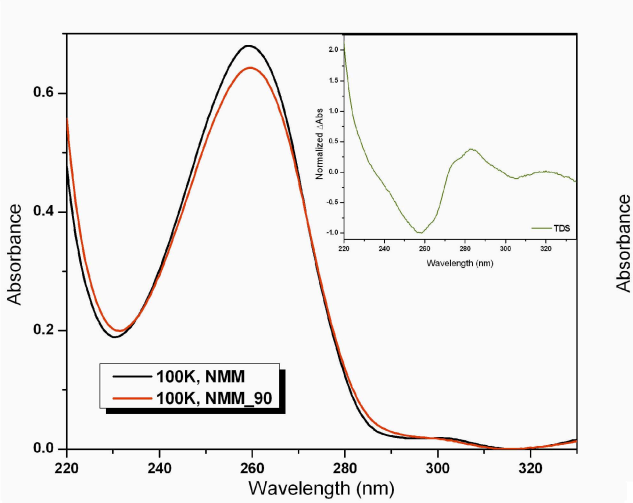


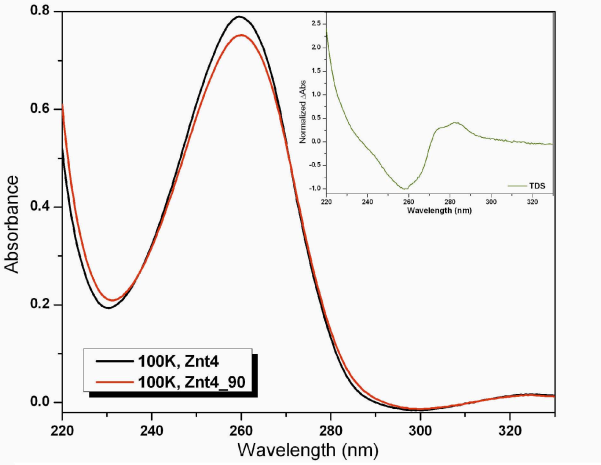
[1]H-NMR spectrum of NMME **(A)** Isomer 2 and **(B)** Isomer 3 in d6-DMSO.

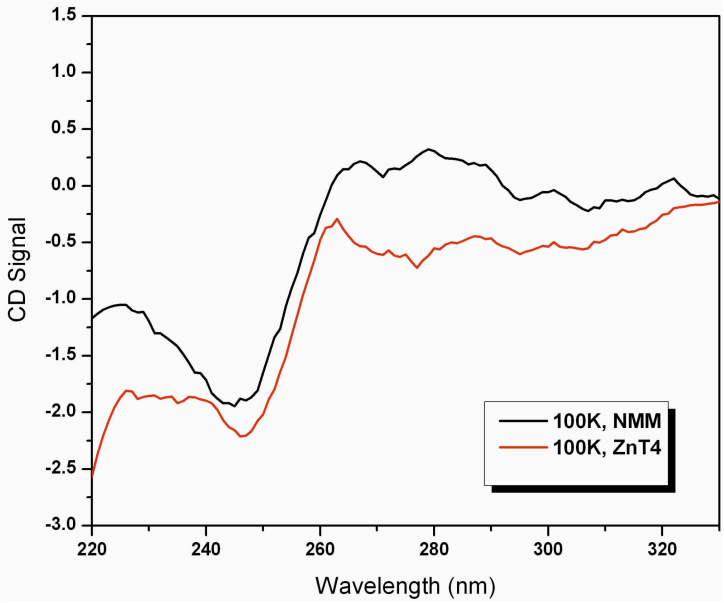


[1]H-NMR spectrum of NMME Isomer 4 in d6-DMSO.

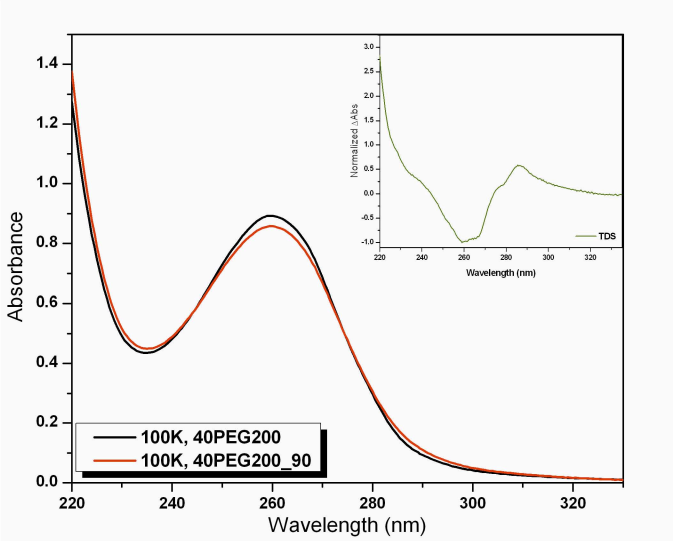
ddiQFP Spectroscopic Data. Absorbance spectra of ddiQFP at both 90°C and RT are shown below with TDS spectra depicted in the insets. [ddiQFP] ranged from 2.3 to 20.5 µM. All measurements were taken in 1mm or 1cm cuvettes, depending on the concentration of the oligo. The buffer used was 10 mM LiCac (pH=7.2) with various other components added (noted in the figure legends). 100K= 10 mM LiCac, 100 mM KCl; 100Na= 10 mM LiCac, 100 mM NaCl.

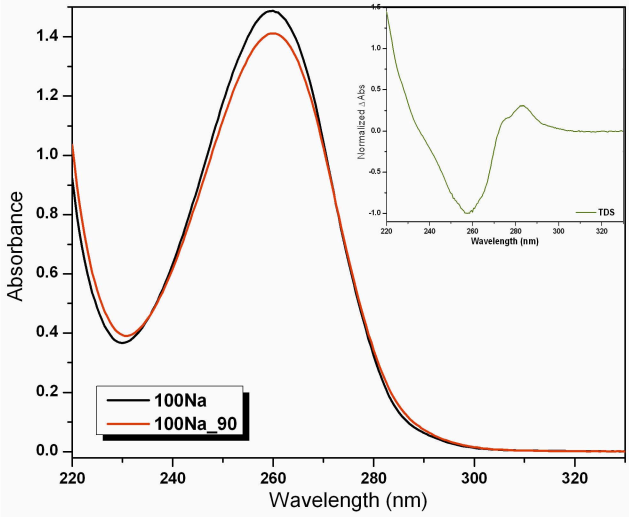


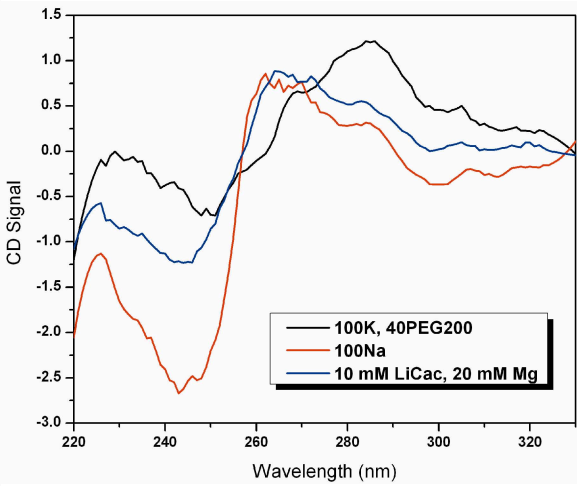


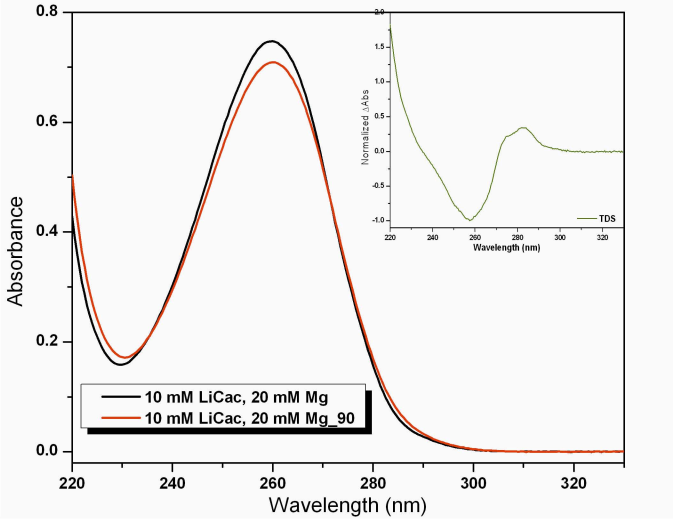


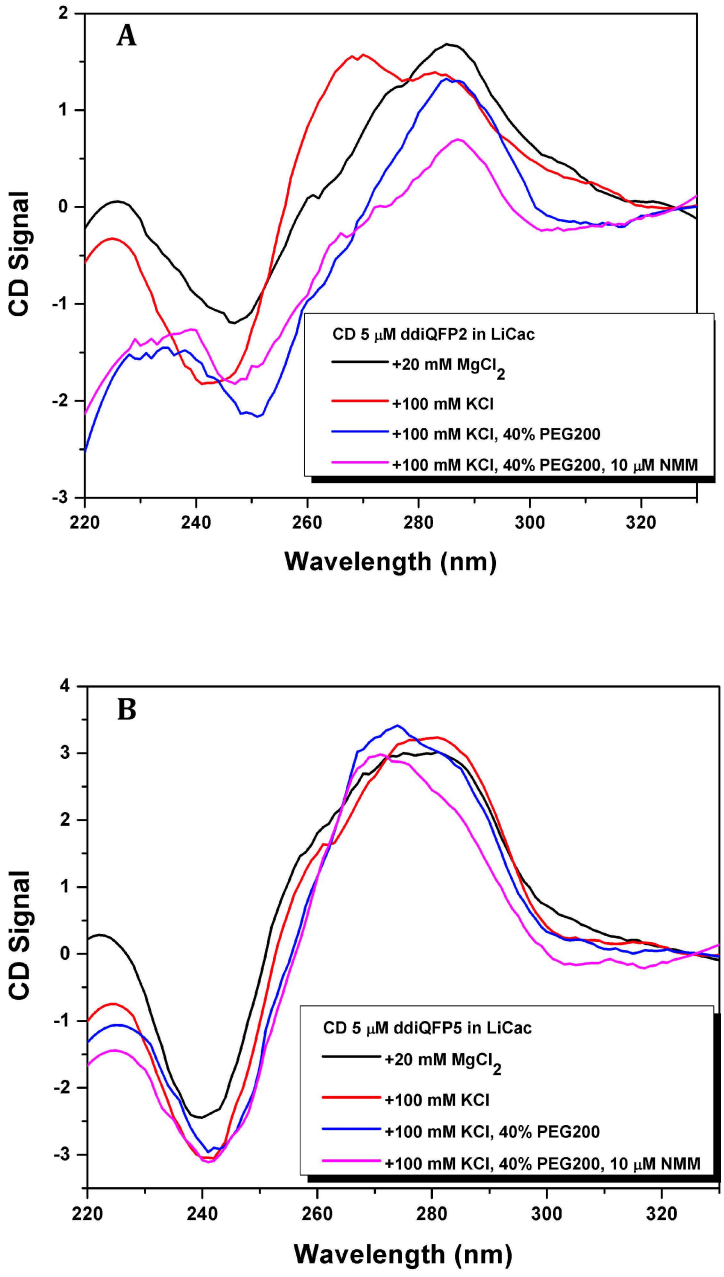
Note: [Porphyrin]:[ddiQFP] = 2:1


Absorbance
Wavelength (nm)
100K, 40PEG200
100K, 40PEG200_90
Normalized ΔAbs
TDS
Wavelength (nm)


Absorbance
Wavelength (nm)
100Na
100Na_90
Normalized ΔAbs
TDS
Wavelength (nm)


CD Signal
Wavelength (nm)
100K, 40PEG200
100Na
10 mM LiCac, 20 mM Mg


Absorbance
Wavelength (nm)
10 mM LiCac, 20 mM Mg
10 mM LiCac, 20 mM Mg_90
Normalized ΔAbs
TDS
Wavelength (nm)



CD spectra of 5 μM **(A)** ddiQFP2 and **(B)** ddiQFP5 taken in 10 mM LiCac, pH=7.2 along with the conditions specified in the legend.

A Copy of the Paper "Interaction of human telomeric DNA with N- methyl mesoporphyrin IX"

Reference:
Nicoludis, J. M.; Barrett, S. P.; Mergny, J. L.; Yatsunyk, L. A., Interaction of human telomeric DNA with N-methyl mesoporphyrin IX. *Nucleic Acids Res.* **2012**, *40, 5432-5447.*

Nucleic Acids Research Advance Access published February 23, 2012

Nucleic Acids Research, 2012, 1–16
doi:10.1093/nar/gks152

Interaction of human telomeric DNA with *N*-methyl mesoporphyrin IX

John M. Nicoludis[1], Steven P. Barrett[1], Jean-Louis Mergny[2,3] and Liliya A. Yatsunyk[1,*]

[1]Department of Chemistry and Biochemistry, Swarthmore College, 500 College Avenue, Swarthmore, PA 19081, USA, [2]University of Bordeaux, ARNA Laboratory, IECB, F-33000 Bordeaux and [3]INSERM, U869, ARNA Laboratory, European Institute of Chemistry and Biology, 2 rue Robert Escarpit, F-33000 Pessac, France

Received November 4, 2011; Revised January 12, 2012; Accepted January 25, 2012

ABSTRACT

The remarkable selectivity of *N*-methyl mesoporphyrin IX (NMM) for G-quadruplexes (GQs) is long known, however its ability to stabilize and bind GQs has not been investigated in detail. Through the use of circular dichroism, UV-visible spectroscopy and fluorescence resonance energy transfer (FRET) melting assay we have shown that NMM stabilizes human telomeric DNA $dAG_3(TTAG_3)_3$ (Tel22) and is selective for its parallel conformation to which it binds in 1:1 stoichiometry with a binding constant of $\sim 1.0 \times 10^5\ M^{-1}$. NMM does not interact with an antiparallel conformation of Tel22 in sodium buffer and is the second example in the literature, after TOxaPy, of a ligand with an excellent selectivity for a specific GQ structure. NMM's stabilizing ability toward predominantly parallel GQ conformation is universal: it stabilizes a variety of biologically relevant G-rich sequences including telomeres and oncogene promoters. The *N*-methyl group is integral for selectivity and stabilization, as the unmethylated analogue, mesoporphyrin IX, does not stabilize GQ DNA in FRET melting assays. Finally, NMM induces the isomerization of Tel22 into a structure with increased parallel component in K^+ but not in Na^+ buffer. The ability of NMM to cause structural rearrangement and efficient stabilization of Tel22 may bear biological significance.

INTRODUCTION

Guanine-rich regions of genomic DNA are suggested to fold into non-canonical secondary structures known as G-quadruplexes (GQ). GQs are formed by π–π stacking of G-quartets, which are composed of four planar guanines held together by Hoogsteen hydrogen-bonds. The presence of monovalent cations at the core channel of the GQ is necessary to balance charges. Sequences from telomeres, oncogene promoters (*c-myc*, *c-myb*, *c-fos*, *c-abl*) (1,2), Ribosomal DNA, minisatellites and immunoglobulin switch regions (3) form GQs *in vitro*. The genomic distribution of these regions makes GQs biomedically relevant (4–7). Data are being accumulated about *in vivo* existence of GQ structures (8–11).

The folding topology of human telomeric repeat sequence, $d(TTAGGG)_n$, under a variety of conditions has been researched extensively (12–18); for a recent review see Phan (19). In total, at least five different intramolecular GQ structures were reported. In a Na^+-rich environment, $dAG_3(TTAG_3)_3$ sequence, Tel22, forms an antiparallel basket type structure (12), whereas in a crystalline form (high K^+ and high DNA) (13) and in PEG-containing solutions (20) Tel22 adopts a parallel GQ conformation (Figure 1A). Other potentially physiologically relevant structures include the mixed hybrid topologies that have both parallel and antiparallel character, named Forms 1, 2 and 3. They are found in solutions of K^+ and mixtures of K^+ and Na^+ (16,21,22) as well as other cations (Mg^{2+} and Ca^{2+}). Forms 1 and 2 are hybrid (3 + 1) GQs with one parallel and three antiparallel strands. Form 3 is an antiparallel GQ with only two G-tetrads. Tel22 in dilute K^+ buffers, used in this work, is believed to fold into Forms 1 and 2, though it is likely that there is an equilibrium with Form 3 and possibly with an unfolded oligonucleotide. The folding of human telomeric DNA into a specific quadruplex or structural transition from one fold to another is believed to occur during telomere formation and function (23). *In vitro* it could be triggered by the presence of divalent cations (24,25), cosolutes (23,26) and a number of small molecule ligands (27–34) including porphyrins (35).

Small molecule ligands have been developed to bind GQ with high affinity, to effectively stabilize GQ structures and to be selective for quadruplex DNA versus other DNA structures, such as abundant duplex DNA (36). Selectivity for one type of quadruplex versus another also needs to be considered in ligand design. GQ structures could form in different parts of the human genome and play both positive and negative roles in human

*To whom correspondence should be addressed. Tel: 610 328 8558; Fax: 610 328 7355; Email: lyatsun1@swarthmore.edu

biology. For example, GQ formation at the RNA component of human telomerase was shown to play a role in telomerase assembly and stimulate telomere maintenance *in vivo* (37). Hence, ligands with broad quadruplex recognition that were originally designed to inhibit telomerase could increase telomerase processivity and genome instability that might lead to secondary tumors (38). With new screening methods aimed at the detection of a ligands' preference for a given GQ structure (39–41), a variety of quadruplex stabilizers were tested and only a few displayed some preference for a given quadruplex type. Recently, the acyclic oligoheteroaryle, TOxaPy, was shown to recognize human telomeric DNA in Na^+, but not in K^+ buffer (42). TOxaPy is the first example of a ligand with specificity for an antiparallel quadruplex fold.

N-methyl mesoporphyrin IX (NMM, Figure 1B), *N*-core methylated non-planar derivative of mesoporphyrin IX, was first identified as a potent inhibitor of Fe^{2+} insertion into protoporphyrin IX by ferrochelatase, the terminal enzyme in heme biosynthetic pathway (43). Following this work, NMM-selective DNA aptamers were developed (44). Interestingly, these aptamers had significant quadruplex forming potential, suggesting GQ as a possible binding structure for NMM. Hence, NMM was one of the first small molecules reported to bind quadruplexes. In later work, NMM was found to be highly selective for GQ DNA versus ssDNA, dsDNA, dsRNA, RNA–DNA hybrid, Z-DNA, triplex and Holliday junctions (45–47). More recently, NMM has been shown to inhibit quadruplex unwinding by the *Escherichia coli* RecQ helicase and by the eukaryotic helicases BLM and Sgs1p, leaving unwinding of Holliday junctions and partial-duplex substrates largely unaffected (47,48). However, NMM is regarded by some as a weak GQ binder because its negative charge at physiological pH prevents favorable electrostatic interactions with DNA. The details of NMM's interaction with GQ DNA and its stabilizing ability toward GQ structures remain largely unexplored. Yet, NMM's ability to interact with GQs is utilized in biology and chemistry, e.g. in fluorescence screening assays (46,49), in microarrays (50), in affinity chromatography (51) or in quantitative Cu(II) detection (52).

In this work, we explore the details of the interactions between NMM and human telomeric DNA, Tel22, and probe NMM's ability to discriminate between various GQ structures using circular dichroism (CD), UV-vis and fluorescence spectroscopies, and gel electrophoresis. We demonstrate that NMM induces the isomerization of Tel22 from a hybrid to a parallel structure in dilute K^+ buffers; NMM effectively stabilizes parallel-stranded GQs with excellent selectivity (>480-fold compared to dsDNA), has a moderate binding constant ($\sim 1.0 \times 10^5\,M^{-1}$) and displays 1:1 binding stoichiometry in its interactions with Tel22. NMM does not alter or stabilize antiparallel GQ structures of Tel22. Along with TOxaPy, NMM is a unique ligand with highly desired specificity for a selected GQ structure. NMM's selectivity for parallel-stranded GQs and its ability to cause a conformational change in GQ structure, make it a promising platform for further ligand development.

MATERIALS AND METHODS

Porphyrins, oligonucleotides and buffers

All porphyrins were purchased from Porphyrin Products (Logan, UT, USA). NMM and 5,10,15,20–tetrakis (*N*-methyl-4-pyridyl) (TMPyP4) porphyrins were dissolved in water at 5–10 mM and stored at −20°C. *N*-methyl protoporphyrin IX (NMP), mesoporphyrin IX dihydrochloride (MIX) and protoporphyrin IX (PIX) were dissolved in DMSO at 5–10 mM and stored at 4°C. For all experiments, porphyrin stock solutions were freshly diluted with desired buffers, and concentrations were determined using extinction coefficients reported in Supplementary Table S1.

The DNA oligonucleotides used in this work (for complete list see Supplementary Table S1) were purchased from Midland (TX, USA) or from Eurogentec (Seraing, Belgium); the fluorescently labeled oligonucleotides 5′-6-FAM-$G_3(TTAG_3)_3$-Dabcyl-3′ (F21D) and 5′-6-FAM-$G_3(TGTG_3)_3$-Dabcyl-3′ (FGD) were purchased from IDT (Coralville, IA, USA); calf-thymus (CT) DNA was purchased from Sigma Aldrich. DNA $dAG_3(TTAG_3)_3$ (hereafter abbreviated as Tel22) was purified by ethanol precipitation, dissolved in water at 1 mM and stored at 4°C. The concentration of DNA was calculated based on absorption at 260 nm at 90°C using extinction coefficients determined from a nearest neighbor approximation (53) and reported in Supplementary Table S1. For Tel22 this coefficient is 228.5 $mM^{-1}cm^{-1}$. Concentration of all quadruplex-forming sequences is expressed per quadruplex. To induce quadruplex formation, stocks of Tel22 were diluted with an appropriate buffer to 25 μM and annealed by heating to 90°C for 10 min, cooling to room temperature during 2–3 h and equilibrating at 4°C for >12 h. Fully equilibrated thermodynamic Tel22/NMM samples were prepared either by annealing Tel22/NMM mixtures as described earlier followed by equilibration at 30°C for 12 h, or by mere mixing of annealed Tel22 with NMM followed by equilibration at 30°C for ~30 h. The time interval necessary for complete equilibration was determined based on preliminary kinetics experiments (data not shown).

The following buffers were used in this work: 10 mM lithium cacodylate, pH 7.2, 5 mM KCl, 95 mM LiCl (hereafter abbreviated as 5K); 10 mM lithium cacodylate, pH 7.2, 50 mM NaCl, 50 mM LiCl (50Na); 10 mM lithium cacodylate, pH 7.2, 100 mM LiCl (100Li).

UV-vis studies

Thermal difference spectra (TDS) (54) were collected on Cary300 (Varian) spectrophotometer equipped with a Peltier-thermostated cuvette holder (temperature accuracy ± 0.3°C) using a 1-cm quartz cuvette. Tel22 samples annealed and fully equilibrated in 5K, 50Na or 100Li buffer with or without 2 or 4 equivalents of NMM were diluted to a concentration of ~4 μM. UV-vis spectra were collected in the range 220–350 nm at 25°C, then the samples were heated to 90°C for 10 min and the spectra were collected again. The difference between the two spectra, TDS, was calculated and normalized to the highest peak for ease of comparison. TDS were collected also for F21D and FGD in order to confirm that the presence of dyes does not interfere with quadruplex formation. In this case, the oligonucleotides were diluted to 10 μM in the appropriate buffer in a 1-mm quartz cuvette and incubated in the UV-vis cuvette holder at 90°C for 10 min after which UV-vis spectra were collected. The heating was turned off, samples were cooled slowly during 4 h and placed at 4°C overnight. After this equilibration the second UV-vis spectrum was collected at 25°C. Data were analyzed as described earlier.

Titrations of NMM with Tel22 in 5K, 50Na and 100Li buffers were carried out in a 1-cm methacrylate or 2-mm quartz cuvettes at 25°C by the stepwise additions of 0.3–0.7 mM Tel22 into a solution of ~2–5 μM NMM. The solution of Tel22 contained an equal amount of NMM to keep porphyrin concentration constant. Titrations were terminated when two to three consecutive additions of Tel22 yielded the same spectra or when [GQ]/[NMM] exceeded 10. Note, for the case of 5K buffer, titrations performed in this manner reflect 'non-equilibrium' condition due to slow rearrangement of original Tel22 structure to a new fold induced by addition of NMM. In order to obtain 'equilibrium' data UV-vis titrations were repeated in a batch method where NMM samples containing increasing amounts of Tel22 were incubated for 30 hours at 30°C. Accurate stoichiometry of NMM/Tel22 complex was determined using continuous variation analysis, also known as a Job plot (55), following closely the protocol published earlier (35). Concentrations of NMM and Tel22 were matched at 18 μM within 5%.

Interactions of NMM with telomeric and non-telomeric GQ forming oligonucleotides were studied by UV-vis. All oligonucleotides were annealed in 5K or 50Na buffers at ~30 μM in the presence of 2.3 μM NMM at 90°C for 10 min, cooled slowly and incubated at 30°C for >12 h. Both the oligonucleotide concentration and high quadruplex to NMM ratio (~13) were chosen to assure that all NMM molecules were bound to GQ DNA. A sample containing 2.3 μM NMM without DNA was treated in identical way. UV-vis spectra were collected on all samples, and red shift and hypochromicity were determined as described earlier (35).

Circular dichroism

CD experiments were performed on an AVIV 410 spectrometer equipped with a Peltier heating unit (temperature accuracy ± 0.3°C). Spectra were collected between 220 and 330 nm, using three to five scans, a 1-nm bandwidth and an averaging time of 1 s at 25°C in 1-cm or 1-mm quartz cuvettes. Some spectra were collected on JASCO J-815 spectropolarimeter with a 2-nm bandwidth, 500 nm min^{-1} scan speed, 1-nm step and five scans. The data were baseline corrected using the CD spectra of a buffer alone in the same cuvette, zeroed by subtracting the average absorbance between 320 and 330 nm and converted into molar ellipticity using the following equation:

$$\Delta\varepsilon = \frac{\theta}{C \times l \times 3.3 \times 10^4},$$

where θ is the CD signal in mdeg, C is concentration in mol L^{-1} and l is a cuvette pathlength in cm. Data were smoothed using a Savitzky–Golay smoothing filter with a 13-point quadratic function.

CD annealing studies were performed to determine if porphyrins induce structural rearrangement of Tel22 in water, 5K, 50Na and 100Li buffers. Samples of 2.5 μM Tel22 were annealed with NMM, NMP or TMPyP4 (1:1 or 1:2 ratio), cooled slowly and placed in a 30°C incubator for 12 h. In parallel, 2.5 μM Tel22 samples were annealed alone cooled down to 30°C, after which porphyrins were added and samples were incubated at 30°C for ~30 h. As NMM was the only porphyrin capable of changing Tel22 structure, its ability to interact in a similar manner with

other GQ-forming sequences (for full list see Supplementary Table S1) was also tested.

Equilibrium CD titration of Tel22 with NMM in 5K buffer were performed in a batch format where individual samples of 2.0 μM Tel22 were annealed with increasing equivalents of NMM (from 0.5 to 12) and incubated at 30°C for >12 h to achieve thermodynamic equilibrium prior to collecting the CD spectra.

Binding constant determination

Singular value decomposition (SVD) analysis of CD and UV-vis titration data was used to verify that binding of Tel22 to NMM is a two-state process. Details of SVD analysis are provided in Supplementary Data and in a recently published protocol (56). Binding constant, K_a, for NMM—Tel22 interactions was extracted using direct fitting of the titration data to the equation:

$$Y = Y_0 - \frac{x}{C_t}(Y_0 - Y_f), \quad (1)$$

where Y_0 and Y_f are the initial and final signal (absorbances at 379 and 399 nm in UV-vis or molar ellipticity at 264 nm in CD), C_t is the total NMM concentration and x is derived from the equilibrium expression:

$$K_a = \frac{[x]}{[C_s - x] \cdot [C_t - x]}, \quad (2)$$

which is based on the simple two-state 1:1 binding reaction: NMM (C_t) + Tel22 (C_s) → Complex (x). The fits yield K_a, which was averaged for all titrations. This data analysis was done in GraphPad Prism software. In addition, the value of the CD signal at 264 nm was normalized and fitted with the Hill equation:

$$\text{Normalized CD signal at 264 nm} = \frac{K_a[\text{NMM}]^n}{1+K_a[\text{NMM}]^n}. \quad (3)$$

Hill coefficient, n, can be interpreted as the average number of binding sites. This model assumes that all binding sites have the same affinity.

Gel electrophoresis

For gel electrophoresis, samples were annealed and equilibrated in 5K buffer at 40 μM Tel22 and increasing concentration of NMM from 0 to 12 equivalents. Samples were run on 20% non-denaturing polyacrylamide gel in 1 × TBE buffer (89 mM Tris–borate, 2 mM EDTA, pH 8.3) with 5 mM KCl (or 10 mM NaCl or 10 mM LiCl) at 220 V at 14–16°C for 4 h. Oligothymidylate markers 5′ dT_n (where n = 6, 9, 15 and 24) were used as internal migration standards (and not necessarily as length markers) along with duplex markers dx_9: 5′-GCGTATCGG + 5′-CCGAT ACGC and dx_{12}: 5′-GCGTGACTTCGG + 5′-CCGAAG TCACGC. Control samples included NMM alone as well as Tel22 to which NMM was added right before loading the gel. DNA bands were visualized by UV-shadowing and using NMM's fluorescence (excitation and emission maxima at 399 and 610 nm, respectively (49); gels were excited with an LED of 470 nm and visualized utilizing a red filter). Part of each gel sample was diluted to a final concentration of 1–2 μM for CD measurements. Data were collected immediately after dilution with the assumption that CD signature observed corresponds to the Tel22 conformation at high concentration of 40 μM used for the gel, and not to the conformation present at 1–2 μM DNA (if these are indeed different).

To determine the molecularity of the quadruplex bound to NMM, Tel22 was mixed either with T_6Tel22 or Tel22T_6 in 1:1 ratio and annealed with 2 or 8 equivalents of NMM. T_6Tel22 and Tel22T_6 were chosen because they have larger size but secondary structure similar to that of Tel22 as determined by CD. Gels were prepared, loaded, run and visualized in the way described earlier.

Fluorescence resonance energy transfer melting studies

Fluorescence resonance energy transfer (FRET) was performed according to established literature procedure (57) using doubly labeled oligonucleotides F21D, 5′-6-FAM-G_3(TTAG_3)$_3$-Dabcyl-3′ and FGD, 5′-6-FAM-G_3 (TGTG_3)$_3$-Dabcyl-3′. Specific experimental details are presented in Supplementary Data.

CD melting studies

Tel22 was prepared at 3–4 μM in 5K buffer alone and with increasing amounts of NMM. Samples were heated from 15 to 95°C, maintained at 95°C for 5 min and cooled to 15°C at the same rate of 0.27 deg min^{-1}. CD melting was monitored at 264 and 295 nm with 0.5-min equilibration time, 1°C step, 4 s averaging time and 2-nm bandwidth. Other quadruplex forming oligonucleotides (~4 μM, for sequences see Table 1) alone and in the presence of 2 equivalents of NMM were melted in an identical way. Melting data were analyzed assuming linear starting and final baselines. Melting temperatures and enthalpies of unfolding were adjusted to get the best fits (58). For data sets where starting or final baselines were not clearly defined (e.g. when $T_{1/2}$ exceeded 80°C) the melting temperatures were determined using the first derivatives. All melting temperatures reported were obtained from melting curves. All data analysis was done in Origin 8.1.

RESULTS

In this report the details of the interaction between NMM and human telomeric GQ DNA, Tel22, was investigated. NMP, the structural analogue of NMM (Figure 1B), was also included in the study. Both NMM and NMP exist in a monomeric form in the concentration range from 1 to 50 μM (Supplementary Figure S1). One needs to be aware that commercial NMM (and NMP) is a mixture of four isomers that differ by the position of the core *N*-methyl group. Results reported in this article are therefore average of the action produced by four NMM isomers.

Structure of Tel22 alone and in the presence of NMM and other porphyrins

The secondary structure of Tel22 alone under various experimental conditions was determined by CD spectroscopy. In 5K buffer the CD spectrum of Tel22 has two

Table 1. Sequences, bathochromic (red-) shifts (>10-fold excess of DNA), melting temperatures ($T_{1/2}$) and (2 eq.) NMM thermal stabilization values ($\Delta T_{1/2}$) for a variety of GQ-forming oligonucleotides in 5K buffer (unless noted otherwise)

Name	Sequence	Red-shift[a] (nm)	$T_{1/2}$ (°C)[b]	$\Delta T_{1/2}$ (°C)[c]	Conformation w/o NMM[d]	Conformation with NMM[d]
Tel22	AGGGTTAGGGTTAGGGTTAGGG	19.3 ± 0.4	48.0 ± 0.8	7 ± 1	M	P
Tel22, 50Na	AGGGTTAGGGTTAGGGTTAGGG	0.0 ± 0.2	51 ± 2	0 ± 2	A	A
26TelG4	AGGGGTTAGGGGTTAGGGGTTAGGGG	18 ± 1	73.5 ± 0.4	6.8 ± 0.6	A/M	M
26TelG4, 50Na	AGGGGTTAGGGGTTAGGGGTTAGGGG	19.5	64.4 ± 0.9	0 ± 1	A	A
Bcl-2	GGGCGCGGGAGGGAATTGGCGGGG	17.2 ± 0.3	48.9 ± 0.8	20. ± 1	P/M	P
cKit1	GGGAGGGCGCTGGGAGGAGGG	18.6 ± 0.9	44.7 ± 0.6	11.4 ± 0.9	P/M	P
cKit2	GGGCGGGCGCGAGGGAGGGG	19.2 ± 0.4	53.7 ± 0.7	13 ± 1	P	P
cMyc	TGAGGGTGGGTAGGGTGGGTAA	20 ± 1	67.5 ± 0.3	14.3 ± 0.3	P/M	P
G4TERT	AGGGGAGGGGCTGGGAGGGC	17.7 ± 0.3	55 ± 2	18 ± 2	P	P
HIF-1α	GGGAGGGAGGGAAGGAGGGAGGGAGGG	18.5 ± 0.7	65 ± 1	1 ± 1	P	P
TBA	GGTTGGTGTGGTTGG	19.5 ± 0.5	38.7 ± 0.8	1 ± 2	A	A
THM	GGGTTGGGTTGGGTTGGG	17.3 ± 0.9	55.7 ± 0.3	23 ± 2	P	P
VEGF	GGGAGGGTTGGGGTGGG	17.3 ± 0.8	57.1 ± 0.8	20.1 ± 0.9	P	P
G8	TGGGGGGGGT	17.8 ± 0.4	n/m	n/m	P	P
G5	TGGGGGT	17.3 ± 0.4	n/m	n/m	P	P

[a]Red-shift values were obtained with a >10-fold excess of DNA.
[b]$T_{1/2}$ values were calculated from CD melting experiments monitoring wavelengths at 264 or 295 nm. NMM:Tel22 ratio was 2:1.
[c]$\Delta T_{1/2}$ values were calculated by subtracting $T_{1/2}$ of DNA alone from the NMM-stabilized value.
[d]P, parallel; A, antiparallel; M, mixed (either mixture of two conformations or hybrid).
The first letter signifies the major conformation.
n/m, not measured.

peaks at 294 and 255 nm, and a trough at 235 nm (Supplementary Figure S2). This signal most likely corresponds to mixed (3 + 1) hybrid structure which concurs with the literature NMR evidence (14,16,18,59), but disagrees with conclusions from a recent CD study where an antiparallel structure was assigned this type of signal (60). It has been shown that in 5 mM K^+ buffer at 25°C >95% of Tel22 was folded into quadruplex (61). In 50Na buffer, Tel22 displays positive peaks at 296 and 246.5 nm and a negative peak at 264 nm (Supplementary Figure S2) indicating an antiparallel structure in agreement with literature (12). Finally, in 100Li buffer no significant secondary structure is detected.

To test the effect of NMM on the structure of Tel22, this porphyrin was added to Tel22 before or after the annealing step, followed by a prolonged incubation at 30°C. In 5K buffer, NMM causes a dramatic increase in the 264 nm signal of Tel22 at the expense of the signal at 295 nm (Figure 2A), suggesting a shift from hybrid to a conformation with substantial parallel component. This NMM-induced structural conversion is slow and requires 12 h when NMM is added during the annealing step and ~30 h at 30°C when NMM is added to Tel22 after annealing (data not shown). Previously, NMM has been shown to shift the equilibrium from an antiparallel to a parallel fold when annealed with the bimolecular telomere sequence dTA GGGUTAGGG in the presence of 100 mM KCl but not when added after annealing step (62). This could be due to inherent differences in the Tel22 and dTAG GGUTAGGGT sequences or to the fact that insufficient incubation time was allowed for the completion of rearrangement in the latter case. The isomerization process induced by NMM does not take place in 50Na or 100Li buffers as shown in Figure 2A indicating an essential role of K^+ ions. Similarly, it was shown previously that NMM does not affect the antiparallel GQ formed by $dG_4T_4G_4$ in 140 mM NaCl (62).

There are similar but smaller changes observed upon annealing of Tel22 in the presence of NMP in 5K buffer (Supplementary Figure S2B). On the other hand, the commonly used porphyrin TMPyP4 was not able to induce the formation of parallel structure of Tel22 under identical condition (Supplementary Figure S2C). Similar to NMM, NMP did not alter Tel22 CD signal in 50Na or in 100Li buffer, while TMPyP4 somewhat decreased the Tel22 CD signal in both cases (data not shown).

TDS were used successfully to identify the presence of GQs by their characteristic signature, such as positive peaks at 243 and 273 nm and a negative peak at 295 nm (54). TDS shown in Figure 2B indicate that Tel22 forms GQ structure in 5K and 50Na buffers. In 100Li, Tel22 displays TDS signal with a peak at 267 nm and a shoulder at 258 nm suggesting a mostly unfolded structure. The addition of NMM to Tel22 in 50Na and 100Li does not change the TDS signature of either, which is consistent with the CD annealing study data. The TDS signature of Tel22 in 5K, however, is altered by the addition of 2 and 4 equivalents of NMM. The troughs at 295 and 262 nm become more prominent, and the intensity of the peak at 273 nm decreases, but the diagnostic maxima and minimum are still present, pointing toward a quadruplex structure. The differences in the TDS signature of Tel22 with and without NMM could be attributed to the different topology of GQ structures (63), corroborating CD evidence.

The ability of NMM to induce structural rearrangement of quadruplex DNA in potassium buffer is more universal and not restricted to Tel22 sequence. Comparable transitions were observed for other related human telomeric sequences such as $dA_3G_3(TTAG_3)_3AA$, $dGG(ATTG_3)_4$, T_6Tel22, Tel22T_6 and $dAG_4(TTAG_4)_3$ (26TelG4),

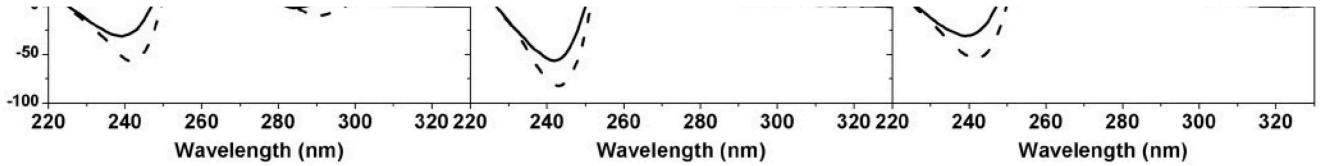


Figure 3. Annealing study of non-telomeric GQ-forming sequences G4TERT, Bcl-2 and cKit1 in the presence of NMM in 5K buffer. Samples of 4–6 μM DNA were annealed at 90°C for 10 min without (solid lines) or with (dotted lines) 2 equivalents of NMM, cooled to 30°C slowly and stored at 30°C for >12 h before collecting CD.

Supplementary Figure S3, as well as non-telomeric quadruplex DNA such as Bcl-2, cKit and G4TERT, Figure 3. These non-telomeric sequences have been shown to form parallel (or hybrid) GQs (64–66) and serve as oncogene promoters: Bcl-2 inhibits cell apoptosis mechanism in B-cell lymphoma cancer cell lines; cKit is a tyrosine kinase receptor that controls cell growth and TERT is the catalytic domain of telomerase. Quadruplexes already in a parallel conformation (THM, c-Myc, HIF-1α and VEGF) were largely unaffected by the addition of NMM (Supplementary Figure S4A), probably due to the presence of maximum possible parallel character in their structures. c-Myc oncogene is a transcriptional regulator of ~15% of human genes and is involved in a variety of cancers (67); VEGF is a promoter gene for vascular endothelial growth factor that is upregulated in a variety of cancers (68); HIF-1α gene encodes the alpha subunit of hypoxia-inducible factor1, which is a transcription factor that regulates response to hypoxia and is essential for tumor angiogenesis (69); and THM is a telomeric sequence from *Tetrahymena thermophila*. The thrombin binding aptamer, TBA, which forms a chair-type antiparallel GQ in 5K buffer (70), was unaffected by NMM as well (Supplementary Figure S4A), probably due to its antiparallel nature in agreement with the data for Tel22 in 50Na buffer.

It is worth noting that weak induced negative signals were observed upon NMM binding to all tested quadruplex structures but TBA in 5K buffer in the visible region of the CD spectra (Supplementary Figure S4C). No induced signal is seen for Tel22/NMM in 50Na. Presence of induced signal is an indication of interaction between NMM and GQ DNA. On the other hand, absence of induced CD signal does not necessarily signify lack of interaction as NMM was shown to interact with 10 different GQ structures but displayed no induced CD signal (49).

Potassium titration of Tel22:NMM 1:2 mixture

NMM-induced structural transition of Tel22 only occurs in the presence of K^+ ion. Therefore we proceeded by testing the effect of increased potassium concentration

on Tel22 structure in the presence of 2 equivalents of NMM (Supplementary Figure S5). As expected, a substantial increase in the CD signal at 264 nm was observed. Fitting the data with the Hill equation yielded a cooperativity coefficient, n, of 1.1 ± 0.1 and a binding constant of $(2.9 \pm 0.3) \times 10^2\,M^{-1}$ [forcing n to be exactly 1 yields a binding constant of $(3.2 \pm 0.2) \times 10^2\,M^{-1}$]. The 1:1 binding model is also confirmed by the fact that excellent fits were obtained with direct fitting [using Equations (1) and (2)], which yielded a similar binding constant of $(3.6 \pm 0.9) \times 10^2\,M^{-1}$. While the fits at 264 nm are excellent, the results should be used with care as the titration profile looks rather complex (Supplementary Figure S5A) and possibly reflects more than simple K^+ binding. In control experiments Tel22 was titrated with KCl in the absence of NMM. CD signal at 264 nm did not increase appreciably (Supplementary Figure S5C). For more details see Supplementary Data.

NMM binding to Tel22 and other GQs studied by UV-vis spectroscopy

Detailed information about NMM binding to quadruplex structure of Tel22 and other sequences was obtained through UV-vis absorbance titrations by monitoring NMM's Soret band at 379 nm. In 5K buffer NMM exhibited a large bathochromic shift of 19.3 ± 0.4 nm and small hypochromicity of 4% upon binding to Tel22 (Figure 4A). While red shift was independent of incubation time, the hypochromicity changed into mild hyperchromicity ($H = -3\%$) when samples were allowed to equilibrate for >24 h at 30°C. The Tel22–NMM complex has an extinction coefficient of $(1.48 \pm 0.10) \times 10^5\,M^{-1}\,cm^{-1}$ at 398.5 nm. When other GQ-forming sequences were titrated into NMM under equilibrium conditions, red shifts of 17–20 nm (Table 1) and small hypochromicities or hyperchromicities of 20 to −21% were observed (Supplementary Table S2 and Supplementary Figure S4B). The large variation in Soret band intensity could possibly indicate differences in binding modes depending on DNA sequence and quadruplex geometry. The Bolton laboratory reported ~20 nm red shift of NMM's Soret band upon addition of 50-fold excess of TBA and $d(G_4T_4G_4)$ accompanied by relatively high hyperchromicity in the latter case (46), results consistent with our data. While red shift and change in Soret band intensity can provide information about the ligand's binding mode when long duplex DNA is concerned (71), no reliable numbers exist for porphyrin binding to quadruplex structures.

The method of continuous variation analysis was used to accurately determine the binding stoichiometry of the Tel22–NMM interaction (Supplementary Figure S6). The plot of absorbance difference versus mole fraction of NMM peaked at 0.50, indicating 1:1 binding. SVD confirmed that the 1:1 binding was a two-state process (data not shown). Direct fit of the titration data yielded a relatively weak binding constant of $(1.0 \pm 0.3) \times 10^5\,M^{-1}$ under equilibrium conditions (Figure 4C). Modest binding constants were also reported for NMM interactions with other quadruplex structures (44–46). Binding parameters of NMM to Tel22 were determined for short incubation times (assuming possible NMM binding to hybrid GQ structure of Tel22) and long incubation times (reflecting equilibrium condition and binding of NMM to presumably parallel GQ structure of Tel22). In both cases the data were rather similar: the same red shift, binding stoichiometry and similar binding constants [$(1.0 \pm 0.3) \times 10^5\,M^{-1}$ and $(3 \pm 2) \times 10^5\,M^{-1}$] and hypochromicities. Therefore NMM binding to hybrid and parallel GQ is rather similar with possible differences in the orientation of the porphyrin core with respect to the DNA axis, as evidenced by the slight difference in hypochromicities.

UV-vis titration experiments were repeated in 50Na (Figure 4B) and 100Li (data not shown) buffers. These titrations did not lead to any red shift and resulted in modest hypochromicity of 19 and 20% for 50Na and 100Li conditions, respectively, at [GQ]/[NMM] ratio of ≥10. Absence of red shift could be interpreted as the inability of antiparallel conformation of Tel22 that exists in 50Na or single stranded DNA that exists in 100Li buffer to interact with NMM. Note, the observed change in hypochromicity might suggest some interactions between NMM and Tel22 that does not involve any degree of porphyrin stacking. This is an important and unexpected

result that demonstrates NMM's selectivity not only for quadruplex structures over other DNA structures (in this case single stranded DNA), but also for parallel GQ topology.

To further test whether binding of NMM to quadruplexes is governed by the choice of cation (Na^+ versus K^+) or by the nature of GQ present, we employed $[dTG_5T]_4$ (G5) which forms a parallel-stranded GQ regardless of buffer conditions. Upon addition of G5 to NMM in 5K or in 50Na buffers similar red shifts of ~18 nm were observed, suggesting the same binding modes and underlying the importance of GQ structure rather than choice of cation.

NMM binding to and isomerization of Tel22 studied by gel electrophoresis and CD

Since CD spectroscopy could only provide indirect evidence of GQ structure but not number or size of DNA species present in solution, the details of NMM binding to Tel22 was further investigated using native PAGE gel electrophoresis. Tel22 samples at 40 μM were annealed with increasing amount of NMM (0.25–12 equivalents) and loaded on the gel as well as scanned on CD in order to get a correlation between the observed gel species and the CD spectra. Tel22 in 5K buffer appears as a single high mobility band on the gel, Figure 5A, suggesting the existence of a single conformation. Addition of NMM to Tel22 with the subsequent equilibration leads to appearance of a new slower migrating band on the PAGE gel and an increase in CD signal at 264 nm in expense of the signal at 295 nm (Figure 5). On the gel the new species seems to be structurally homogeneous. The CD spectrum of the resulting GQ is consistent with a parallel propeller-loop structure. Gel and CD data indicate that Tel22 undergoes structural rearrangement in the presence of NMM. The rearrangement is slow as the addition of NMM to Tel22 right before loading the gel (last three lanes on Figure 5A) does not lead to the formation of Tel22–NMM complex. The CD spectra of these samples resemble that of Tel22 alone (Supplementary Figure S7C).

In order to determine the distribution of NMM among observed DNA species, PAGE gels were also visualized using NMM's fluorescence (Supplementary Figure S7). NMM fluoresces weakly in the free state but its fluorescence increases in the presence of a quadruplex (46). Our results suggest that NMM acts on Tel22 in a stoichiometric fashion (single turn-over event) and remains bound to the final isomerized Tel22. On the other hand, it is not bound to the starting hybrid Tel22 structure. We cannot exclude the possibility that NMM is bound to the original Tel22 structure, but either dissociates from it during electrophoresis; or its florescence is completely quenched by the DNA.

The slow conversion of Tel22 structure in the presence of NMM as well as slower migration of resulting complex on PAGE gel could suggest that the structural transition reflects formation of bimolecular quadruplex. To test this possibility and determine unambiguously the molecularity of NMM-induced Tel22 structure we performed strand mixing experiments. T_6Tel22 was chosen to have larger size but similar secondary structure to that of Tel22 and interact with NMM in a similar manner, Supplementary Figure S8. Tel22 was mixed with T_6Tel22, annealed with NMM and run on the 20% native PAGE gel. Two bands are expected for monomolecular complex and three [$(Tel22)_2$–NMM; (Tel22 × T_6Tel22)–NMM and $(T_6Tel22)_2$–NMM] for bimolecular. Two bands are seen on the gel, Figure 5B, clearly indicating monomolecular nature of quadruplex core in Tel22–NMM complex. The Trent group observed similar conversion of Tel22 structure induced by 50% of acetonitrile (72) and suggested based on NMR line-width and relaxation rate constants that the newly formed GQ is most likely monomolecular supporting our conclusion.

CD data from gel samples, Figure 5C, were analyzed in order to determine binding parameters and compare them to those obtained from UV-vis titrations (see above). While in UV-vis the focus is on NMM, in CD it is on DNA. SVD analysis was employed first to determine the number of significant spectral species present in the NMM–Tel22 mixture (Supplementary Figure S9). The analysis indicates that NMM binding to Tel22 under equilibrium is a simple two-state process reflecting an NMM-induced gain in Tel22's parallel character at the expense of antiparallel structure. Direct fitting of the data with 1:1 binding model, Figure 5D, yielded binding constant of $(0.7 \pm 0.1) \times 10^5\,M^{-1}$ in good agreement with results from UV-vis titrations.

Gel experiments were repeated in Na^+ and Li^+ containing buffers (Supplementary Figure S10). In 50Na Tel22 alone ran as a single band with a slightly slower migration than in 5K buffer, reflecting the difference in quadruplex conformation. In 100Li, Tel22 alone was heterogeneous and ran as major and minor bands with slow mobility suggesting structures with large hydrodynamic radius, possibly single-stranded conformation, in agreement with CD and TDS results. Addition of NMM to Tel22 in 50Na or 100Li buffers caused no change in the mobility of Tel22 or NMM's fluorescence, confirming the UV-vis results that NMM does not interact with the antiparallel (in 50Na) or single-stranded (in 100Li) conformation of Tel22, and underscoring the importance of K^+ ions.

Stabilizing ability and selectivity of NMM and its analogues determined by FRET melting assay

FRET melting assay is the benchmark technique used to test stabilizing ability and selectivity of quadruplex ligands. We employed FRET with labeled human (F21D) and yeast (FGD) telomeric DNA to test NMM and NMP in comparison to non-methylated analogues, MIX and PIX. To confirm that under experimental conditions F21D and FGD form GQ structures, CD and TDS spectra were collected (Supplementary Figure S11). Both structures displayed a quadruplex-specific TDS signature (54). The CD spectrum of F21D closely resembles that of

Figure 5. Isomerization of Tel22 in the presence of NMM. (**A**) Gel electrophoresis experiments. Non-denaturing gel was prepared at 20% acrylamide supplemented with 5 mM KCl. Tel22 strand concentration was 40 μM. Tel22 samples were annealed with NMM (0.25–12 equivalents), cooled to 30°C and incubated at 30°C for >12 h, except for the last three lanes (before size marker lanes) where NMM was added to Tel22 right before loading the gels. (**B**) Determination of the molecularity of Tel22–NMM complex. Samples of Tel22 were mixed with T_6Tel22 in 1:1 ratio and annealed with 2 and 8 equivalents of NMM. Tel22 and T_6Tel22 strand concentration was 40 μM. Gels were visualized using UV-shadowing. Strand mixing experiments were also repeated with Tel22 and Tel22T_6 yielding similar results (data not shown). (**C**) CD wavelength scans for the Tel22–NMM samples from the gel in (A) at 25°C. Samples were diluted to 1.9 μM final Tel22 concentration with 5K buffer. Isoelliptic points at 252 and 281 nm are marked on the graph. (**D**) Molar ellipticity at 264 nm and fit to 1:1 binding model that yielded K_a of $(0.7 \pm 0.1) \times 10^5\,M^{-1}$.

Tel22 in 5K buffer. FGD has significantly more parallel character than F21D, displaying a major peak at 265 nm. Collectively, TDS and CD results indicate that the fluorescent probes do not interfere with the formation GQ structures of F21D and FGD.

The stability of GQ in the presence of porphyrins was characterized using $\Delta T_{1/2}$ parameter that reflects a change in the melting temperature as compared to GQ alone. First, concentration dependence of $\Delta T_{1/2}$ was determined for NMM, NMP and their non-methylated analogues,

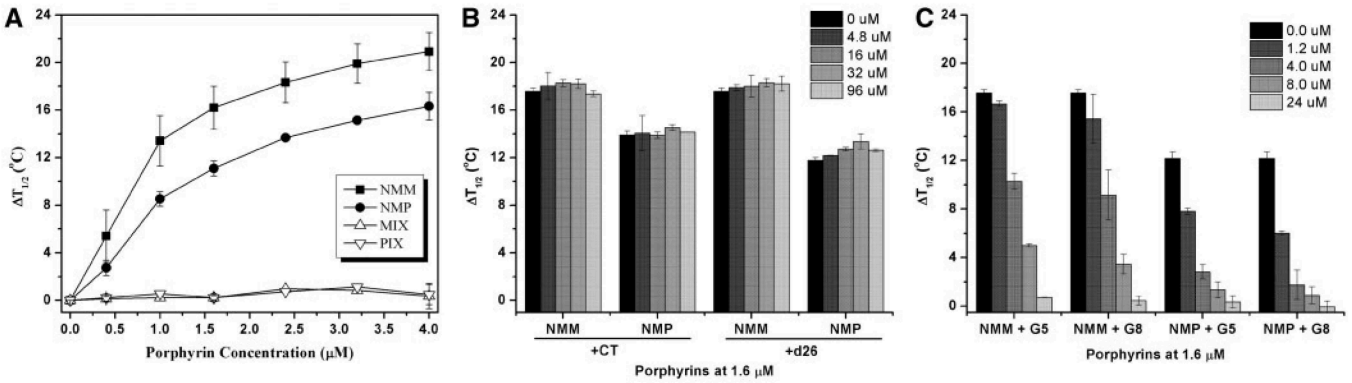


Figure 6. FRET melting of F21D in the presence of porphyrins and competitors. (**A**) Increase in the stabilization temperature, $\Delta T_{1/2}$, of 0.2 μM F21D as a function of porphyrin concentration. (**B**) Stabilization of F21D by NMM or NMP in the presence of duplex DNA competitors, ds26 or CT or (**C**) parallel stranded GQ competitors $[dTG_5T]_4$, G5, and $[dTG_8T]_4$, G8. Concentrations of porphyrins and F21D were fixed at 1.6 and 0.2 μM, respectively. Amount of competitor is indicated in the figure legend. All experiments were performed in 5K buffer.

Table 2. Stabilization temperature, $\Delta T_{1/2}$ (°C), for porphyrins under study in FRET melting assays

DNA/porphyrin	NMM	NMP	MIX	PIX	TMPyP4	ZnTMPyP4
F21D, 5K	14 ± 2	8.5 ± 0.6	0.2 ± 0.3	0.2 ± 0.3	34 ± 2	23 ± 1
FGD, 5K	12.5 ± 0.3	9.8 ± 0.3	0.8 ± 0.3	0.2 ± 0.3	31 ± 2	22.1 ± 0.9
F21D, 50Na	0.4 ± 0.5	0.3 ± 0.3	–	–	30.6 ± 0.3	16.4 ± 0.7

The melting temperatures were recorded at 0.2 μM DNA and 1 μM porphyrins.

MIX and PIX in 5K buffer (Figure 6A for F21D and Supplementary Figure S12A for FGD). The plots show dose-dependent behavior with stabilization temperatures of 14 ± 2°C for F21D + NMM, 12.5 ± 0.3°C for FGD + NMM, 8.5 ± 0.6°C for F21D + NMP and 9.8 ± 0.3°C for FGD + NMP at 1 μM ligand concentration in 5K buffer (Table 2). Under saturating porphyrins concentrations the values of $\Delta T_{1/2}$ for F21D were yet higher, >20°C for NMM and >16°C for NMP. This excellent stabilization of human telomeric DNA by NMM (and NMP) could be due to porphyrins' binding to DNA or could result from the NMM-induced structural conversion of F21D to a more stable conformation. Overall, NMM displayed stabilizing abilities superior to those of NMP and both porphyrins stabilized human telomeric DNA, F21D, to a greater extent than yeast telomeric DNA, FGD (except at concentrations <0.5 μM). MIX and PIX, on the other hand, did not stabilize either of the GQs. This finding underlines the importance of the core *N*-methyl group for GQ stabilization and suggests its role in porphyrins interactions with GQ DNA.

F21D in 50Na buffer forms an antiparallel quadruplex according to CD and TDS (Supplementary Figure S11). In FRET assays both NMM and NMP were unable to stabilize this type of GQ structure (Supplementary Figure S13). These results clearly indicate NMM's and NMP's preference for parallel over antiparallel GQ DNA in agreement with UV-vis and CD results described earlier.

FRET melting assay allows facile determination of ligand's selectivity. The effect of two dsDNA models, CT and ds26, on $T_{1/2}$ of F21D and FGD in the presence of fixed amount of NMM or NMP (1.6 μM) is shown in Figure 6 for F21D and Supplementary Figure S12 for FGD. The data demonstrate exceptional selectivity of both porphyrins as the $T_{1/2}$ of F21D and FGD did not decrease upon addition of up to 480 equivalents of competitors (as compared to the labeled DNA) setting the selectivity ratio at least at 480. On the other hand, the stabilizing ability of widely used TMPyP4 toward F21D decreased substantially in the presence of competitors (Supplementary Figure S14), in agreement with literature (40). These results corroborate earlier UV-vis data for NMM and TMPyP4 binding to CT DNA, where no change in the position and intensity of NMM's Soret band was detected upon addition of 50-fold excess of CT DNA, while TMPyP4 displayed substantial hypochromicity and red shift under identical conditions (46).

FRET assays represent a convenient technique to test porphyrins' binding modes. Specifically to determine if NMM could stabilize F21D by interacting with loops or through intercalation, the non-loop containing GQ of different length $[dTG_5T]_4$ (G5) and $[dTG_8T]_4$ (G8) were used as competitors in FRET. These quadruplexes form parallel GQ structure in both Na^+ and K^+ (73,74). The $T_{1/2}$ of F21D stabilized by NMM decreased substantially in the presence of large amounts of G5 or G8 as shown in Figure 6C. Similar results were obtained for FGD, Supplementary Figure S12C. The data indicate that both G5 and G8 compete effectively for NMM binding underlying once again exceptional selectivity of NMM toward parallel GQ. Since both G5 and G8 at equal

concentration decrease the $T_{1/2}$ of F21D by virtually the same amount, NMM's intercalation could be excluded, as G5 and G8 have different numbers of intercalation sites. Loop binding could also be excluded as neither G5 nor G8 have loops. However this statement is based on the assumption that NMM's binding mode is the same for Tel22 and G5/G8 quadruplexes, which might not be true. Selectivity ratios could be determined from FRET competition assays (for details see Supplementary Data) and based on values obtained, NMM's (and NMP's) preference for specific GQ follows the order FGD (5K) > F21D (5K) > G8 (5K) ≈ G5 (5K) >> F21D (50Na). This higher selectivity of NMM toward FGD and F21D as compared to G5 and G8 might suggest different modes of binding. More importantly, NMM displays clear discrimination against antiparallel GQ structure of F21D in 50Na.

Our FRET data taken together indicate that, against common belief, NMM and NMP have rather good GQ stabilizing properties, >480-fold selectivity for GQ over dsDNA, and remarkable selectivity for parallel versus antiparallel GQ structures of human telomeric DNA. The latter finding is extremely important as up to date only TOxaPy is reported to clearly favor a specific GQ structure (42) making NMM a second ligand in this class, but with opposite binding preference.

Stabilizing ability of NMM determined via CD melting

We next conducted CD melting of Tel22 alone and in the presence of NMM in order to ensure that stabilization of F21D by NMM observed in FRET was due to NMM's interaction with quadruplex and not with the fluorophores. Data were monitored at 264 and 295 nm, CD maxima for parallel and antiparallel GQ, respectively. In the absence of NMM $T_{1/2}$ for Tel22 is 48.0 ± 0.8°C as determined by CD at 295 nm, while that for F21D is 44.9 ± 0.6°C as determined from FRET. Slight destabilization of a quadruplex due to fluorophore-labeling has been observed previously (75). In the presence of NMM Tel22 adopts predominantly parallel conformation whose melting temperature (determined from CD signal at 264 nm) displays dose-dependent behavior (Figure 7, inset) observed earlier in FRET (Figure 6A). At the same time, $T_{1/2}$ determined from CD data at 295 nm showed no dependence on NMM concentration. In the presence of 20 equivalents of NMM, Tel22 is stabilized by 13.9°C. It is worth noting, however, that this stabilization temperature is calculated relative to $T_{1/2}$ of hybrid conformation of Tel22 (monitored at 295 nm) as in the absence of NMM Tel22 does not adopt parallel GQ structure and CD signal at 264 nm does not display temperature dependence.

In order to extract thermodynamic parameters from melting profiles, the reversibility of Tel22 folding/unfolding process was tested by recording melting and cooling curves which displayed small hysteresis of 5.2 ± 0.5°C (at 0.27°C min^{-1} temperature change rate). The hysteresis was independent of NMM concentration indicating that NMM most likely does not participate in the rate-determining step of Tel22 folding. The hysteresis indicates slow kinetics of Tel22 folding in the presence of NMM, also seen in annealing studies. Because hysteresis was rather small, the thermodynamic parameters were evaluated: $\Delta H = 261 \pm 17$ kJ mol^{-1} and $\Delta G(298\,\mathrm{K}) = -27 \pm 1$ kJ mol^{-1} (for heating curves). These data, while only estimates due to hysteresis, are in excellent agreement with thermodynamic parameters for unfolding of human telomeric sequences as well as other intramolecular GQs in the absence of ligands (76,77).

NMM's stabilizing abilities were also determined with respect to other quadruplex forming sequences and the results are summarized in Table 1. As in the case of Tel22 small hysteresis of 2–6°C is seen in these experiments. Two equivalents of NMM stabilized different quadruplex structures to a different degree. Specifically, structures with substantial antiparallel component were minimally stabilized:

by 0 ± 2°C for Tel22 in 50Na, by 1 ± 2°C for TBA in 5K and by 0 ± 1°C for 26TelG4 in 50Na. Structures with both parallel and antiparallel components, 26TelG4 and Tel22 in 5K buffer were moderately stabilized by ~7°C. Quadruplexes that contain predominantly parallel character (with the exception of HIF-1α) were stabilized rather effectively. VEGF, THM, G4TERT and Bcl-2 displayed the highest stabilization temperatures of ≥18°C, Table 1. Interestingly these structures also caused the largest hyperchromicity of NMM's Soret band in UV-vis titrations (Supplementary Table S2). These results yet again highlight NMM's good stabilizing ability and its preference for parallel over antiparallel GQ DNA.

Surprisingly, NMM displays no apparent stabilization of HIF-1α structure ($\Delta T_{1/2} = 1 \pm 1$°C) in spite of the parallel nature of this quadruplex and its ability to bind NMM as judged by a red-shift of 18.5 nm in UV-vis. HIF-1α has six GGG stretches and can form more than one stable GQ under experimental conditions. The biphasic nature of the HIF-1α melting curve (Supplementary Figure S15) indicates that at least two quadruplexes are present (a thermodynamic and a kinetic GQ). The thermodynamic GQ is more prominent under equilibrium conditions. Upon cooling a monophasic curve is observed which may be attributed to the formation of a single kinetic GQ. The melting process is accompanied by a substantial hysteresis of 17 ± 2°C. Addition of NMM leads to monophasic melting curve and small hysteresis of 1.8 ± 0.4°C, indicating that NMM favors a single (thermodynamic) GQ structure. While NMM is not required for the formation of thermodynamic product, it certainly speeds up its formation.

DISCUSSION

The main goals of this work were to provide detailed characterization of NMM's binding to human telomeric DNA, to determine its general GQ stabilizing ability, and to clarify which structural elements confer its unusual selectivity toward GQ DNA. NMM-induced structural rearrangement of Tel22 and NMM's ability to selectively interact with Tel22 in K^+ buffer but not in Na^+ buffer were discovered during this work, characterized and presented here as well.

NMM-induced structural rearrangement of Tel22

NMM, and to a lesser extent NMP, displayed the ability to convert Tel22 in 5K buffer to a new structure whose CD signature resembles that of a parallel GQ DNA. This transition requires potassium ions as no structural changes were observed in 100Li or 50Na buffers or in water. NMM was also able to enhance the parallel conformation of several non-telomeric GQs, including G4TERT, Bcl-2 and cKit1, highlighting the breadth of this ability. Similarly, it was reported earlier that NMM created 'its own binding site' on DNA aptamer possibly by rearranging the original DNA structure as the methylation protection pattern differed substantially in the presence and absence of NMM and was K^+ dependent in the former case (44). As NMP caused structural rearrangement of Tel22 to a much lesser extent than NMM, the peripheral substituents on the porphyrin core might be important. There are steric and electronic differences between the ethyl groups of NMM and vinyl groups of NMP. In addition, NMP is less soluble in water than NMM.

A number of other ligands have been reported to induce structural rearrangement of human telomeric DNA (78). The majority of these ligands induced transitions to antiparallel or mixed hybrid structures and only four to a parallel conformation. The list includes a conjugate of anthracene and polyamine (79), octacationic quaternary ammonium Zn(II) phtalocyanine (80), Ni(II) salphen complex with cyclic amine head groups (33) and ZnTMPyP4 (35). In all respects, NMM acts as a very potent driver of Tel22 structural rearrangement, equal or superior to the other four molecules reported in the literature.

Details of NMM binding to and stabilization of Tel22

The NMM–Tel22 complex consists of one DNA strand and one NMM molecule binding between which is characterized by modest K_a of $\sim 1.0 \times 10^5\,M^{-1}$. The value of binding constant is in a good agreement with literature data: equilibrium dialysis experiments with a variety of GQ structures yielded NMM's K_a of 10^4–$10^5\,M^{-1}$ (40). NMM binding to G-rich DNA aptamer was an order of magnitude tighter, $1–2 \times 10^6\,M^{-1}$ (44). Finally, NMM binding constants estimated based on the NMM's inhibition of GQ unwinding by Sgs1 and BLM DNA helicases was $10^6\,M^{-1}$ (47). Difference in GQ structure or technique of K_a determination might be responsible for this modest variation in the value of binding constants.

In spite of relatively weak binding, NMM and NMP showed excellent stabilizing ability toward fluorescently labeled human telomeric DNA in 5K buffer, >20°C for NMM and >16°C for NMP under saturating conditions. For NMM, good stabilizing ability was also confirmed in CD melting studies of Tel22 in 5K buffer with $\Delta T_{1/2}$ of 13.9°C at [NMM]/[Tel22] = 20. The lack of GQ stabilization by both MIX and PIX indicate that the *N*-methyl group of the porphyrins is integral for their association with GQ DNA. This methyl group can alter the interactions between Tel22 and NMM in at least two ways. It can interact directly with the quadruplex or it could lead to the distortion of NMM's macrocyle (81), which, in turn, can confer NMM's remarkable selectivity.

When compared to other porphyrins, NMM's and NMP's GQ stabilizing abilities are somewhat weaker (Table 2). This trend was expected as both NMM and NMP are negatively charged at physiological pH and could not engage in electrostatic interactions with the DNA backbone. However, with the added advantage of exceptional selectivity [(45,46) and here], the stabilizing abilities of NMM and NMP are rather good. In fact, at the highest concentration of the duplex competitor ds26 used (480-fold excess), the stabilization of F21D caused by TMPyP4 ($\Delta T_{1/2} = 10.0$°C) is lower than that caused by both NMM ($\Delta T_{1/2} = 18.2$°C) and NMP ($\Delta T_{1/2} = 15.9$°C). The same is true in the case of FGD. This means that

under physiological conditions NMM and NMP will perform far better than TMPyP4.

The unique structural specificity of NMM–GQ interaction raises questions about NMM's binding mode. No structural data is available for NMM binding to GQ and the information about the binding modes is derived solely based on the spectroscopic data. UV-vis data and FRET competition experiments with dsDNA and parallel quadruplexes of different length ($[dTG_5T]_4$ and $[dTG_8T]_4$) indicate that, as expected but rarely tested, true intercalation is not likely. Groove binding could be excluded as well, because parallel stranded propeller GQ, preferred by NMM, has three of its grooves blocked by loops.

Other possible binding modes that require consideration are end-stacking and loop binding. End-stacking was observed in the NMR structure of TMPyP4-bound *c-myc* promoter (82) but loop binding was reported for the only crystal structure of TMPyP4-bound human telomeric DNA (83). As for NMM, the end-stacking binding mode is compatible with the fact that parallel GQ of Tel22 has easily accessible G-tetrads on both ends of its core (13), whereas the end G-tetrads of antiparallel Tel22 are crowded with loops (12). In addition, NMM binds to an antiparallel structure of TBA (although NMM does not stabilize it) possibly because TBA forms chair type antiparallel GQ whose end G-tetrads are not obstructed by loops (70). In view of NMM's remarkable selectivity, it is very possible that the interaction is more specific than simple end-stacking of conjugated systems. Parallel and antiparallel structures of Tel22 have no loop type overlap, therefore NMM's selectivity could be explained by combination of end-stacking and specific loop recognition. Loops could be considered promising targets for development of GQ-structure specific ligands. At the same time, FRET experiments suggest that NMM is well capable of binding to parallel stranded quadruplexes $[dTG_5T]_4$ and $[dTG_8T]_4$ that have no loops, therefore loop binding, while possible, might not be essential for NMM interactions with GQ DNA. Based on spectroscopic data alone, binding of NMM to quadruplex structures probably involves both end-stacking and loop-binding to a different extent for different quadruplexes. Structural data therefore are highly desirable as they can provide detailed molecular view of NMM's binding to quadruplex DNA.

Selectivity of NMM toward GQ versus dsDNA and toward different GQ structures

Greater than 10-fold selectivity of NMM toward variety of GQ structures versus a total of 46 single, double and triple stranded DNA, RNA and DNA–RNA hybrids has been demonstrated by the Chaires and Bolton groups (40,45,46). Biological studies of helicase inhibition by NMM suggested selectivity for GQ DNA versus Holliday junctions or partial duplex substrate to be >30-fold (47). Our work supports these findings and demonstrates exceptional selectivity of NMM and NMP (~480-fold) for quadruplex versus dsDNA in FRET assays. It is important to note, that a variety GQ ligands have good to excellent selectivity for GQ DNA versus dsDNA (36,78). This selectivity, while absolutely necessary, is probably not sufficient for a promising drug candidate whose mode of action involves GQ stabilization. Selectivity of a ligand for a specific GQ target must be considered. This is a challenging task due to heterogeneity of quadruplex structures on one hand and to the presence of the same tetrad structure in each quadruplex on the other hand.

NMM displays unexpected preference for a parallel GQ structure of Tel22 in K^+ buffer and no interaction with antiparallel Tel22 in Na^+ buffer. Three experimental results support this conclusion. First, no red shift was observed in UV-vis titrations of NMM with Tel22 annealed in 50Na buffer. Second, NMM failed to stabilize F21D in FRET melting assays performed under the same buffer condition. Lastly, NMM was not able to alter the antiparallel structure of Tel22 in 50Na (shown by CD and gel electrophoresis) because binding of NMM to the original structure is necessary for the rearrangement. Similarly, no binding of NMM to human telomeric DNA in Na^+-rich buffer was detected in competition dialysis experiments by Chaires group (40). In addition, NMM failed to alter or stabilize antiparallel structure of TBA in 5K buffer. However, red shift of 19.5 nm was observed in UV-vis titrations suggesting possible interaction. Antiparallel structure of TBA lacks diagonal loop (all loops are lateral) (70) and has an open G-tetrad available for NMM binding via end-stacking.

Contrary to our observation of NMM's selectivity for a specific GQ conformation, equilibrium dialysis experiments in a Na^+-rich buffer using 14 quadruplex and i-motif structures demonstrated only 2-fold selectivity. In these experiments NMM showed the highest preference for the parallel G-wire formed by 5′-$dG_{10}T_4G_{10}$ and, surprisingly, for the antiparallel quadruplex 5′-$d(G_4T_4)_3$ (40). To explain the latter case we conducted preliminary annealing experiments and observed NMM-induced structural conversion of 5′-$d(G_4T_4)_3$ in 50Na buffer to quadruplex with substantially increased parallel component (data not shown). The time used for equilibrium dialysis experiment is sufficient for such a transition to take place.

The first reported quadruplex ligand capable of clear discrimination between parallel and antiparallel quadruplex of human telomeric repeat is TOxaPy (42). Interestingly, with respect to Tel22, TOxaPy behaves in a manner opposite to NMM. It stabilizes F21T (where TAMRA has been used instead of Dabcyl) in Na^+-rich buffer ($\Delta T_m = 10.8°C$), but not in K^+ rich buffer ($\Delta T_m < 1°C$, at 1 μM ligand) (42). However, both ligands bind the parallel structure of $[d(TG_5T)]_4$ regardless of buffer (K^+ or Na^+). When other properties are compared, TOxaPy displays more efficient binding to quadruplex DNA ($K_a = 5 \times 10^6$ M versus $K_a = \sim 1 \times 10^5$ M for NMM), but lower stabilizing ability and selectivity than NMM. This comparison is only qualitative as experiment type and exact conditions were similar but not identical.

NMM thus represents a second complementary example of GQ stabilizer with clear preference for a specific GQ structure or structure type. It is, therefore, a

promising platform for development of efficient cancer therapeutics or probes for detection and isolation of specific GQ structures. NMM's structural elements present convenient points for synthetic modifications. Side chains and core nitrogens could be chemically modified with the hope of preparing compounds that will retain NMM's exceptional selectivity, but will have improved quadruplex binding and stabilization properties.

CONCLUSION

The all parallel propeller-loop structure of the human telomeric repeat, which has previously been shown to form only in high potassium buffer, under crowding conditions, or in the crystalline state, becomes kinetically and thermodynamically favorable in dilute, low potassium environments when Tel22 is complexed with NMM. At the same time NMM is completely insensitive to Na^+ form of Tel22 and displays low binding to TBA and 26TelG4, both of which have substantial antiparallel component. NMM therefore is a second example, after TOxaPy, of GQ-ligand that has clear preference for a unique GQ structure, a property necessary for being a successful drug candidate. These two compounds have opposite preference: NMM favors parallel structures, whereas TOxaPy favors antiparallel GQs. NMM displays an excellent stabilizing ability toward human, yeast and *T. thermophila* telomeric sequences and a variety of oncogene promoters in dilute K^+ buffer. All of these sequences contain substantial parallel component either present already or induced by NMM. While the stabilizing ability of NMM is weaker than that of TMPyP4, its selectivity is far superior making it a better GQ-stabilizer under the physiological excess of duplex DNA. The question remains about the structural details of the NMM–Tel22 complex and the kinetics and mechanism of NMM-induced structural rearrangement of Tel22 in K^+ buffer. Our group is currently pursuing this work. Results, taken together, indicate that NMM presents an interesting platform toward developing a promising cancer drug that acts via structural transformation and subsequent stabilization of specific GQ targets.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1 and 2, Supplementary Figures 1–15, Supplementary Methods and Supplementary References [84–87].

ACKNOWLEDGEMENTS

The authors thank Dr. Johnson from University of Pennsylvania for inspiring us to work with NMM, Dr. Chaires from University of Louisville for his help with data fitting for K_a determination and Dr. Pasternack from Swarthmore College for helpful discussions.

FUNDING

To L.A.Y.: Camille and Henry Dreyfus Faculty Start-up Award; Cottrell College Science Award from Research Corporation (grant number 7843); Swarthmore funds. To J.L.M.: Association pour la Recherche sur le Cancer (programme libre ARC); Région Aquitaine; Fondation pour la Recherche Médicale (FRM); INCa and ANR grants (F-DNA and G4-TOOLBOX). Funding for open access charge: Swarthmore College and INSERM.

Conflict of interest statement. None declared.

REFERENCES

1. Siddiqui-Jain,A., Grand,C.L., Bearss,D.J. and Hurley,L.H. (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl Acad. Sci. USA*, **99**, 11593–11598.
2. Rankin,S., Reszka,A.P., Huppert,J., Zloh,M., Parkinson,G.N., Todd,A.K., Ladame,S., Balasubramanian,S. and Neidle,S. (2005) Putative DNA quadruplex formation within the human *c-kit* oncogene. *J. Am. Chem. Soc.*, **127**, 10584–10589.
3. Sen,D. and Gilbert,W. (1988) Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature*, **334**, 364–366.
4. Huppert,J.L. and Balasubramanian,S. (2005) Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.*, **33**, 2908–2916.
5. Huppert,J.L. and Balasubramanian,S. (2007) G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.*, **35**, 406–413.
6. Todd,A.K., Johnston,M. and Neidle,S. (2005) Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res.*, **33**, 2901–2907.
7. Ribeyre,C., Lopes,J., Boulé,J.-B., Piazza,A., Guédin,A., Zakian,V.A., Mergny,J.-L. and Nicolas,A. (2009) The yeast Pif1 helicase prevents genomic instability caused by G-quadruplex-forming CEB1 sequences *in vivo*. *PLoS Genet*, **5**, e1000475.
8. Paeschke,K., Capra,J.A. and Zakian,V.A. (2011) DNA replication through G-quadruplex motifs Is promoted by the saccharomyces cerevisiae Pif1 DNA helicase. *Cell*, **145**, 678–691.
9. Lopes,J., Piazza,A., Bermejo,R., Kriegsman,B., Colosio,A., Teulade-Fichou,M.-P., Foiani,M. and Nicolas,A. (2011) G-quadruplex-induced instability during leading-strand replication. *EMBO J*, **30**, 4033–4046.
10. Maizels,N. (2006) Dynamic roles for G4 DNA in the biology of eukaryotic cells. *Nat. Struct. Mol. Biol.*, **13**, 1055–1059.
11. Lipps,H.J. and Rhodes,D. (2009) G-quadruplex structures: in vivo evidence and function. *Trends Cell Biol.*, **19**, 414–422.
12. Wang,Y. and Patel,D.J. (1993) Solution structure of the human telomeric repeat d[$AG_3(T_2AG_3)_3$] G-tetraplex. *Structure*, **1**, 263–282.
13. Parkinson,G.N., Lee,M.P.H. and Neidle,S. (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*, **417**, 876–880.
14. Luu,K.N., Phan,A., Kuryavyi,V., Lacroix,L. and Patel,D.J. (2006) Structure of the human telomere in K+ solution: an intramolecular (3+1) G-quadruplex scaffold. *J. Am. Chem. Soc.*, **128**, 9963–9970.
15. Phan,A.T., Luu,K.N. and Patel,D.J. (2006) Different loop arrangements of intramolecular human telomeric (3+1) G-quadruplexes in K+ solution. *Nucleic Acids Res.*, **34**, 5715–5719.
16. Ambrus,A., Chen,D., Dai,J., Bialis,T., Jones,R.A. and Yang,D. (2006) Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. *Nucleic Acids Res.*, **34**, 2723–2735.

17. Zhang,Z., Dai,J., Veliath,E., Jones,R.A. and Yang,D. (2010) Structure of a two-G-tetrad intramolecular G-quadruplex formed by a variant human telomeric sequence in K+ solution: insights into the interconversion of human telomeric G-quadruplex structures. *Nucleic Acids Res.*, **38**, 1009–1021.
18. Xu,Y., Noguchi,Y. and Sugiyama,H. (2006) The new models of the human telomere d[AGGG(TTAGGG)3] in K+ solution. *Bioorg. Med. Chem.*, **14**, 5584–5591.
19. Phan,A.T. (2010) Human telomeric G-quadruplex: structures of DNA and RNA sequences. *FEBS J.*, **277**, 1107–1117.
20. Heddi,B. and Phan,A.T. (2011) Structure of human telomeric DNA in crowded solution. *J. Am. Chem. Soc.*, **133**, 9824–9833.
21. Li,J., Correia,J.J., Wang,L., Trent,J.O. and Chaires,J.B. (2005) Not so crystal clear: the structure of the human telomere G-quadruplex in solution differs from that present in a crystal. *Nucleic Acids Res.*, **33**, 4649–4659.
22. Dai,J., Carver,M. and Yang,D. (2008) Polymorphism of human telomeric quadruplex structures. *Biochimie*, **90**, 1172–1183.
23. Miyoshi,D., Nakao,A. and Sugimoto,N. (2002) Molecular crowding regulates the structural switch of the DNA G-quadruplex. *Biochemistry*, **41**, 15017–15024.
24. Miyoshi,D., Nakao,A., Toda,T. and Sugimoto,N. (2001) Effect of divalent cations on antiparallel G-quartet structure of d(G4T4G4). *FEBS Lett.*, **496**, 128–133.
25. Miyoshi,D., Nakao,A. and Sugimoto,N. (2003) Structural transition from antiparallel to parallel G-quadruplex of $d(G_4T_4G_4)$ induced by Ca^{2+}. *Nucleic Acids Res.*, **31**, 1156–1163.
26. Hänsel,R., Löhr,F., Foldynová-Trantírková,S., Bamberg,E., Trantírek,L. and Dötsch,V. (2011) The parallel G-quadruplex structure of vertebrate telomeric repeat sequences is not the preferred folding topology under physiological conditions. *Nucleic Acids Res.*, **39**, 5768–5775.
27. De Cian,A. and Mergny,J.-L. (2007) Quadruplex ligands may act as molecular chaperones for tetramolecular quadruplex formation. *Nucleic Acids Res.*, **35**, 2483–2493.
28. Han,H., Cliff,C.L. and Hurley,L.H. (1999) Accelerated assembly of G-quadruplex structures by a small molecule. *Biochemistry*, **38**, 6981–6986.
29. Pivetta,C., Lucatello,L., Paul Krapcho,A., Gatto,B., Palumbo,M. and Sissi,C. (2008) Perylene side chains modulate G-quadruplex conformation in biologically relevant DNA sequences. *Bioorg. Med. Chem.*, **16**, 9331–9339.
30. Gonçalves,D.P.N., Rodriguez,R., Balasubramanian,S. and Sanders,J.K.M. (2006) Tetramethylpyridiniumporphyrazines - a new class of G-quadruplex inducing and stabilizing ligands. *Chem. Commun.*, 4685–4687.
31. Zhang,H., Xiao,X., Wang,P., Pang,S., Qu,F., Ai,X. and Zhang,J. (2009) Conformational conversion of DNA G-quadruplex induced by a cationic porphyrin. *Spectrochim. Acta A Mol. Biomol. Spectrosc.*, **74**, 243–247.
32. Zhou,J. and Yuan,G. (2007) Specific recognition of human telomeric G-quadruplex DNA with small molecules and the conformational analysis by ESI mass spectrometry and circular dichroism spectropolarimetry. *Chem. Eur. J.*, **13**, 5018–5023.
33. Arola-Arnal,A., Benet-Buchholz,J., Neidle,S. and Vilar,R. (2008) Effects of metal coordination geometry on stabilization of human telomeric quadruplex DNA by square-planar and square-pyramidal metal complexes. *Inorg. Chem.*, **47**, 11910–11919.
34. Gonçalves,D.P.N., Ladame,S., Balasubramanian,S. and Sanders,J.K.M. (2006) Synthesis and G-quadruplex binding studies of new 4-N-methylpyridinium porphyrins. *Org. Biomol. Chem.*, **4**, 3337–3342.
35. Bhattacharjee,A.J., Ahluwalia,K., Taylor,S., Jin,O., Nicoludis,J.M., Buscaglia,R., Chaires,J.B., Kornfilt,D.J.P. and Yatsunyk,L.A. (2011) Induction of G-quadruplex DNA structure by Zn(II) 5,10,15,20-tetrakis(N-methyl-4-pyridyl)porphyrin. *Biochimie*, **93**, 1297–1309.
36. Monchaud,D. and Teulade-Fichou,M.P. (2008) A hitchhiker's guide to G-quadruplex lilgands. *Org. Biomol. Chem.*, **6**, 627–636.
37. Sexton,A.N. and Collins,K. (2011) The 5' guanosine tracts of human telomerase RNA are recognized by the G-quadruplex binding domain of the RNA helicase DHX36 and function to increase RNA accumulation. *Mol. Cell. Biol.*, **31**, 736–743.
38. De Cian,A., Cristofari,G., Reichenbach,P., De Lemos,E., Monchaud,D., Teulade-Fichou,M.-P., Shin-ya,K., Lacroix,L., Lingner,J. and Mergny,J.-L. (2007) Reevaluation of telomerase inhibition by quadruplex ligands and their mechanisms of action. *Proc. Natl Acad. Sci. USA*, **104**, 17347–17352.
39. Lacroix,L., Séosse,A. and Mergny,J.-L. (2011) Fluorescence-based duplex–quadruplex competition test to screen for telomerase RNA quadruplex ligands. *Nucleic Acids Res.*, **39**, e21.
40. Ragazzon,P. and Chaires,J.B. (2007) Use of competition dialysis in the discovery of G-quadruplex selective ligands. *Methods*, **43**, 313–323.
41. Tran,P.L.T., Largy,E., Hamon,F., Teulade-Fichou,M.-P. and Mergny,J.-L. (2011) Fluorescence intercalator displacement assay for screening G4 ligands towards a variety of G-quadruplex structures. *Biochimie*, **93**, 1288–1296.
42. Hamon,F., Largy,E., Guédin-Beaurepaire,A., Rouchon-Dagois,M., Sidibe,A., Monchaud,D., Mergny,J.-L., Riou,J.-F., Nguyen,C.-H. and Teulade-Fichou,M.-P. (2011) An acyclic oligoheteroaryle that discriminates strongly between diverse G-quadruplex topologies. *Angew. Chem. Int. Ed.*, **50**, 8745–8749.
43. De Matteis,F., Gibbs,A.H. and Smith,A.G. (1980) Inhibition of protohaem ferro-lyase by N-substituted porphyrins. *Biochem. J.*, **189**, 645–648.
44. Li,Y., Geyer,C.R. and Sen,D. (1996) Recognition of anionic porphyrins by DNA aptamers. *Biochemistry*, **35**, 6911–6922.
45. Ren,J. and Chaires,J.B. (1999) Sequence and structural selectivity of nucleic acid binding ligands. *Biochemistry*, **38**, 16067–16075.
46. Arthanari,H., Basu,S., Kawano,T.L. and Bolton,P.H. (1998) Fluorescent dyes specific for quadruplex DNA. *Nucleic Acids Res.*, **26**, 3724–3728.
47. Huber,M.D., Lee,D.C. and Maizels,N. (2002) G4 DNA unwinding by BLM and Sgs1p: substrate specificity and substrate-specific inhibition. *Nucleic Acids Res.*, **30**, 3954–3961.
48. Wu,X. and Maizels,N. (2001) Substrate-specific inhibition of RecQ helicase. *Nucleic Acids Res.*, **29**, 1765–1771.
49. Paramasivan,S. and Bolton,P.H. (2008) Mix and measure fluorescence screening for selective quadruplex binders. *Nucleic Acids Res.*, **36**, e106.
50. Hershman,S.G., Chen,Q., Lee,J.Y., Kozak,M.L., Yue,P., Wang,L.-S. and Johnson,F.B. (2008) Genomic distribution and functional analyses of potential G-quadruplex-forming sequences in Saccharomyces cerevisiae. *Nucleic Acids Res.*, **36**, 144–156.
51. Smith,J.S. and Johnson,F.B. (2010) G-quadruplex DNA: methods and protocols. In: Baumann,P. (ed), *Methods in Molecular Biology,* Vol. 608. Humana Press, NYC, pp. 207–221.
52. Qin,H., Ren,J., Wang,J. and Wang,E. (2010) G-quadruplex facilitated turn-off fluorescent chemosensor for selective detection of cupric ion. *Chem. Commun.*, **46**, 7385–7387.
53. Cantor,C.R., Warshaw,M.M. and Shapiro,H. (1970) Oligonucleotide interactions. 3. Circular dichroism studies of the conformation of deoxyoligonucleotides. *Biopolymers*, **9**, 1059–1077.
54. Mergny,J.-L., Li,J., Lacroix,L., Amrane,S. and Chaires,J.B. (2005) Thermal difference spectra: a specific signature for nucleic acid structures. *Nucleic Acids Res.*, **33**, 1182–1192.
55. Job,P. (1928) Formation and stability of inorganic complexes in solution. *Ann. Chim.*, **9**, 113–203.
56. Gray,R.D. and Chaires,J.B. (2011) *Current Protocols in Nucleic Acid Chemistry*. John Wiley & Sons, Inc., Hoboken NJ, pp. 17.14.11–17.14.16.
57. De Cian,A., Guittat,L., Kaiser,M., Saccà,B., Amrane,S., Bourdoncle,A., Alberti,P., Teulade-Fichou,M.-P., Lacroix,L. and Mergny,J.-L. (2007) Fluorescence-based melting assays for studying quadruplex ligands. *Methods*, **42**, 183–195.
58. Ramsay,G.D. and Eftink,M.R. (1994) Analysis of multidimensional spectroscopic data to monitor unfolding of proteins. *Methods Enzymol.*, **240**, 615–645.
59. Gray,R.D., Li,J. and Chaires,J.B. (2009) Energetics and kinetics of a conformational switch in G-quadruplex DNA. *J. Phys. Chem. B*, **113**, 2676–2683.
60. Renciuk,D., Kejnovska,I., Skolakova,P., Bednarova,K., Motlova,J. and Vorlickova,M. (2009) Arrangement of human

telomere DNA quadruplex in physiologically relevant K^+ solutions. *Nucleic Acids Res.*, **37**, 6625–6634.
61. Gray,R.D. and Chaires,J.B. (2011) Linkage of cation binding and folding in human telomeric quadruplex DNA. *Biophys. Chem.*, **159**, 205–209.
62. Paramasivan,S., Rujan,I. and Bolton,P.H. (2007) Circular dichroism of quadruplex DNAs: applications to structure, cation effects and ligand binding. *Methods*, **43**, 324–331.
63. Karsisiotis,A.I., Hessari,N.M.A., Novellino,E., Spada,G.P., Randazzo,A. and Webba da Silva,M. (2011) Topological characterization of nucleic acid G-quadruplexes by UV absorption and circular dichroism. *Angew. Chem. Int. Ed.*, **50**, 10645–10648.
64. Dai,J., Dexheimer,T.S., Chen,D., Carver,M., Ambrus,A., Jones,R.A. and Yang,D. (2006) An intramolecular G-quadruplex structure with mixed parallel/antiparallel G-strands formed in the human BCL-2 promoter region in solution. *J. Am. Chem. Soc.*, **128**, 1096–1098.
65. Palumbo,S.L., Ebbinghaus,S.W. and Hurley,L.H. (2009) Formation of a unique end-to-end stacked pair of G-quadruplexes in the hTERT core promoter with implications for inhibition of telomerase by G-quadruplex-interactive ligands. *J. Am. Chem. Soc.*, **131**, 10878–10891.
66. Phan,A.T., Kuryavyi,V., Burge,S., Neidle,S. and Patel,D.J. (2007) Structure of an unprecedented G-quadruplex scaffold in the human c-kit promoter. *J. Am. Chem. Soc.*, **129**, 4386–4392.
67. González,V. and Hurley,L.H. (2010) The c-MYC NHE III1: function and regulation. *Annu. Rev. Pharmacol. Toxicol.*, **50**, 111–129.
68. Sun,D., Guo,K. and Shin,Y.-J. (2011) Evidence of the formation of G-quadruplex structures in the promoter region of the human vascular endothelial growth factor gene. *Nucleic Acids Res.*, **39**, 1256–1265.
69. De Armond,R., Wood,S., Sun,D., Hurley,L.H. and Ebbinghaus,S.W. (2005) Evidence for the presence of a guanine quadruplex forming region within a polypurine tract of the hypoxia inducible factor 1α promoter. *Biochemistry*, **44**, 16341–16350.
70. Macaya,R.F., Schultze,P., Smith,F.W., Roe,J.A. and Feigon,J. (1993) Thrombin-binding DNA aptamer forms a unimolecular quadruplex structure in solution. *Proc. Natl Acad. Sci. USA*, **90**, 3745–3749.
71. Pasternack,R.F., Gibbs,E.J. and Villafranca,J.J. (1983) Interactions of porphyrins with nucleic acids. *Biochemistry*, **22**, 2406–2414.
72. Miller,M.C., Buscaglia,R., Chaires,J.B., Lane,A.N. and Trent,J.O. (2010) Hydration is a major determinant of the G-quadruplex stability and conformation of the human telomere 3′ sequence of $d(AG_3(TTAG_3)_3)$. *J. Am. Chem. Soc.*, **132**, 17105–17107.
73. De Cian,A., DeLemos,E., Mergny,J.-L., Teulade-Fichou,M.-P. and Monchaud,D. (2007) Highly efficient G-quadruplex recognition by bisquinolinium compounds. *J. Am. Chem. Soc.*, **129**, 1856–1857.
74. Lu,M., Gou,Q. and Kallenbach,N.R. (1992) Structure and stability of sodium and potassium complexes of dT_4G_4 and dT_4G_4T. *Biochemistry*, **31**, 2455–2459.
75. Alberti,P. and Mergny,J.-L. (2003) DNA duplex–quadruplex exchange as the basis for a nanomolecular machine. *Proc. Natl Acad. Sci. USA*, **100**, 1569–1573.
76. Chaires,J.B. (2010) Human telomeric G-quadruplex: thermodynamic and kinetic studies of telomeric quadruplex stability. *FEBS J.*, **277**, 1098–1106.
77. Mergny,J.-L., Phan,A.-T. and Lacroix,L. (1998) Following G-quartet formation by UV-spectroscopy. *FEBS Letters*, **435**, 74–78.
78. Georgiades,S.N., Abd Karim,N.H., Suntharalingam,K. and Vilar,R. (2010) Interaction of metal complexes with G-quadruplex DNA. *Angew. Chem. Int. Ed.*, **49**, 4020–4034.
79. Rodriguez,R., Pantoş,G.D., Gonçalves,D.P.N., Sanders,J.K.M. and Balasubramanian,S. (2007) Ligand-driven G-quadruplex conformational switching by using an unusual mode of interaction. *Angew. Chem. Int. Ed.*, **46**, 5405–5407.
80. Ren,L., Zhang,A., Huang,J., Wang,P., Weng,X., Zhang,L., Liang,F., Tan,Z. and Zhou,X. (2007) Quaternary ammonium zinc phthalocyanine: inhibiting telomerase by stabilizing G-quadruplexes and inducing G-quadruplex structure transition and formation. *ChemBioChem*, **8**, 775–780.
81. Senge,M.O., Kalisch,W.W. and Runge,S. (1997) *N*-Methyl derivatives of highly substituted porphyrins – the combined influence of both core and peripheral substitution on the porphyrin conformation. *Liebigs Annalen*, **1997**, 1345–1352.
82. Phan,A.T., Kuryavyi,V., Gaw,H.Y. and Patel,D.J. (2005) Small-molecule interaction with a five-guanine-tract G-quadruplex structure from the human MYC promoter. *Nat. Chem. Biol.*, **1**, 167–173.
83. Parkinson,G.N., Ghosh,R. and Neidle,S. (2007) Structural basis for binding of porphyrin to human telomeres. *Biochemistry*, **46**, 2390–2397.
84. Haq,I., Chowdhry,B.Z. and Chaires,J.B. (1997) Singular value decomposition of 3-D DNA melting curves reveals complexity in the melting process. *Eur. Biophys. J.*, **26**, 419–426.
85. Henry,R.W. and Hofrichter,J. (1992) Singular Value Decomposition: Application to the analysis of experimental data. In: Brand,L. and Johnson,M.L. (eds), *Methods in Enzymology*, Vol. 210. Academic Press, Waltham, MA, pp. 129–191.
86. Gray,R.D. and Chaires,J.B. (2008) Kinetics and mechanism of K^+- and Na^+-induced folding of models of human telomeric DNA into G-quadruplex structures. *Nucleic Acids Res.*, **26**, 4191–4203.
87. Scolaro,L.M., Castriciano,M., Romeo,A., Patane,S., Cefali,E. and Allegrini,M. (2002) Aggregation behavior of Protoporphyrin IX in aqueous solutions: clear evidence of vesicle formation. *J. Phys. Chem. B*, **106**, 2453–2459.

A Copy of the Paper
"Optimized end-stacking provides specificity of N-methyl mesoporphyrin IX for human telomeric G-quadruplex DNA"

Reference:
Nicoludis, J.M.; Miller, S.T.; Jeffrey, P.D.; Barrett, S. P.; Rablen, P.R.; Lawton, T.J.; Yatsunyk, L.A., Optimized end-stacking provides specificity of N-methyl mesoporphyrin IX for human telomeric G-quadruplex DNA. *J. Am. Chem. Soc.* **2012**, *134, 20446–20456.*

Optimized End-Stacking Provides Specificity of *N*-Methyl Mesoporphyrin IX for Human Telomeric G-Quadruplex DNA

John M. Nicoludis,[†,∥] Stephen T. Miller,[†] Philip D. Jeffrey,[‡] Steven P. Barrett,[†] Paul R. Rablen,[†] Thomas J. Lawton,[§] and Liliya A. Yatsunyk*[,†]

stabilized by a central monovalent metal cation. Various forms of the human telomeric GQs have been structurally characterized by NMR and X-ray crystallography.[1] In the crystalline state, human telomeric DNA preferentially folds into an all-parallel conformation with three TTA propeller loops (Figure 1b).[2] Under dilute solution conditions in Na^+ buffer, telomeric DNA folds into an antiparallel conformation,[3] and in K^+ buffer it adopts three mixed-polarity conformations depending on the exact sequence.[1] In solution, the all-parallel propeller loop GQ structure can be induced by polyethylene glycol (PEG), ethanol, high DNA concentration ([nucleoside] > 100 mM),[4,5] or *N*-methyl mesoporphyrin IX (NMM).[6]

GQ structures can be stabilized by small molecule ligands[7,8] including porphyrins.[9] Ligands that can selectively stabilize GQs, but not double-stranded DNA (dsDNA), may have potential applications as pharmaceuticals for cancer therapy.

architecture of human telomeric quadruplexes (both DNA and RNA) in complexes with ligands, of which 11 were determined by X-ray crystallography. These ligands include a set of tetra-substituted naphthalene diimide compounds,[15,16] nickel and copper salphen complexes,[17] an acridine compound,[18] berberine,[19] BRACO-19,[20] and a well-studied porphyrin, TMPyP4.[21] With the exception of TMPyP4, the ligands stack onto a G-tetrad. Available crystal structures have revealed some elements required for efficient quadruplex stabilization by a small molecule; however, the limited number of crystal structures and, in some cases, their relatively low resolution have prevented generalization of these features.[20,22−24] In addition,

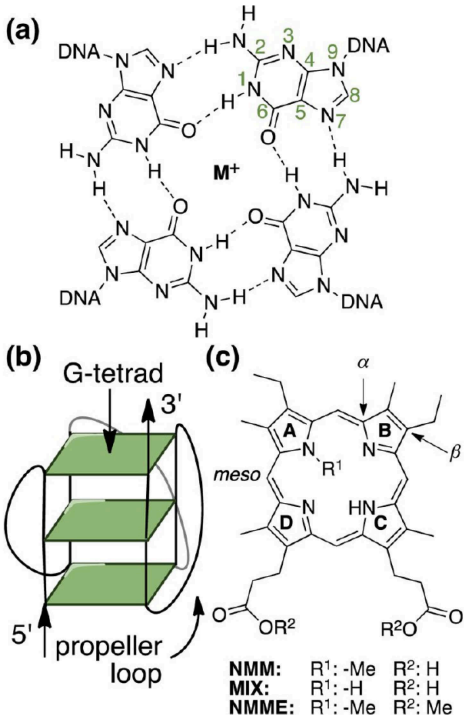


Figure 1. Structure of a G-tetrad and NMM. (a) Schematic representation of a G-tetrad including numbering of the atoms of one guanine. (b) A parallel G-quadruplex. (c) Structure of *N*-methyl mesoporphyrin IX, NMM. Only one of the four regioisomers of NMM

most of the crystallized ligands have only moderate selectivity for GQs over dsDNA. Structural information on highly GQ-selective ligands, their mode of binding to quadruplex DNA, and knowledge of molecular features required for selectivity will be invaluable for future ligand design.

Recently, our group conducted an in-depth biophysical investigation of interactions between NMM, shown in Figure 1c, and human telomeric DNA, dAGGG(TTAGGG)$_3$, Tel22.[6] First identified as a GQ binder through a SELEX approach to find porphyrin aptamers,[25] NMM displayed exceptional selectivity for GQ versus dsDNA in competition dialysis and fluorescence studies[22,26] and, based on this property, is currently widely used in chemistry and biology.[27−29] Even more interestingly, our laboratory was able to demonstrate that NMM has the unique ability to recognize parallel-stranded but not antiparallel GQ structures.[6] The molecular basis for such selectivity requires further investigation. It is important to note that the commercially available NMM is a racemic mixture of eight isomers: four regioisomers, which differ in the position of the *N*-Me group, each of which has a pair of enantiomers that differ in the side of the macrocycle from which the *N*-Me group protrudes. The enantiomers do not easily interconvert at ordinary temperatures; the free energy barrier to racemization is calculated to be 54.3 kcal mol^{-1} (full details of the calculation appear in the Supporting Information). All of the reported biological and biochemical studies involving NMM were performed using this isomer mixture. To better inform these studies, our laboratory set out to characterize the interaction of Tel22 with the isomer mixture of NMM, rather than focus on separate isomers.

In this work, we present the crystal structure of the NMM−Tel22 complex solved in two space groups, $P2_12_12$ and $P6$, at resolutions of 1.65 and 2.15 Å, respectively (Figure 2). The $P2_12_12$ structure is the highest resolution structure of the human telomeric GQ DNA reported to date. Tel22 forms a dimer of 5′-5′ stacked intramolecular parallel quadruplexes with NMM bound to the 3′-terminal G-tetrads. The structure displays optimized surface complementarity between NMM and the 3′ G-tetrad, which results from GQ-induced distortion of the porphyrin macrocycle. The *N*-methyl (*N*-Me) group of NMM points directly into the center of the quadruplex core and is aligned with the column of potassium ions. The binding of NMM to duplex or antiparallel GQ is inhibited due to steric clashes involving this *N*-Me group. Results of our work, thus, provide framework for understanding how NMM interacts with Tel22 and will inform the design of highly selective GQ ligands.

90 °C for 10 min, slowly cooled to 30 °C, and incubated at this

dx.doi.org/10.1021/ja3088746 | *J. Am. Chem. Soc.* 2012, 134, 20446−20456

temperature for 30 h. Crystals were grown using hanging drop vapor diffusion methods at 4 °C. Crystals in the $P2_12_12$ space group (crystal form **1**) were obtained using the well solution of 0.05 M lithium cacodylate buffer, pH 7.2, 0.40 M ammonium sulfate, 0.05 M KCl, 0.01 M $CaCl_2$, and 15% PEG400; crystals in the *P6* space group (crystal form **2**) were obtained using the well solution of 0.05 M lithium cacodylate buffer, pH 7.2, 0.625 M ammonium acetate, 0.20 M KCl, and 15% PEG400. One microliter of sample was mixed with 0.5 or 0.75 μL of well solution for crystal forms **1** and **2**, respectively. Crystals grew in a week as hexagonal or rhombic prisms (Figure 2a); they were harvested and cryoprotected in a solution of the mother liquor with an additional 15% PEG400, bringing total PEG400 concentration to 30%.

Data were collected at NSLS Beamline X29 at Brookhaven National Laboratories using an ADSC 315r detector (for crystal form **1**) and at LS-CAT sector 21 line G at Argonne National Laboratories using a MAR 300 CCD (for crystal form **2**). Diffraction data were processed with MOSFLM.[31] The structures were solved via molecular replacement using PHENIX[32] with the DNA G-quadruplex from PDB entry 3T5E as the search model. Models were built using iterative cycles of building in Coot[33] and refinement in PHENIX followed by final refinement in REFMAC5.[34] The ligand was placed using the automated LigandFit PHENIX module and subsequently refined using PHENIX and REFMAC5. Refinement statistics and final geometries are given in Table 1. For details on Tel22 hydration sphere, see Supporting Information Figure S2.

unit cell dimensions		
a, *b*, *c* (Å)	50.755, 41.980, 42.384	64.200, 64.200, 42.360
α, β, γ (deg)	90, 90, 90	90, 90, 120
resolution (Å)	42.38−1.65	42.36−2.15
(highest resolution shell)	(1.74−1.65)	(2.27−2.15)
R_{merge} (%) overall	9.9 (46.2)	6.6 (63.8)
I/σ	11.1 (2.8)	12.8 (2.5)
completeness (%)	99.83 (98.2)	98.8 (84.4)
redundancy	8.0 (6.3)	7.9 (6.7)
refinement		
resolution (Å)	42.38−1.65	42.36−2.15
reflections	10 820	5056
R_{work}/R_{free} (%)	22.21/26.18	23.79/26.21
no. of atoms	570	518
ions	3	3
water	53	7
overall *B*-factor (Å^2)	34.223	53.195
rms deviations		

Molecular graphics and structure alignments were produced using PyMOL.[35] For alignment, PyMOL's built-in Pair_Fit function was used, which aligned all 465 atoms of native Tel22 (PDB ID 1KF1) to the structures of the Tel22−ligand complexes. In the case of BMSG-SH-3 (PDB ID 3SC8), the 21 nucleotides (447 atoms) were aligned with nucleotides 2−22 of 1KF1. Structure factors and coordinates for NMM−Tel22 complex in $P2_12_12$ and *P6* space groups have been deposited in the Protein Data Bank with accession codes of 4FXM and 4G0F, respectively.

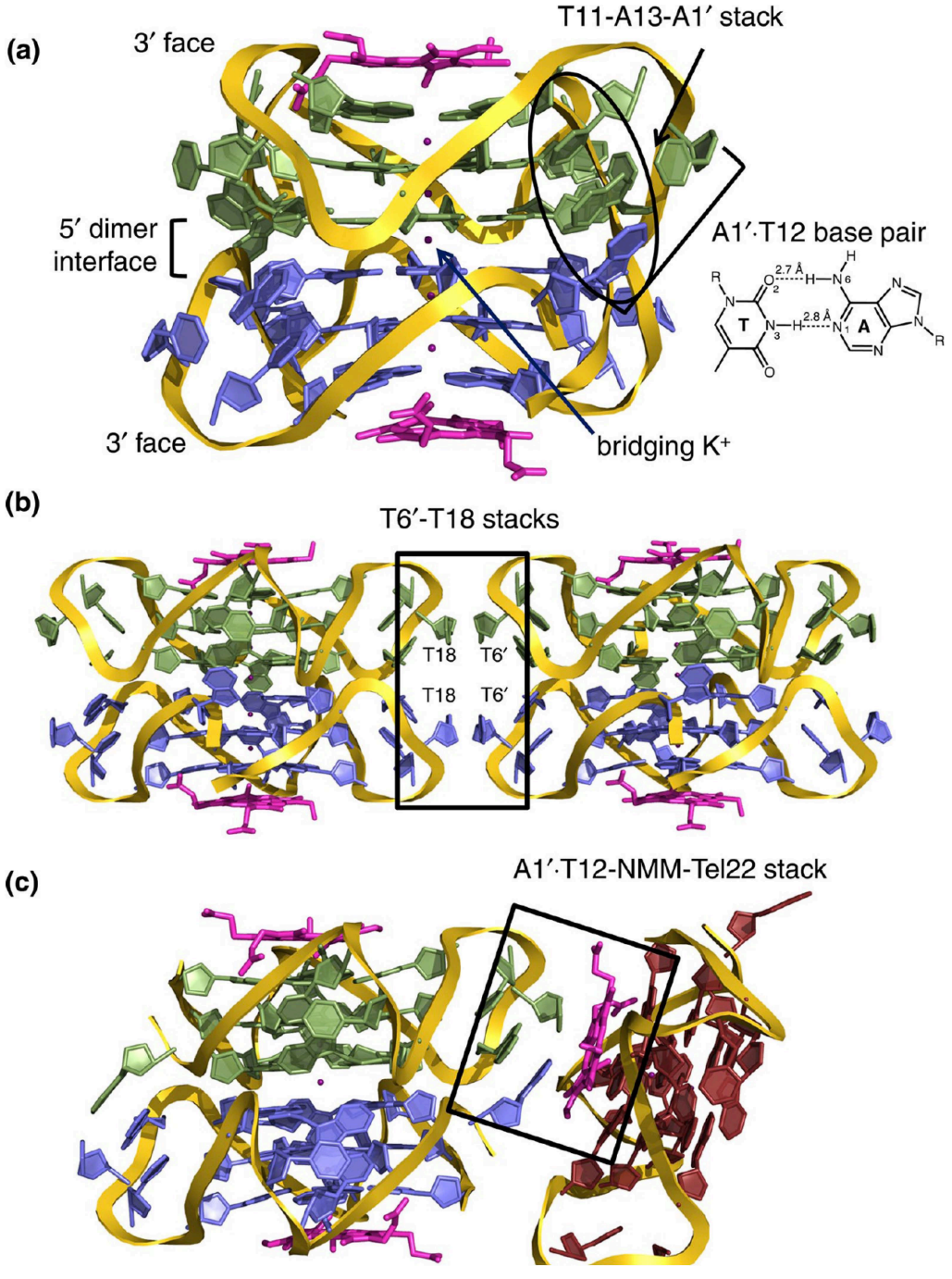
Calculations. Energy-minimization calculations were used to obtain the solution- and gas-phase structures of NMM and mesoporphyrin IX (MIX). The X-ray structure coordinates of NMM from **1** were used as a starting point, with hydrogen atoms added at idealized positions. The initial MIX structure was generated by replacing the methyl group of NMM with a hydrogen atom. The geometry was optimized without constraint in the gas phase and in the presence of a simulated water solvent at B3LYP/6-31G(d).[36] All stationary points were verified via the calculation of vibrational frequencies. The calculation in the presence of a simulated solvent was performed using the Polarizable Continuum Model with integral equation formalism.[37] This model treats the solvent as a continuum characterized by a dielectric constant and having a cavity to accommodate the solute. The transition structure for NMM racemization was located using a manually generated initial guess and was verified using a frequency calculation (one imaginary frequency, with the correct motion) and intrinsic reaction coordinate (IRC) following (IRC pathways led to the two enantiomers of NMM).[38] All calculations were carried out using the Gaussian 09 package.[39]

Normal-Coordinate Structural Decomposition. NMM distortion was calculated using the normal-coordinate structural decomposition (NSD) engine.[40] In short, this software quantifies the in-plane and out-of-plane distortions of a porphyrin macrocycle from its structural coordinates (24 atoms). The out-of-plane porphyrin distortion can be represented as a linear combination of the six lowest-frequency normal deformations: saddled (B_{2u}), ruffled (B_{1u}), domed (A_{2u}), waved in *x* ($E_{g(x)}$) and in *y* directions ($E_{g(y)}$), and propeller (A_{1u}). The total out-of-plane deviation (D_{oop}) is calculated as the root-sum-square (rss) of the component deviations (d_{oop}) from an idealized planar D_{4h}-symmetric porphyrin.

Principal Component Analysis. Out-of-plane displacements for individual atoms of the NMM macrocycle and the G-tetrads were calculated using principal component analysis (PCA) in MATLAB (Mathworks). In this method, the three-dimensional coordinates of the NMM macrocycle (24 atoms) and the G-tetrads (11 atoms of each guanine group) were transformed into standard axis, where the first two components represent the in-plane vectors (the least-squares plane) and the third represents the normal vector, which signifies the out-of-plane deviations of individual atoms. Like in NSD, the D_{oop} was calculated as the rss of the deviations of individual atoms from the mean NMM or G-tetrad plane.

FRET Melting Assays. Fluorescence resonance energy transfer (FRET) melting assays were performed according to the established literature procedure[41] using doubly labeled oligonucleotide F21D, 5′-6-FAM-$G_3(TTAG_3)_3$-Dabcyl-3′, purchased from IDT. FRET competition experiments were conducted to establish the ability of NMME to differentiate between GQ and dsDNA and utilized a self-complementary 26-mer oligonucleotide 5′-CAATCGGATC-GAATTCGATCCGATTG-3′ (ds26, Midland) as a duplex competitor. FRET experiments were run in 5K buffer (10 mM lithium cacodylate, pH 7.2, 5 mM KCl, and 95 mM LiCl) using 0.2 μM of F21D.

UV−Vis and CD Studies. UV−vis spectra were collected on a Cary 300 (Varian) spectrophotometer equipped with a Peltier-thermostatted cuvette holder. UV−vis titration studies were performed by the stepwise addition of Tel22 into a solution of NMM at pH 5.8 and 8.6 as described previously.[6] CD experiments were performed on an AVIV 410 spectrometer equipped with a Peltier heating unit (temperature accuracy ±0.3 °C). CD annealing studies were performed using NMM−Tel22 and NMME−Tel22 in 5K buffer as described previously.[6] Additional experimental details can be found in the Supporting Information.



shape (Figure 2). Discussions in this Article focus on crystal form **1** unless otherwise noted.

In both crystal forms, Tel22 adopts a parallel propeller loop three G-tetrad quadruplex also observed in the structures of human telomeric DNA alone or with a variety of ligands.[2,15,16] This topology is most likely a result of the high DNA and K^+ concentrations required for crystal growth as well as NMM's influence on Tel22 structure. Specifically, NMM was shown to cause a structural transition of Tel22 to a parallel fold under dilute DNA and K^+ conditions.[6] Therefore, the parallel GQ core topology was expected for NMM−Tel22 crystals.

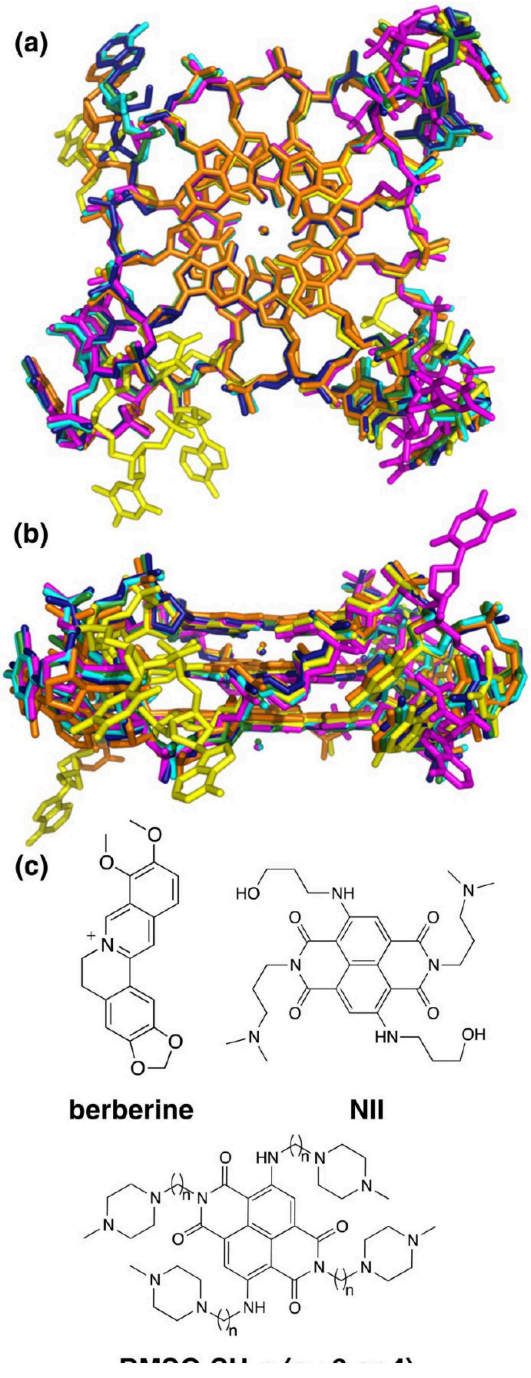
Two symmetry related Tel22 quadruplexes are arranged into a dimer, shown in Figure 3a, which is stabilized by the 5′−5′ $\pi-\pi$ stacking of G-tetrads and the bridging K^+ ion, bringing the total K^+ count to five ions per dimer. A similar dimer is also observed in the structure of Tel22 alone,[2] and complexed with

dx.doi.org/10.1021/ja3088746 | J. Am. Chem. Soc. 2012, 134, 20446−20456

naphthalene diimide ligands BMSG-SH-3 and -4.[16] The dimer is further stabilized by an A1′·T12 reverse Watson–Crick base pair (the prime (′) notation signifies that the two bases belong to separate oligonucleotide strands) with hydrogen-bonding distances of 2.7 Å for the N6 A1′···O2 T12 and 2.8 Å for the N1 A1′···N3 T12 interaction. In addition, A1′ is part of an A1′–A13–T11 π–π stacked system that further stabilizes the dimer. Although the multiplicity of interactions between the monomers in the Tel22 dimer strongly hints at its stability, previous PAGE experiments have demonstrated that the NMM–Tel22 complex is monomeric in solution.[6] The observed discrepancy could originate from the drastically different experimental conditions. For crystallization, the concentration of Tel22 was 0.69–0.80 mM, and that of K^+ ions was 50 mM. In contrast, in solution studies, only 40 μM Tel22 in 5 mM KCl was loaded on a PAGE gel. To resolve this discrepancy, UV–vis and CD melting studies were performed on Tel22 samples (without NMM, see the Supporting Information for details) in the concentration range from 0.020 to 0.62 mM in the presence of 5 and 50 mM KCl. The results indicate that Tel22 is monomolecular under these conditions (Figure S3). Thus, the Tel22 dimer observed in the crystalline state is most likely due to crystal packing forces as its formation leads to a significantly reduced exposed quadruplex surface. Two neighboring dimers interact with each other via two pairs of stacked T6′–T18 bases (Figure 3b) whose planes are separated by 3.3 ± 0.1 Å. These bases are derived from the propeller TTA loops.

The Tel22 dimer is capped on both ends by NMM as shown in Figure 3a leading to a 1:1 binding stoichiometry in agreement with our earlier spectroscopic data.[6] NMM stacks onto 3′ G-tetrad, and its *N*-Me group points into the quadruplex core where it aligns with the column of K^+ ions. The other face of NMM interacts with an A1′·T12 base pair of a nearby dimer, as shown in Figure 3c. There are two choices in the stacking orientation of the A1′·T12 base pair with respect to NMM that are related by a 180° rotation around an axis perpendicular to the porphyrin plane. The choice of orientation determines the space group and leads to a specific three-dimensional arrangement of molecules in the crystal (Figure 2b and c). The stacking between the A1′·T12 base pair and NMM likely results from the crystal packing and has no (or minimal) impact on biologically significant interaction of NMM with 3′ G-tetrad as evidenced by the similar geometry of NMM in crystal forms **1** and **2** (see below).

Comparison of Native and Ligand-Bound Tel22 G-Quadruplex DNA Structures. First, the structures of the Tel22 GQs from crystal forms **1** and **2** were compared to each other. The root-mean-square deviation (rmsd) of the 465 atoms is small, 0.332 Å, suggesting a nearly identical fold. A variety of other ligands have been crystallized with the same DNA sequence: berberine,[19] naphthalene diimide ligands BMSG-SH-3 and -4,[16] and 2,7-bis[3-(dimethylamino)propyl]-4,9-bis[(3-hydroxypropyl)amino]benzo[*lmn*]-[3,8]-phenanthroline-1,3,6,8(2*H*,7*H*)-tetrone (called NII here).[15] All ligands bind to the 3′ G-tetrad of Tel22 GQ; in addition, berberine and NII intercalate into the 5′-5′ stacked Tel22 dimer and another pair of NII intercalates between loop nucleotides. These six unique Tel22 structures were aligned using PyMOL (Figure 4 and Figure S4), and the rmsd values, referenced to native Tel22 structure (PDB 1KF1), are listed in Table 2. The low values of the rmsd indicate that Tel22 in all complexes adopts a similar quadruplex fold independent of ligand binding. The larger rmsd values seen for Tel22 in complex with the NII (2.989 Å) and BMSG-SH-3 (2.242 Å) arise mainly from variation in the TTA propeller loops. When structures of quadruplex cores are compared (using coordinates of 12 guanines), the rmsd values decrease to 0.550–0.896 Å (Table 2). The positions of the K^+ ions within the quadruplex core, the distances between the carbonyl oxygens within G-tetrads (O6–O6 distance), and the Hoogsteen hydrogen-bonding distances (N1–O6 and N2–N7) for the ligand



dx.doi.org/10.1021/ja3088746 | J. Am. Chem. Soc. 2012, 134, 20446–20456

binding 3′ G-tetrad are nearly the same for all structures (Figure S4). These results indicate that the quadruplex core geometry is well maintained in all complexes.

It was interesting to compare the mode of porphyrin binding to human telomeric quadruplex in NMM−Tel22 and TMPyP4−$(dTAG_3TTAG_3)_2$ complexes,[21] Figure S5. Unlike NMM, TMPyP4 does not interact directly with a G-tetrad; rather one molecule of TMPyP4, aligned with GQ axis, is sandwiched between A−T base pairs of two dimers, and another molecule, perpendicular to GQ axis, binds to the loops via stacking with four T bases. We believe that the ruffled geometry of TMPyP4 together with the pyridine substituents that are perpendicular to the macrocycle prevents TMPyP4 from stacking directly onto G-tetrad.

Specific Interactions between NMM and Tel22. In the crystal form **1**, the core of NMM and its *N*-Me group are well-defined, while peripheral groups, especially the propionates, are largely disordered as judged by the observed electron density and the value of the temperature (B) factors, Figure 5. In the crystal form **2**, the core *N*-Me group is not visible in the electron density map, but the NMM macrocycle is still well-defined. Its domed distortion (see below) strongly suggests that the *N*-Me group points toward the 3′ G-tetrad. Our crystallographic data indicate that there is one possible way for NMM to bind to the Tel22 quadruplex with respect to the *N*-Me group (Figure 5), but it is not immediately clear why this orientation is preferred. It is possible that a single NMM isomer crystallizes with Tel22 or that multiple isomers bind Tel22 with identical orientations of the *N*-Me group but different placements of peripheral substituents. In an effort to identify unbiased density for the peripheral substituents of NMM, an iterative build omit map was calculated using PHENIX.[42] While the position of the *N*-Me was still clear in this map, no interpretable density for peripheral substituents was observed. Additional maps were generated in ARP/wARP[43] using the atom update and refinement module and starting models both with and without NMM; again, no clear density for side groups was observed. Thus, at this point, it is impossible to distinguish thermal motion of a single NMM isomer bound to Tel22 from multiple occupancies of more than one NMM isomer.

Using the Ligand Expo program,[44] a previously defined NMM isomer with component identifier MMP was chosen as the most appropriate NMM model, based on some evidence for the location of the propionate groups in electron density map (Figure S6). This isomer is the (*S*)-enantiomer of the regioisomer with the *N*-Me group attached to ring A (Figure 1c). It was placed automatically using the LigandFit PHENIX module in an effort to remove bias in ligand placement. The NMM model used originates from the structure of a ferrochelatase, terminal enzyme in heme biosynthesis that is responsible for iron insertion into hemes.[45]

Using our NMM−Tel22 model, we examined the specific interactions between Tel22 and NMM's macrocycle, *N*-Me group, and side chains. The macrocycle of NMM is located approximately 3.6 Å from the 3′ G-tetrad, a distance consistent with efficient $\pi-\pi$ stacking. Another porphyrin, TMPyP4, was modeled 4.2 Å above G-tetrad in its NMR structure with the *c-myc* promoter.[46] This significantly longer stacking distance could be explained by the nearly perpendicular orientations of pyridyl groups in TMPyP4 relative to the porphyrin core and steric clashes that would result from bringing this porphyrin closer to the G-tetrad. Nearly planar ligands such as berberine, two different metal−salphen complexes, and a naphthalene diimide compound are located approximately 3.4−3.6 Å away from the terminal G-tetrad.[16,17,19] Thus, the exceptional selectivity of NMM for GQ DNA does not result from its stronger stacking interaction with the 3′ G-tetrad of Tel22 as compared to other ligands.

The *N*-Me group of NMM is bent 44.8° away from the mean porphyrin plane pointing toward the 3′ G-tetrad. It is aligned with the column of K^+ ions, leading to the observed off-center position of the NMM macrocycle toward G22 (Figure 5b). Interestingly, the napahthalene diimide ligands, BMSG-SH-3[16] and BRACO-19,[20] are also stacked asymmetrically onto the human telomeric DNA. In the case of BRACO-19 complex, the off-center position is due to the optimized interactions between the cationic ring nitrogen atom of the ligand and the potassium ion channel; in the case of BMSG-SH-3, it is due to the interaction between the side-chain methyl-piperazine groups of the ligand and Tel22 grooves.

As far as peripheral substituents are concerned, clear electron density is visible for one propionate carboxylate group at 1σ. This group may participate in a hydrogen-bonding network with two backbone phosphates (in one case through a water molecule) as shown in Figure S6a, and, thus, may be protonated under crystallization conditions. The other propionate, which is only visible at 0.5σ, may interact with a 3′ sugar oxygen through a water molecule, Figure S6b. Thus, our current model supports, at least to some degree, the stereospecificity of NMM−Tel22 interactions, and Tel22's preference for one NMM isomer. It is important to note that there is no density connecting modeled carboxylate groups to the porphyrin core and the observed electron density could be equally fitted with water molecules. To resolve ambiguity concerning the stereospecificity of Tel22−NMM interactions,

we are in the process of separating the isomers of NMM. Structural information on separated NMM isomers may shed light on the positions of the peripheral groups and their interactions with the GQ. If Tel22 displays a preference for one isomer over another, it might become important to use separate isomers for current applications of NMM.[27−29]

Structural Basis for Selectivity of NMM for Parallel Quadruplex DNA. Structural features of the NMM−Tel22 complex provide possible explanations for the unique selectivity of NMM for the parallel versus antiparallel topology observed in our earlier biophysical study.[6] Importantly, most quadruplex ligands do not display such selectivity, but rather recognize a broad spectrum of quadruplex structures. Such ligands, while useful in some applications, could lead to undesired effects such as an increase in genome instability or in telomerase processivity if used as drugs.[47]

Parallel and antiparallel GQs formed by Tel22 in K^+ and Na^+ buffers, respectively, differ in loop positions and, possibly, positions of the metal ions within the GQ core. The terminal 3′ G-tetrad in parallel Tel22 GQ is unobstructed and provides an excellent place for ligand binding. In contrast, in the antiparallel basket form of Tel22, both terminal G-tetrads are hindered either by lateral loops that form an obstructing A·T Watson−Crick base pair or by the diagonal loop that runs across the terminal G-tetrad and whose bases stack onto it.[3] The blocked terminal G-tetrads of the antiparallel quadruplex call into question the binding mode of ligands with broad quadruplex specificity. It may be possible for a ligand to displace the obstructing bases of the loops or to bind to other features of the G-quadruplex. This latter explanation has been invoked for TOxaPy, an acyclic mimic of telomestatin.[48] In contrast to NMM, TOxaPy is selective for the antiparallel over the parallel Tel22 GQ. TOxaPy most likely binds to GQ grooves that are present in the antiparallel form of the human telomeric GQ but that are blocked by the external propeller loops in the parallel form. The comparison between TOxaPy and NMM indicates that selectivity for specific GQ topologies is possible and that proper GQ ligand design requires understanding of the molecular details of ligand−GQ interactions.

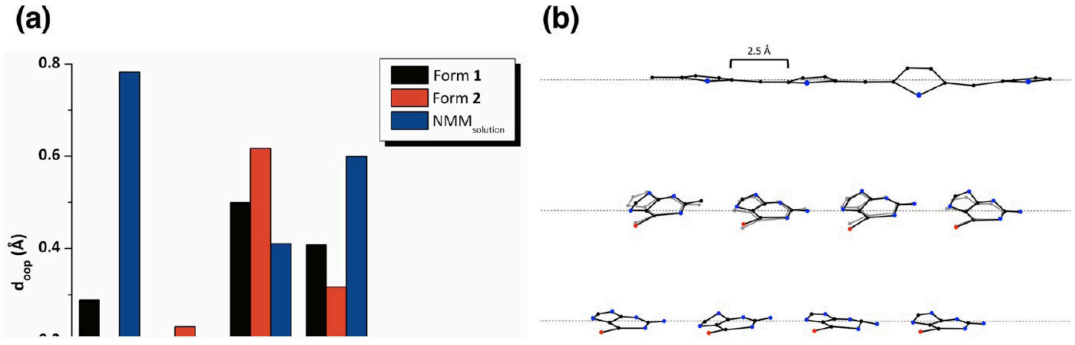
Potassium ions are known to be positioned between the G-tetrads as is observed here (Figure 3). Such positioning creates sufficient space to accommodate the *N*-Me group, and the distance between this group and K^+ is 3.7 Å. Na^+ ions have a smaller ionic radius as compared to K^+ ions and could be positioned within the G-tetrad, coplanar with the guanines. This is indeed observed in the crystal structures of the $d(TG_4T)$ quadruplex where the inner Na^+ ions are almost equidistant from the two G-tetrads, while the outermost Na^+ is positioned within a terminal 3′ G-tetrad.[49−52] This Na^+ ion is also coordinated to a water molecule in the axial position. If such positioning holds true in the antiparallel structure of Tel22 (even without considering the axial water molecule), the Na^+ ions would be sufficiently high up in the channel to cause steric clashes with the *N*-Me group, precluding NMM binding to this form of GQ and thus leading to the observed parallel versus antiparallel GQ selectivity.

The Role of the *N*-Me Group and NMM's Selectivity. In the crystal structure of Ni(II)- and Cu(II)-salphen compounds with a human telomeric GQ, the metal atoms are located almost directly above the ion channel.[17] Similarly, in the BRACO-19 structure, the positively charged ring nitrogen is in line with the potassium ions.[20] In both cases, the observed positioning can be justified by the cationic nature of the metal ion or nitrogen, which mimics the K^+ ions. The position of the *N*-Me group of NMM directly above the potassium channel of G-quadruplex is perplexing. The ion channel might merely provide a space for the *N*-Me group, allowing the porphyrin to stack onto the 3′ G-tetrad without steric clashes (note, other closest contacts of *N*-Me group are with the carbonyl oxygens, 3.3, 3.3, 3.4, and 3.5 Å). In contrast, such interaction is not possible with the duplex base pairs, thus leading to the GQ versus dsDNA selectivity observed in FRET melting studies[6] and in equilibrium dialysis experiments.[53]

The *N*-Me group plays another important role; it distorts the planarity of the NMM macrocycle, making this molecule more susceptible to further distortions induced by interacting biomolecules.[45] It is commonly accepted that nonplanarity of a porphyrin can be induced either by peripheral or core substitution or by interactions with proteins.[54] The amount and type (saddled, ruffled, domed, waved, and propeller) of porphyrin distortion can be quantified using the NSD method pioneered by Shelnutt.[40] The total out-of-plane deviation (D_{oop}) of the NMM macrocycle is 0.727 Å in **1** and 0.747 Å in **2** (Table 3 and Table S2). The NSD analysis indicates that NMM is predominantly domed, with 34.6% and 43.7% domed character for **1** and **2**, respectively, Figure 6a. Other significant nonplanar deformations are wave(x) for **1** (28.3%) and **2** (22.5%) and saddled for **1** (20.0%). Doming deformation of the porphyrin can also be seen in the clothesline plot shown in Figure 6b; all of the pyrrole nitrogen atoms of NMM face downward, whereas the β-carbon atoms are distorted slightly upward. As expected, the pyrrole ring that hosts *N*-Me group is the most distorted.

Out-of-plane distortions are also observed in the G-tetrads, especially the one interacting with NMM, Figure 6b. The carbonyl groups of the 3′ G-tetrad (e.g., the O6 group) point down toward the center of the quadruplex, while the peripheral atoms, N2, C2, N3, C4, and N9, point up, forming an inverted dome-like shape with a total D_{oop} of 1.85 Å. The middle tetrad is significantly less distorted (D_{oop} = 1.08 Å), and the 5′ G-tetrad that forms the dimer interface is nearly planar (D_{oop} = 0.49 Å). The distortion of the 3′ G-tetrad allows for optimized $\pi-\pi$ stacking with nonplanar NMM, but it does not result from the interaction of this G-tetrad with NMM. A similar degree and type of distortion is seen in the 3′ G-tetrad of native Tel22 (Figure 6b, gray line, rmsd 0.241 Å) and Tel22 complexed with other ligands.

The observed domed geometry of NMM is either already present in its solution structure or is induced by its binding to



the dome-shaped 3′ G-tetrad of Tel22. It is expected that NMM in solution is nonplanar due to the steric strain imposed by the *N*-Me group, although, to the best of our knowledge, there is no experimental structure of NMM alone. Therefore, we calculated energy-minimized structures of NMM in the gas and solution phases starting from the NMM's coordinates in the NMM−Tel22 crystal form **1** (Figure S7a). Alone, NMM displays significant saddled, domed, and wave(x) out-of-plane deviations (Figure S7d and Table S2). A similar geometry for NMM (albeit with overall higher deformation) was reported in a resonance Raman study.[55] When NMM structure is compared to that bound to Tel22, the former is significantly more nonplanar (D_{oop} = 1.068 vs 0.727) and saddled (43.2% vs 20.0%), and significantly less domed (22.7% vs 34.6%), Figure 6a. Thus, geometry of NMM changes upon its binding to Tel22, possibly, to optimize its $\pi-\pi$ stacking with the distorted 3′ G-tetrad. NSD analysis of NMM bound to wild-type[45] and H183A mutant[56] of *Bacillus subtilis* ferrochelatase indicates a predominantly saddled conformation of NMM (31.0% for wild-type and 40.1% for H138A mutant, Figure S7e) required to fit the binding pocket of this protein. Therefore, ferrochelatase, similar to Tel22, imposes a specific geometry on NMM.

The observed flexibility of NMM is facilitated by the presence of the *N*-Me group, which causes nonplanar deviation of the porphyrin macrocycle and decreased conjugation. Further distortion of NMM macrocycle, required for efficient binding to Tel22, comes at low energetic cost of only 0.7 kcal mol^{-1} (see Supporting Information). MIX (Figure 1c) has the same macrocycle structure as NMM, but lacks the *N*-Me group. This porphyrin was unable to stabilize Tel22 in FRET melting assays.[6] To understand why this is the case, energy minimization was performed on NMM structure with the *N*-Me group replaced by a hydrogen atom as a model for MIX (Figure S7b). The resulting structure is practically planar (D_{oop} = 0.046) and is approximately 4.3 kcal mol^{-1} more stable than "frozen" MIX, a structure that is forced to stay in the geometry adopted by NMM within the crystal structure (see the Supporting Information). Therefore, a large energy investment is required to distort the macrocycle of MIX to match the geometry of the 3′ G-tetrad of Tel22. In conclusion, the crystal

dx.doi.org/10.1021/ja3088746 | *J. Am. Chem. Soc.* 2012, 134, 20446−20456

structure of the NMM−Tel22 complex questions the commonly accepted requirement for a good GQ ligand to have a planar extended aromatic core, as the interaction of such ligands with Tel22 may be less efficient due to the ligand's inability to optimize its $\pi-\pi$ stacking with the distorted terminal G-tetrad.

The Role of the Propionate Groups. Under physiological conditions, the propionates of NMM are likely deprotonated and may repel the DNA backbone, opposing the stabilizing $\pi-\pi$ stacking and weakening the overall interaction between NMM and Tel22. This argument was used to explain NMM's relatively modest binding constant of $\sim 10^5$ M^{-1}.[6] Interestingly, the model of NMM used in this work to fit the observed electron density suggests that propionates may be involved in hydrogen bonding with the GQ backbone and that one of the carboxylates may be protonated, Figure S6a. To understand the role of propionates in NMM binding to Tel22, we performed CD annealing and FRET melting studies on the dimethyl ester of NMM (NMME). NMME is a neutral derivative of NMM and was chosen for these experiments because it cannot be deprotonated. Similar to NMM, NMME induced structural transition from a mixed hybrid to a parallel conformation of Tel22 in CD annealing studies (Figure 7a). In FRET melting assays, NMME exhibited stabilization of human telomeric DNA comparable to that of NMM and, just like NMM, displayed excellent selectivity toward GQ versus dsDNA (Figure 7b and c). The only advantage NMM has over NMME is its water solubility. These results taken together suggest that the possible negative charge of the propionate groups does not contribute significantly to, or detract significantly from, the binding or selectivity of NMM or its ability to induce structural conversion of Tel22.

To further test the effect of protonation state of the propionates on the strength of NMM interactions with Tel22, we determined the binding affinity at three different pH values: $(1.0 \pm 0.3) \times 10^5$ M^{-1} at pH 7.2,[6] $(1.2 \pm 0.1) \times 10^5$ M^{-1} at pH 5.8, and $(0.3 \pm 0.1) \times 10^5$ M^{-1} at pH 8.6; see Figure S8. The binding affinity remains virtually unchanged in the pH range 5.8−7.2 and drops only slightly at pH 8.6, suggesting that either the protonation state of the propionates does not change in this pH range or that the protonation state does not significantly affect the affinity of NMM for Tel22 in agreement with the conclusion reached above.

It is commonly accepted that increased ionic strength weakens an interaction between a ligand and a biological molecule if this interaction, at least in part, is electrostatic in nature. To test the effect of ionic strength on NMM−Tel22 interactions, we performed CD wavelength and melting experiments under ionic strength from 15 to 410 mM and a constant amount of KCl, 5 mM. The data indicate that the stabilization temperature, $\Delta T_{1/2}$, due to the presence of NMM is independent of the ionic strength up to 110 mM, and then it decreases only slightly (from 7 to 6 °C), Figure S9. Taken together, all biochemical data point to the predominantly nonelectrostatic nature of the NMM−Tel22 interactions.

sort observed for the majority of GQ-stabilizers. Rather, it relies on optimized $\pi-\pi$ stacking and complementarity between the *N*-Me group and a "hole" left by the K^+ ion channel in the center of the Tel22 GQ structure. It is not entirely clear if the latter interaction is merely an avoidance of steric clashes, or a true attraction, based, possibly, on the favorable electron density distribution between the ligand and Tel22 GQ. This specific recognition of the *N*-Me group by the interacting G-tetrad can explain the selectivity of NMM for GQ versus dsDNA and for parallel versus antiparallel GQ folds. Neither duplex DNA nor antiparallel GQ (possibly with a Na^+ ion within the terminal G-tetrad) have sufficient space to accommodate the *N*-Me group of NMM and thus cannot bind to this ligand.

Our previous biochemical characterization of NMM binding to Tel22 was insufficient to explain why MIX, a planar derivative of NMM, could not effectively stabilize Tel22.[6] The model of the NMM−Tel22 complex presented here suggests that the energetic barrier to the structural adjustment of MIX required for its optimal interaction with Tel22 is too high (4.3 kcal mol^{-1} for MIX versus 0.7 kcal mol^{-1} for NMM).

It has been shown previously that GQ DNA can undergo structural changes to optimize ligand binding, as in the case of the X-ray structure of BRACO-19 with a bimolecular GQ[20] and the NMR structure of *c-myc* GQ with a quindoline ligand.[57] Here, we report for the first time that a ligand, NMM, is also capable of adjusting its geometry (from saddled to domed) upon GQ binding. The complementarity of the surface of the terminal 3′ G-tetrad with NMM's macrocycle and its *N*-Me group leads to observed selectivity and to a 1:1 binding stoichiometry. Although it is widely believed that the extended aromaticity of the ligand is vital for its $\pi-\pi$ interactions with a G-tetrad, our data suggest that the ability of the ligand to adjust its geometry to closely match that of the interacting G-tetrad is equally important. Telomestatin, the gold standard of a selective GQ ligand, has an extended, but not rigid, aromatic system that may explain its efficient $\pi-\pi$ overlap with a quadruplex.

We have previously described that NMM induced structural isomerization of Tel22 in low K^+ conditions from a mixed-hybrid to a parallel structure.[6] Here we showed that Tel22 imposed domed geometry on NMM. Combined, these data indicate that both Tel22 and NMM have the ability to change geometry/topology in order to take part in their very special interactions. Both NMM and Tel22 undergo "adaptive binding", a term frequently used in the aptamer field. It will

small yet critical structural differences will help uncover biological roles of various porphyrin compounds.

interaction between propionates and Tel22 GQ, structure and nonplanar deviations in NMM, MIX, and transition state of NMM, a UV−vis titration of NMM with Tel22 at pH 8.6, CD experiments that test ionic strength effect on NMM−Tel22, and coordinates for all calculated porphyrin structures. This material is available free of charge via the Internet at http://pubs.acs.org.

(6) Nicoludis, J. M.; Barrett, S. P.; Mergny, J. L.; Yatsunyk, L. A. *Nucleic Acids Res.* **2012**, *40*, 5432.
(7) Monchaud, D.; Teulade-Fichou, M. P. *Org. Biomol. Chem.* **2008**, *6*, 627.
(8) Georgiades, S. N.; Abd Karim, N. H.; Suntharalingam, K.; Vilar, R. *Angew. Chem., Int. Ed.* **2010**, *49*, 4020.
(9) Shi, D.-F.; Wheelhouse, R. T.; Sun, D.; Hurley, L. H. *J. Med. Chem.* **2001**, *44*, 4509.
(10) De Cian, A.; Lacroix, L.; Douarre, C.; Temime-Smaali, N.; Trentesaux, C.; Riou, J.-F.; Mergny, J.-L. *Biochimie* **2008**, *90*, 131.
(11) Liu, Y.; Lin, C.; Li, H.; Yan, H. *Angew. Chem., Int. Ed.* **2005**, *44*, 4333.
(12) Qin, H.; Ren, J.; Wang, J.; Wang, E. *Chem. Commun.* **2010**, *46*, 7385.
(13) Alzeer, J.; Vummidi, B. R.; Roth, P. J. C.; Luedtke, N. W. *Angew. Chem., Int. Ed.* **2009**, *48*, 9362.
(14) Lubitz, I.; Zikich, D.; Kotlyar, A. *Biochemistry* **2010**, *49*, 3567.
(15) Parkinson, G. N.; Cuenca, F.; Neidle, S. *J. Mol. Biol.* **2008**, *381*, 1145.
Moriarty, N. W.; Zwart, P. H.; Hung, L. W.; Read, R. J.; Adams, P. D. *Acta. Crystallogr., Sect. D: Biol. Crystallogr.* **2008**, *64*, 61.
(43) Langer, G.; Cohen, S. X.; Lamzin, V. S.; Perrakis, A. *Nat. Protoc.* **2008**, *3*, 1171.
(44) Feng, Z.; Chen, L.; Maddula, H.; Akcan, O.; Oughtred, R.; Berman, H. M.; Westbrook, J. *Bioinformatics* **2004**, *20*, 2153.
(45) Lecerof, D.; Fodje, M.; Hansson, A.; Hansson, M.; Al-Karadaghi, S. *J. Mol. Biol.* **2000**, *297*, 221.
(46) Phan, A. T.; Kuryavyi, V.; Gaw, H. Y.; Patel, D. J. *Nat. Chem. Biol.* **2005**, *1*, 167.
(47) De Cian, A.; Cristofari, G.; Reichenbach, P.; De Lemos, E.; Monchaud, D.; Teulade-Fichou, M.-P.; Shin-ya, K.; Lacroix, L.; Lingner, J.; Mergny, J.-L. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 17347.
(48) Hamon, F.; Largy, E.; Guédin-Beaurepaire, A.; Rouchon-Dagois, M.; Sidibe, A.; Monchaud, D.; Mergny, J. L.; Riou, J. F.; Nguyen, C. H.; Teulade-Fichou, M. P. *Angew. Chem., Int. Ed.* **2011**, *50*, 8745.
(49) Phillips, K.; Dauter, Z.; Murchie, A. I. H.; Lilley, D. M. J.; Luisi, B. *J. Mol. Biol.* **1997**, *273*, 171.

(50) Laughlan, G.; Murchie, A.; Norman, D.; Moore, M.; Moody, P.; Lilley, D.; Luisi, B. *Science* **1994**, *265*, 520.
(51) Lee, M. P. H.; Parkinson, G. N.; Hazel, P.; Neidle, S. *J. Am. Chem. Soc.* **2007**, *129*, 10106.
(52) Creze, C.; Rinaldi, B.; Haser, R.; Bouvet, P.; Gouet, P. *Acta Crystallogr., Sect. D* **2007**, *63*, 682.
(53) Ragazzon, P.; Chaires, J. B. *Methods* **2007**, *43*, 313.
(54) Ma, J.-G.; Zhang, J.; Franco, R.; Jia, S.-L.; Moura, I.; Moura, J. J. G.; Kroneck, P. M. H.; Shelnutt, J. A. *Biochemistry* **1998**, *37*, 12431.
(55) Jarzecki, A. A.; Spiro, T. G. *J. Phys. Chem. A* **2004**, *109*, 421.
(56) Karlberg, T.; Hansson, M. D.; Yengo, R. K.; Johansson, R.; Thorvaldsen, H. O.; Ferreira, G. C.; Hansson, M.; Al-Karadaghi, S. *J. Mol. Biol.* **2008**, *378*, 1074.
(57) Dai, J.; Carver, M.; Hurley, L. H.; Yang, D. *J. Am. Chem. Soc.* **2011**, *133*, 17673.
(58) A. Shelnutt, J.; Song, X.-Z.; Ma, J.-G.; Jia, S.-L.; Jentzen, W.; Medforth, C. *Chem. Soc. Rev.* **1998**, *27*, 31.

Supporting Information

Optimized end-stacking provides specificity of N-methyl mesoporphyrin IX for human telomeric G-quadruplex DNA

John M. Nicoludis[1], Stephen T. Miller[1], Philip D. Jeffrey[2], Steven P. Barrett[1], Paul R. Rablen[1], Thomas J. Lawton[3], Liliya A. Yatsunyk[1*]

[1]Department of Chemistry and Biochemistry, Swarthmore College, 500 College Ave., Swarthmore, PA 190813; [2]Department of Molecular Biology, Princeton University, Lewis Thomas Laboratory, Washington Road, Princeton, New Jersey 08544-1014; [3]Depts. of Molecular Biosciences and of Chemistry, Northwestern University, Evanston, IL 60208

Determination of extinction coefficient for NMME

A stock solution of NMME of precisely known concentration, 0.575 mM, was made in DMSO and used to prepare 18 samples that ranged in concentration from 1 to 50 μM. The UV-vis absorbances of these samples were measured and plotted *vs.* their concentrations yielding an extinction coefficient of 1.156×10^5 $M^{-1}cm^{-1}$ at 410.5 nm (Figure S1).

Concentration dependence of the melting transition studied through UV-vis and CD melting

Tel22 samples were prepared in the concentration range from 20 to 620 μM in 5K buffer (to mimic biochemical studies) and 50K buffer (to mimic crystallization conditions). Buffer compositions, 5K: 10 mM lithium cacodylate, 7.2, 5 mM KCl, 95 mM LiCl; and 50K: 10 mM lithium cacodylate, 7.2, 50 mM KCl, 50 mM LiCl. Samples were placed in cuvettes with 1.0, 0.2, or 0.1 cm pathlength depending on strand concentration and melted. UV-vis melting was monitored at 295 and 350 nm. The former wavelength is sensitive to the GQ folding state, while the latter wavelength was used as a reference to monitor instrument performance. The temperature was measured with the temperature sensor inserted in the cuvette holder right next to DNA samples. The temperature change rate was 0.2 °C min^{-1}, averaging time was 2.00 s, temperature interval was 1.0 °C, and spectral bandwidth was 2.0 nm. The value of signal at 350 nm was subtracted from each data set. In CD Tel22 in a potassium buffer displays maximum absorbance in 294 nm. Thus CD melting was monitored at 294 nm with 1.0 min equilibration

time, 1°C step, 10 s averaging time, and 1.0 nm bandwidth. CD data were collected on samples of Tel22 with concentration up to 150 μM, as samples with higher concentration had prohibitively high absorbance at 260 nm (leading to high Dyno voltage and noisy data). Melting and cooling data were superimposable. Therefore, melting data were analyzed assuming a two-state model with constant enthalpy, ΔH.[1] Starting and final baselines (assuming to be linear), melting temperature, T_m, and enthalpy of unfolding were adjusted to get the best fits. All reported T_m s were obtained from both melting and cooling curves.

Samples of Tel22 were also prepared in the presence of 2 equivalents of NMM. Melting of these samples was not possible using UV-vis because of NMM's interference with the 295 nm signal. In CD NMM-Tel22 complex displays maximum absorbance at 264 nm in a potassium buffer.[2] Thus melting of this complex using CD was not possible as DNA absorbs strongly at 260 nm interfering with the measurements.

Part of each sample prepared for melting was diluted in 50K buffer for CD wavelength measurements. Data were collected immediately after dilution with the assumption that CD signature observed corresponds to the Tel22 conformation in the original undiluted solution.

Influence of ionic strength on NMM-Tel22 interactions

The effect of ionic strength on NMM-Tel22 interaction was tested in CD wavelength and melting experiments. For these studies five sets of samples were prepared at 15, 60, 110, 210, and 410 mM ionic strengths. Each sample contained 2-5 μM Tel22 with or without 2 equivalents of NMM. The desired ionic strength was achieved by supplementing the samples with LiCl. CD wavelength data were collected on each sample. Then the samples were melted and the signals at 295 nm and 264 nm were monitored. Data processing was done as described above using Origin 8.1.

Additional details on crystallization conditions

Well conditions that produced diffraction quality crystals always contained 0.05 M lithium cacodylate buffer, pH 7.2, and 15 % PEG400. Other components required for crystal growth included ammonium acetate or sulfate at 0.40 - 0.70 M and KCl at 0.15 - 0.2 M. A lower concentration of KCl (0.05 M) supplemented with 0.01-0.10 M $CaCl_2$ or $MgCl_2$ also produced diffraction quality crystals.

Hydration sphere in NMM-Tel22 structure

There were 53 and seven well-defined water molecules in crystal forms 1 and 2, respectively (Figure S2). These water molecules are positioned along the four grooves and within the three loops of the GQ. Water molecules interact with phosphate groups and with N2 and N3 of the guanine bases that line the grooves. These groove-bound waters could in principle be replaced by ligand side chains. For example, methyl-piperazine side chains of BMSG-SH-3 and -4 form hydrogen bonds with phosphate backbone groups in quadruplex grooves leading to improved selectivity of these ligands.[3] We hypothesize that structure-guided chemical modification of NMM's propionates can lead to this type of quadruplex recognition and may improve binding affinity.

Measurement of the *N*-Me tilt angle

In this work, the tilt angle of the *N*-Me group was calculated by projecting the coordinates of the *N*-Me carbon onto a plane that is parallel to the mean porphyrin plane and intersects the coordinates of the nitrogen of the N-substituted pyrrole. The three sets of coordinates (N-*Me* carbon, nitrogen atom of the N-substituted pyrrole, and the projection point) make a right triangle which can be used to find the *N*-Me tilt angle through trigonometry. Other papers sometimes report tilt angle of the *N*-alkyl bearing pyrrole ring, which relates to the tilt angle of the *N*-Me group described here but is usually smaller.

Principal component analysis

Principal component analysis (PCA) was used to transform the coordinates of atoms in NMM and G-quartets to a standard three-dimensional axis. The first two components are equivalent to the projection of an object onto its least squares plane. This projection provides *x* and *y* coordinates of the object atoms centered at the point (0,0). The third component represents the deviations of the object atoms from its least squares plane, which are also known as the out-of-plane (*oop*) deviations. To create a line plot of the atoms *vs.* the *oop* deviations, the conversion of the *x* and *y* coordinates to a one-dimensional variable is needed. Thus, *x* and *y* were converted to an angle (θ) such that $\theta = arctan(y/x)$. The resulting angles of the object were translated (not up to scale) in order to offset the G-quartets according to the right-hand helicity of the G-quadruplex.

Calculation of the energy required to distort the porphyrin macrocycle

A fully relaxed structure for NMM was generated using coordinates for NMM taken from crystal form 1, then adding hydrogen atoms where appropriate, and finally conducting a B3LYP/6-31G(d) geometry optimization. A corresponding partially frozen structure was obtained by forcing the two dihedral angles Φ_1 and Φ_2 (see figure below) to remain at the values obtained from the crystal structure (Φ_1= 29.1° and Φ_2 = -20.8°). Replacing the *N*-Me group of the NMM structure with a hydrogen atom, followed by B3LYP/6-31G(d) geometry optimization, yielded the fully relaxed MIX structure. Finally, the partially frozen MIX structure was obtained by forcing the two dihedral angles Φ_1 and Φ_2 to take on the values in the NMM crystal structure, while allowing all other geometric variables to vary freely. The results of these calculations are shown in Table S1.

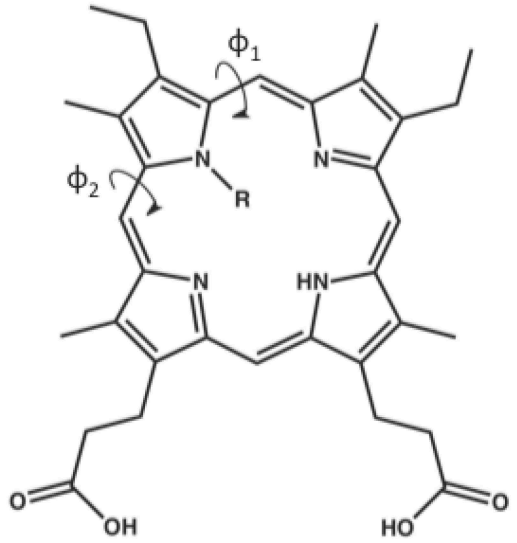


Table S1: Results of the energy calculations for NMM and MIX

Ligand	Structure	**ΔE (water)**, kcal mol^{-1}
NMM	Relaxed to Φ_1= 24.0° and Φ_2 = -22.4°	0.0
NMM	Frozen at Φ_1 = 29.1° and Φ_2 = -20.8°	0.7
MIX	Relaxed to Φ_1= 0.7° and Φ_2 = -0.4°	0.0
MIX	Frozen at Φ_1 = 29.1° and Φ_2 = -20.8°	4.3

Variations in Φ_1 and Φ_2 were chosen because these angles uniquely define the bend of the most distorted pyrrole ring; the dihedral angles Φ_1 and Φ_2 connect the most distorted pyrrole ring, which bears the *N*-Me group in NMM, to the rest of the porphyrin. The data indicate that 0.7 kcal mol^{-1} of energy is

required to distort NMM from its relaxed geometry to that found in crystal form 1. A significantly higher amount of energy, 4.3 kcal mol^{-1}, is required for similar distortion of MIX. All calculations were carried out using the Gaussian 09 package.[4]

Calculation of the transition state for NMM racemization

NMM exists as a mixture of four regioisomers, each of which has a pair of enantiomers, the latter differing in whether the N-Me group protrudes from one face of the macrocycle or the other. We employed density functional calculations to locate the transition structure for racemization of NMM (isomer with the N-Me group attached to ring A, Figure 1) and to compute the corresponding free energy barrier. The racemization process requires the N-methyl group to squeeze through the middle of the porphyrin ring leading to a significant non-planarity of the macrocycle (D_{oop} = 1.773). NSD analysis indicates that the transition structure is mainly doomed (61.9 %) and saddled (33.5%), with little contribution from other components, Table S2 and Figure S7c, S7f. To improve the quality of the energy calculation, we used B2PLYPD/6-311+G(d,p)m single point calculations, at the B3LYP/6-31G(d) optimized geometries and with B3LYP/6-31G(d) thermodynamic corrections. The calculated free energy barrier is 52.3 kcal mol^{-1} (at 298 K) in the gas phase, or 54.3 kcal mol^{-1} in the presence of a simulated water solvent, indicating that racemization is essentially impossible at room temperature, and would require temperatures in excess of 500 K to occur on a timescale shorter than decades.

Table S2: Principal components of porphyrin non-planarity determined using NSD Engine

Ligand	D_{oop} (Å)	*sad* (B_{2u})	*ruf* (B_{1u})	*dom* (A_{2u})	*wav(x)* ($E_{g(x)}$)	*wav(y)* ($E_{g(y)}$)	*pro* (A_{1u})	$\Sigma(\mathbf{d}_{oop})$
NMM 1	0.727	0.289	0.150	0.499	0.408	0.084	0.014	1.444
NMM 2	0.747	0.109	0.231	0.617	0.317	0.104	0.033	1.410
NMM_{sol}	1.068	0.782	0.004	0.410	0.600	0.008	0.005	1.809
NMM_{gas}	1.044	0.717	0.067	0.436	0.618	0.024	0.009	1.870
NMM_{2Q3J}	1.839	1.428	0.542	0.327	0.240	0.936	0.093	3.566
NMM_{1C1H}	1.581	0.998	0.622	0.720	0.769	0.079	0.030	3.218
MIX_{frozen}	0.728	0.289	0.150	0.500	0.408	0.084	0.014	1.445
MIX_{sol}	0.046	0.044	0.005	0.008	0.007	0.006	0.004	0.074
MIX_{gas}	0.050	0.042	0.018	0.012	0.012	0.009	0.006	0.099
NMM_{TS}	1.773	0.842	0.028	1.558	0.031	0.056	0.001	2.517
$TMPyP4_{2HRI}$	1.238	0.145	1.104	0.403	0.107	0.340	0.044	2.143

* *Note, the % deviation for each component can be calculated by dividing the value of the coefficient for each component by $\Sigma(d_{oop})$ and multiplying the result by 100%*
* *TS – transition state*
* *NSD in the case of TMPyP4 was done on the porphyrin stacked onto the A·T base pair, and not the one that is perpendicular to GQ axis*

Figure S1

Figure S1. Beer Law plot for *N*-methylmesoporphyrin IX dimethyl ester (NMME) in DMSO. The extinction coefficient determined from the slope of the line is 1.156×10^5 M^{-1} cm^{-1} at 410.5 nm. Data indicates no aggregation.

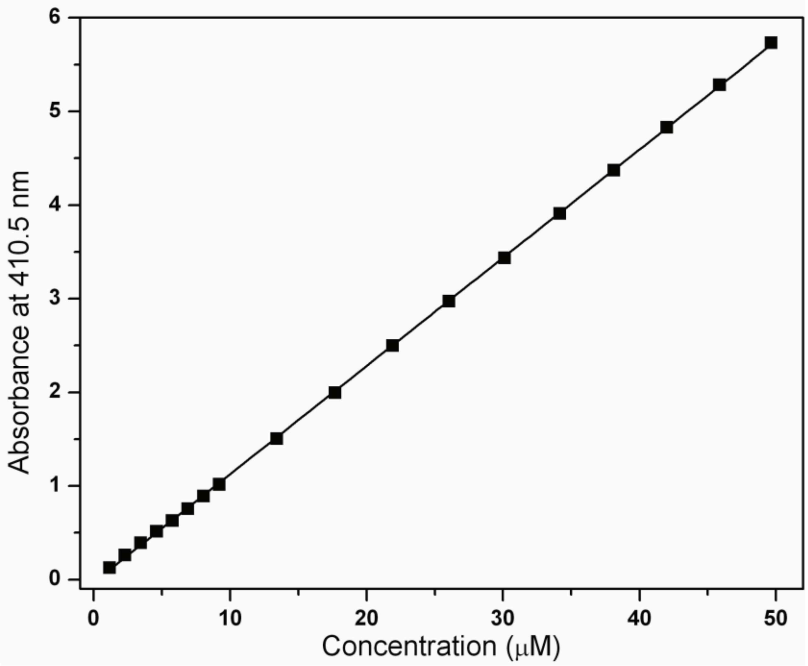


Figure S2

Figure S2. Hydrogen-bonding network of water molecules in the Tel22-NMM structure. Water molecules are shown as red spheres. Hydrogen bonds are shown for lengths between 3.6 Å at a 0° angle to 3.2 Å at a 63° angle. Water molecules are mainly found in the grooves and within the propeller loops and are hydrogen bonded to phosphate OP1 as well as guanines N2 and N3 atoms.

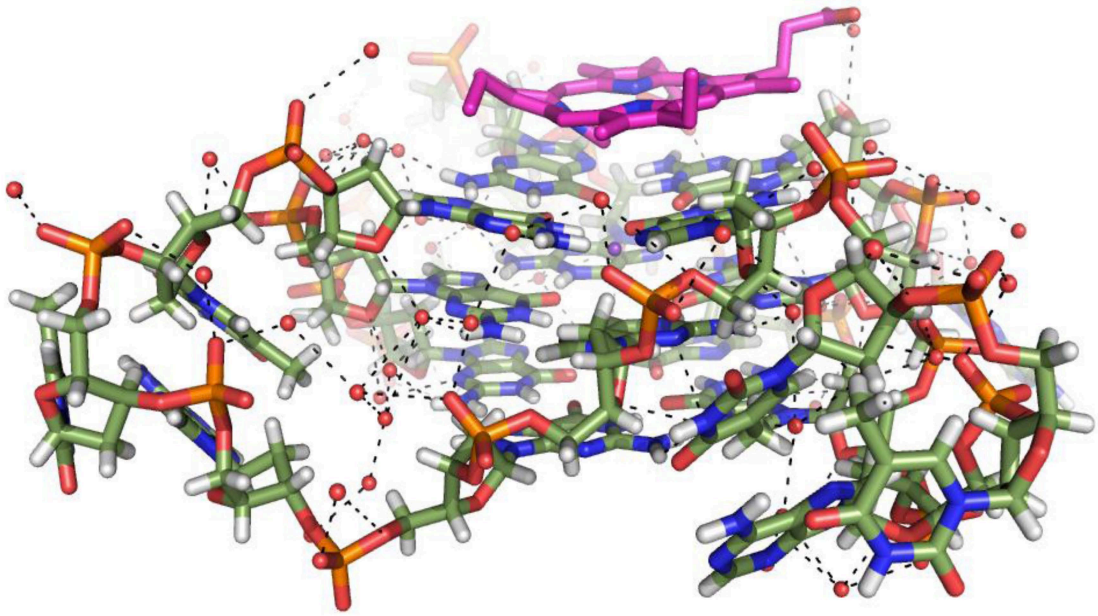


Figure S3

Figure S3. Concentration dependence of melting transition for Tel22 determined in UV-vis and CD experiments by monitoring signal at 295 nm. (a) UV-vis melting of 209 to 624 μM Tel22 samples in 5K buffer. Data demonstrates no dependence of melting temperature on Tel22 strand concentration. (b) Reversible CD melting of 148 μM Tel22 in 50K buffer. (c) CD wavelength scans for Tel22 samples in 50K buffer as function of concentration. (d) CD wavelength scans for the NMM-Tel22 samples in 50K buffer as function of concentration.

Melting and cooling curves for Tel22 samples at concentrations of 20, 40, 100, 150, 200, 300, 450, 500, and 620 μM monitored by CD and UV-vis at 295 nm were all fit as described in experimental section yielding reproducible results. Because melting of Tel22 was reversible we were able to extract thermodynamic parameters, such as enthalpy of unfolding, ΔH, free energy of unfolding, ΔG, and T_m. In 50K buffer Tel22 displays T_m of 336.3 ± 1.0 and ΔH of 192 ± 10 kJ mol^{-1}; in 5K buffer Tel22 displays T_m of 323.8 ± 1.2 and ΔH of 173 ± 11 kJ mol^{-1}. The free energy of unfolding at 298 K increases with increasing cation concentration from 14 kJ mol^{-1} in 5 mM K^+ solution to 22 kJ mol^{-1} in 50 mM K^+ solution. This finding is in agreement with the suggestion by Chaires that folding of Tel22 (and GQs in general) is driven by cation binding.[5]

The independence of T_m on Tel22 concentration and reversibility of the melting process indicate unequivocally that Tel22 is *monomolecular* in the studied concentration range, up to 0.62 mM. This conclusion is consistent with the results of Vorlickova's laboratory that observed no changes in the CD spectrum of Tel22 up to 1.8 mM in 10 mM potassium phosphate, 7.2 buffer with 150 mM KCl.[6] The crystals of Tel22 were grown from the solution containing 0.50 mM DNA.[7] According to above results, at this concentration Tel22 is expected to be monomeric, yet, dimer was observed in the crystal structure.[7] This dimer formation, then, is a result of the crystal packing forces and most likely bares no biological significance.

In our study, crystals of NMM-Tel22 complex were grown from solution containing 0.69-0.80 mM Tel22 and showed the presence of 5′-5′ dimers of Tel22. While experiments with the NMM-Tel22 complex, similar to the one presented here for Tel22, proved to be much more challenging due to porphyrin absorbance at 295 nm that coincides with DNA absorbance in UV-vis, we can extrapolate our conclusion from the Tel22 system to the NMM-Tel22 system and suggest that the dimer is due, most likely, to crystal packing forces and is not biologically significant. In support of this conclusion, we do

not observe significant difference in the CD signature of Tel22 alone or of the NMM-Tel22 complex as function of concentration in the studied concentration range (Figure S3c and S3d).

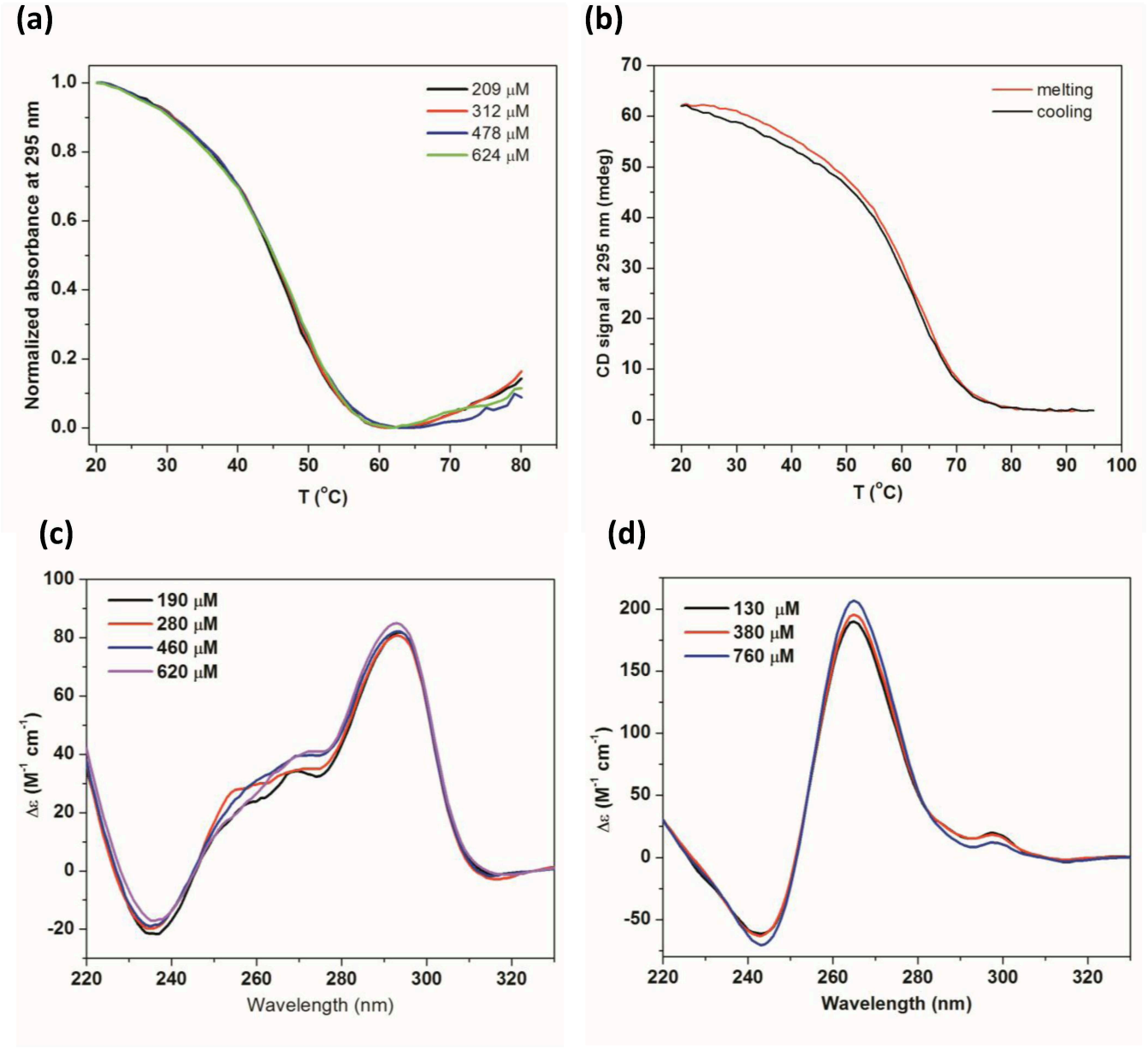


Figure S4. Comparison of G-quartets in a variety of the reported Tel22 structures. (a) Overlay of Tel22 crystal structures (see Figure 4 for color coding). All structures have similar positions of the K^+ ions (spheres) in the GQ central core. (b) Overlay of the top, ligand-facing G-quartets of the Tel22 structures shown in (a). (c) Representation of the O6-O6 interatomic distances and Hoogsteen hydrogen-bonding distances for N1-O6 and N2-N7. (d) Average interatomic O6-O6 distances for all G-tetrads and (e) Hoogsteen hydrogen-bonding distances measured for the top G-quartet for the set of Tel22 structures shown in (a).

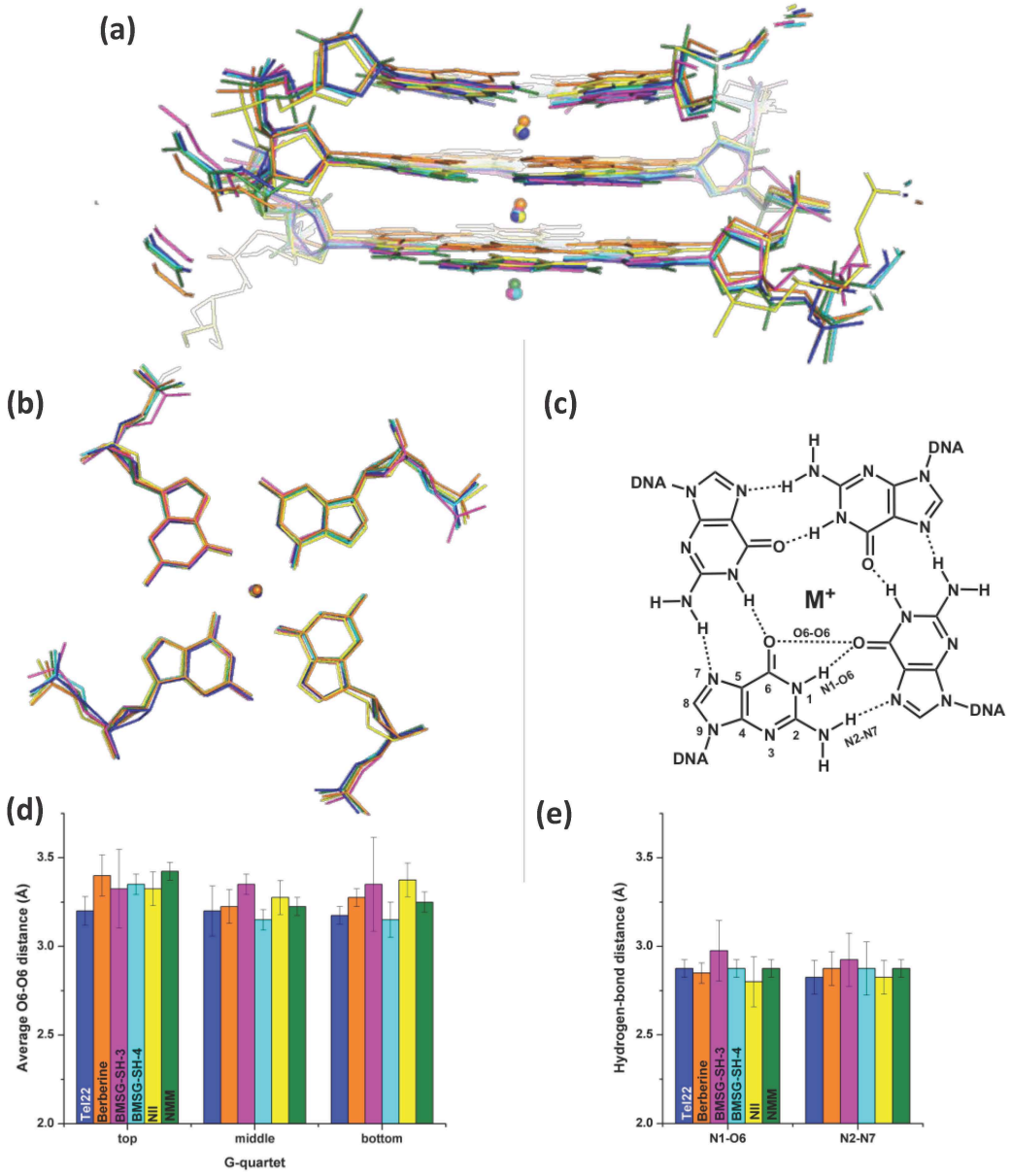


Figure S5. NMM and TMPyP4 binding to human telomeric DNA G-quadruplex. (a) NMM-Tel22 complex and (b) TMPyP4-$(dTAG_3TTAG_3)_2$ complex.[8] The binding of NMM and TMPyP4 to human GQ is vastly different. NMM (magenta) stacks directly onto the 3′ G-tetrad of Tel22 in a 1:1 ratio and interacts with a reverse A-T base pair on its other face, Figure 3c. TMPyP4 (orange) binds to a bimolecular GQ, $(dTAG_3TTAG_3)_2$ in a 1:2 ratio *via* two modes: it stacks onto a reverse Watson-Crick T1·A8 base pair. This base pair is stacked onto an A-A-T triplex, which interacts directly with a terminal G-tetrad. The other side of this ligand is engaged in identical interactions. The second TMPyP4 molecule is oriented perpendicular to GQ axis and binds to loops *via* four T bases, two on each side of the ligand. G is shown in green, A in red and T in blue. As for the geometry, NMM is predominantly doomed, while TMPyP4 is mainly ruffled, Table S2.

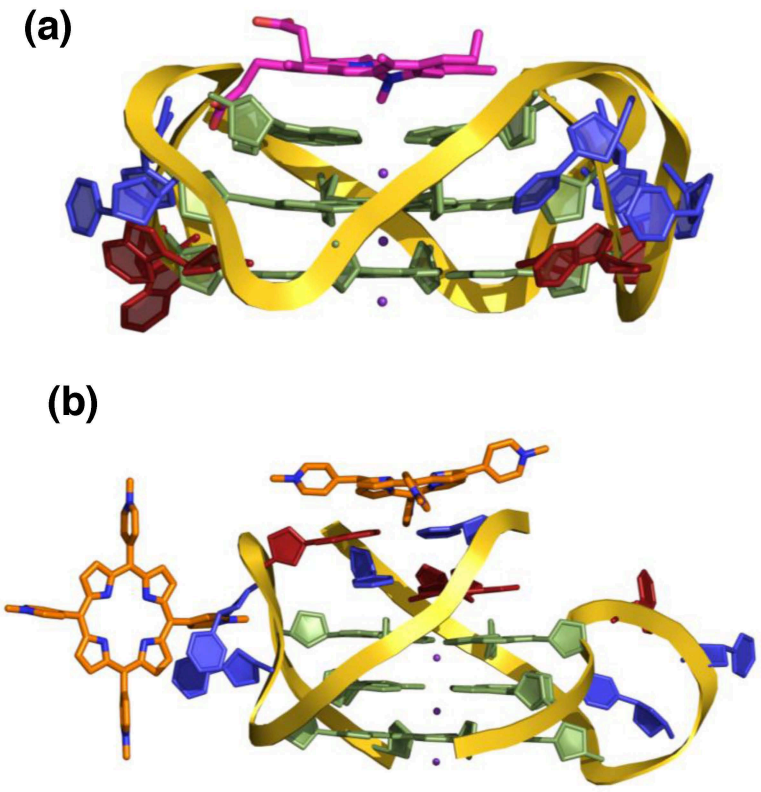


Figure S6

Figure S6. Details of **interactions between NMM's** propionic side chains and Tel22. In the observed electron density there is some evidence for the location of the propionate groups suggesting binding (or at least higher population) of a single isomer. Both propionates are likely involved in hydrogen bonding with the quadruplex backbone. (a) The first propionate may participate in a hydrogen bonding with two backbone phosphates (in one case through a water molecule), and, thus, may be protonated under crystallization conditions. (b) The second propionate may interact with a 3′ sugar oxygen through a water molecule. All distance are shown in Å; electron density in (a) is shown at 1σ, and that in (b) is shown at 0.5σ.

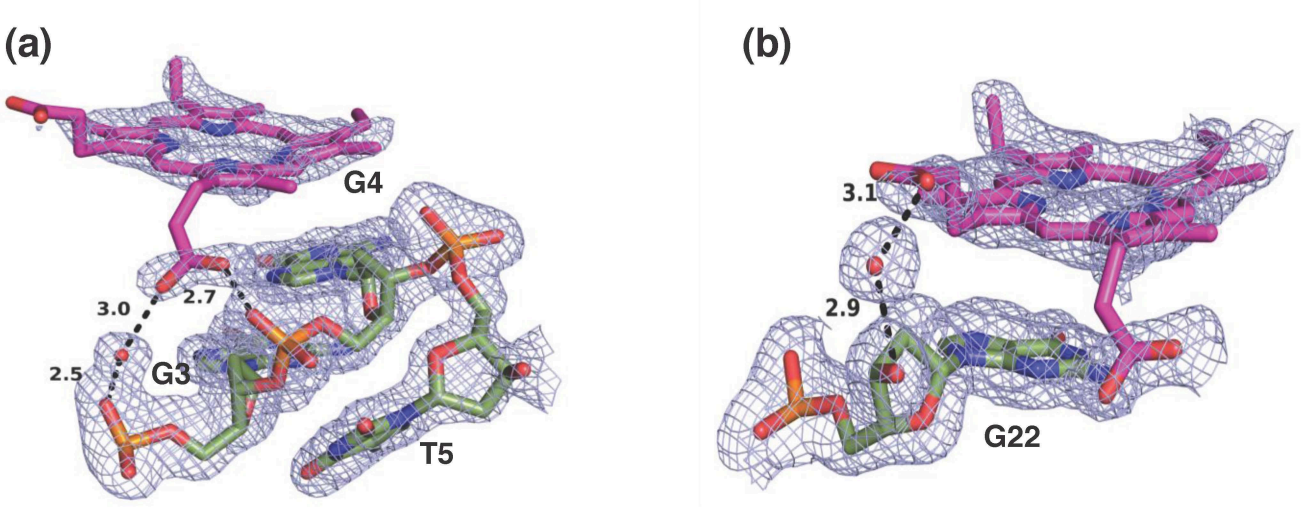


Figure S7

Figure S7. Structure and deviations for NMM and MIX and transition state of NMM. Energy-minimized structure of (a) NMM and (b) MIX in solution state. (c) Geometry of transition state of NMM. (d) NSD results for energy minimized NMM structure in the gas phase (red) and in solution phase (black). Note the great similarity between the two structures: the overall distortion, D_{oop}, is 1.07 and 1.04 Å; domed deviation from planarity is 22.7% and 23.3%, and saddled deviation is 43.2% and 38.3% for solution and gas phase, respectively. In the discussion solution phase NMM structure is used, unless otherwise noted. (e) NSD results for NMM from wild-type (PDB ID 1C1H) and H138A mutant (PDB ID 2Q3J) of ferrochelatase. The NMM structure bound to ferrochelatase is significantly more distorted than that bound to Tel22; its *N*-Me group forms larger angle with the average porphyrin plane, 68.5° *vs* 44.8 ° for NMM complexed with Tel22. (f) NSD result for transition state of NMM in comparison NMM in solution state. NSD results display the six lowest-frequency normal deformations.

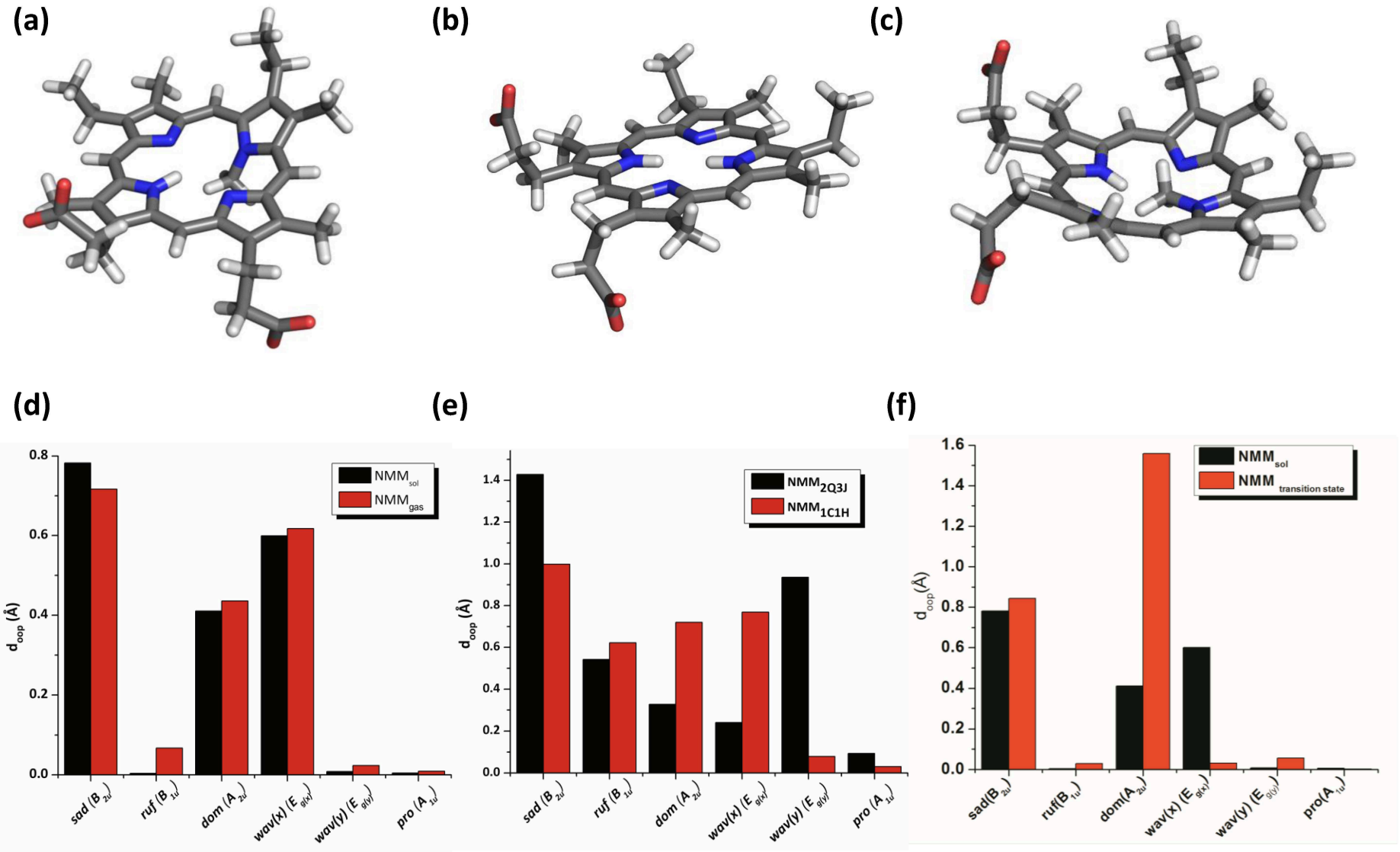

(a)
(b)
(c)
(d)
(e)
(f)
NMM sol
NMM gas
NMM 2Q3J
NMM 1C1H
NMM transition state
d oop (Å)
sad (B 2u)
ruf (B 1u)
dom (A 2u)
wav(x) (E g(x))
wav(y) (E g(y))
pro (A 1u)
sad(B 2u)
ruf(B 1u)
dom(A 2u)
wav(x) (E g(x))
wav(y) (E g(y))
pro(A 1u)

Figure S8

Figure S8. UV-vis titration of NMM with Tel22 at pH 8.6 (a) Representative UV-vis absorption spectra of 11.1 μM NMM titrated with 0.84 mM Tel22 in 10 mM TrisHCl buffer, pH 8.6, 5 mM KCl, and 95 mM LiCl. Data were collected in 2 mm quartz cuvette. The final [GQ]/[NMM] ratio was 16.1. (b) A plot of NMM's absorbances at 379 and 399.5 nm as a function of added Tel22. Solid line represents fit to 1:1 binding model, while the dashed lines represent 95 % confidence interval. Note, similar experiments were performed at pH = 5.8 using the following buffer: 10 mM lithium cacodylate, pH 5.8, 5 mM KCl, and 95 mM LiCl.

Binding constants for NMM-Tel22 are $(1.2 \pm 0.1)\times10^5$, $(1.0 \pm 0.3)\times10^5$, and $(0.3 \pm 0.1)\times10^5$ M^{-1} at pH of 5.8, 7.2 and 8.6, respectively. Note, small decrease in binding affinity at pH 8.6 could be explained either by the change in the protonation state of NMM or by the difference in buffer compositions – lithium cacodylate buffer was used to achieve pH of 5.8 and 7.2, but TrisHCl buffer was used to achieve pH of 8.6. In the former case, increase in pH leads to higher deprotonation state of NMM propionates and, thus, results in decrease biding between negatively charged DNA and NMM.

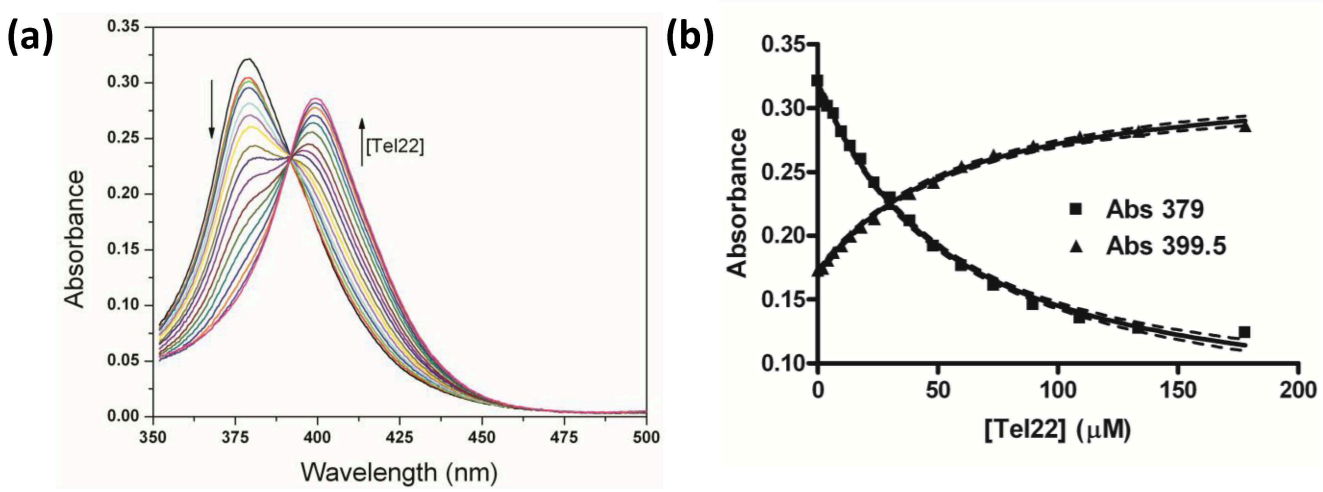


Figure S9

Figure S9. Influence of ionic strength on NMM-Tel22 interaction. (a) Annealing of 5.0 μM Tel22 alone in 10 mM lithium cacodylate buffer, pH 7.2, with 5 mM KCl and the amount of LiCl required to bring the total ionic strength to 15, 60, 110, 210, and 410 mM. (b) Annealing of 5.0 μM of Tel22 with 10.0 μM NMM under conditions described in (a). (c) Stabilization of 5.0 μM Tel22 by 10.0 μM NMM under conditions described in (a). Stabilization temperature was determined *via* CD melting experiments.

The CD signal for Tel22 alone decreased to some extent at 270 nm and 292 nm with concurrent shift of the peak from 292 to 294 nm upon increase in the ionic strength from 15 to 210 mM. In the presence of NMM under the same conditions, the CD signal at 264 nm increased, while that at 294 nm decreased – both changes are consistent with an increase in the parallel character of the Tel22-NMM complex. This data, together with a small red shift in the UV-vis spectra of NMM-Tel22 mixtures (not shown), suggest that the efficiency of NMM-induced conformational change in Tel22 (from a hybrid to a parallel conformation) increases with increasing ionic strength.

The stability of Tel22 alone or in the presence of NMM increased slightly when ionic strength was increased to 210 mM (Table S3), but the stabilization temperature $\Delta T_{1/2}$ (the difference between melting temperature of Tel22 in complex with NMM and that of Tel22 alone) remained practically unchanged, with a small decrease at high ionic strength (Figure S7c). Therefore, the strength of NMM-Tel22 interaction is not highly sensitive to ionic strength, suggesting a predominantly non-electrostatic nature of this interaction.

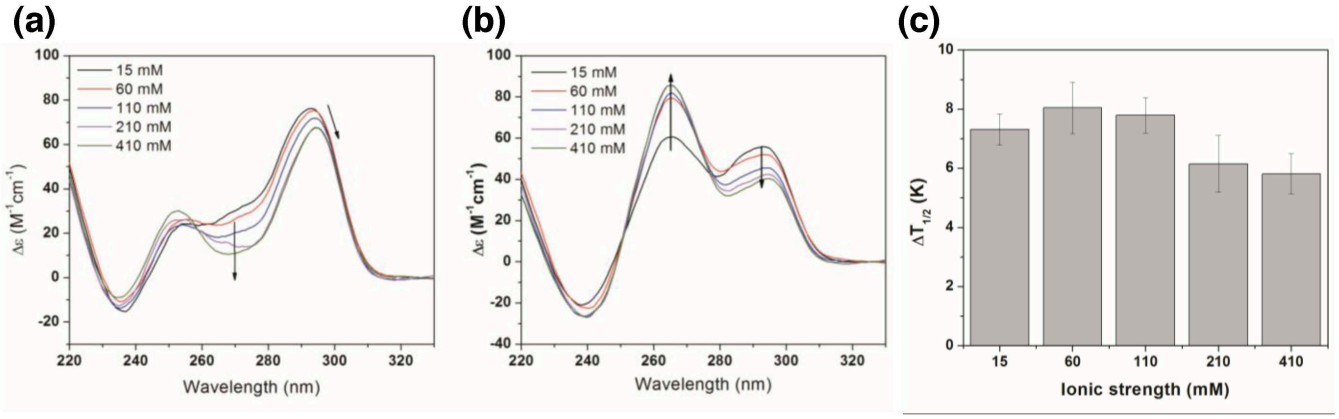


Table S3. Data for a representative experiment that tested the influence of ionic strength on NMM-Tel22 interaction

Ionic strength, mM	$T_{1/2}$, °C for Tel22	$T_{1/2}$, °C for Tel22+NMM	$\Delta T_{1/2}$, °C
15	45.0 ± 0.4	52.3 ± 0.3	7.3 ± 0.5
60	46.1 ± 0.8	54.1 ± 0.4	8.0 ± 0.9
110	46.8 ± 0.5	54.5 ± 0.3	7.8 ± 0.6
210	48.8 ± 0.9	54.9 ± 0.4	6.2 ± 1.0
410	47.6 ± 0.6	53.4 ± 0.4	5.8 ± 0.7

References

(1) Ramsay, G. D.; Eftink, M. R. *Methods Enzymol.* 1994, *240*, 615.
(2) Nicoludis, J. M.; Barrett, S. P.; Mergny, J.-L.; Yatsunyk, L. A. *Nucleic Acids Res.* 2012, *40*, 5432.
(3) Collie, G. W.; Promontorio, R.; Hampel, S. M.; Micco, M.; Neidle, S.; Parkinson, G. N. *J. Am. Chem. Soc.* 2012, *134*, 2723.
(4) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; ; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Wallingford CT, 2009; Vol. Revision A.02.
(5) Gray, R. D.; Chaires, J. B. *Biophysical Chemistry* 2011, *159*, 205.
(6) Renciuk, D.; Kejnovska, I.; Skolakova, P.; Bednarova, K.; Motlova, J.; Vorlickova, M. *Nucleic Acids Res.* 2009, *37*, 6625.
(7) Parkinson, G. N.; Lee, M. P. H.; Neidle, S. *Nature* 2002, *417*, 876.
(8) Parkinson, G. N.; Ghosh, R.; Neidle, S. *Biochemistry* 2007, *46*, 2390.