FRET-MC: a fluorescence melting competition assay for studying G4 structures in vitro

Yu Luo, Anton Granzhan, Daniela Verga, Jean-Louis Mergny

FRET-MC: a fluorescence melting competition assay for studying G4 structures *in vitro*

Yu Luo[1,2], Anton Granzhan[1], Daniela Verga[1*] & Jean-Louis Mergny[2*]

1. Université Paris Saclay, CNRS UMR9187, INSERM U1196, Institut Curie, 91400 Orsay, France.
2. Laboratoire d'Optique et Biosciences, Ecole Polytechnique, CNRS, Inserm, Institut Polytechnique de Paris, 91128 Palaiseau, France.

* Authors to whom correspondence may be addressed: daniela.verga@curie.fr; jean-louis.mergny@inserm.fr

Revised version, December 10, 2020

Abstract

G-quadruplexes (G4) play crucial roles in biology, analytical chemistry and nanotechnology. The stability of G4 structures is impacted by the number of G-quartets, the length and positions of loops, flanking motifs, as well as additional structural elements such as bulges, capping base pairs, or triads. Algorithms such as G4Hunter or Quadparser may predict if a given sequence is G4-prone by calculating a quadruplex propensity score; however, experimental validation is still required. We previously demonstrated that this validation is not always straightforward, and that a combination of techniques is often required to unambiguously establish whether a sequence forms a G-quadruplex or not. In this article, we adapted the well-known FRET-melting assay to characterize G4 in batch, where the sequence to be tested is added, as an unlabeled competitor, to a system composed of a dual-labeled probe (F21T) and a specific quadruplex ligand. PhenDC3 was preferred over TMPyP4 because of its better selectivity for G-quadruplexes. In this so-called FRET-MC (melting competition) assay, G4-forming competitors lead to a marked decrease of the ligand-induced stabilization effect (ΔT_m), while non-specific competitors (*e.g.*, single- or double-stranded sequences) have little effect. Sixty-five known sequences with different typical secondary structures were used to validate the assay, which was subsequently employed to assess eight novel sequences that were not previously characterized.

Keywords: G-quadruplex; FRET-melting; UV-melting; G-quartet; DNA structure

1. Introduction

G-quadruplexes (G4) are four-stranded nucleic acid structures adopted by G-rich DNA and RNA sequences. G4 result from the stacking of two or more G-quartets (also called G-tetrads), which are formed by four guanine bases interacting through Hoogsteen hydrogen bonds [1]. G4 structures have been widely used to design biosensors to detect specific small molecules [2-5] and to control the assembly of supramolecular DNA complexes [6-9]. G4 structures also exist *in vivo* and play important roles in cells [10, 11]. For example, G4s can be formed at human telomeres and G4 ligands may interfere with telomeric functions, leading to telomere shortening and/or uncapping [12-14]. G4s are also found in the promoter regions of genes critical in cancer, including *KRAS*, *BCL2* and *VEGF* [15-17]: BCL2 plays an essential role in cell survival; VEGF is a key angiogenic growth factor which contributes to angiogenesis and tumor progression, and KRAS is one of the most frequently mutated oncogenes in many signal transduction pathways, relevant for different types of human carcinomas [16].

Given the importance of G4 structures in biology and nanotechnology [18], algorithms such as G4Hunter [19, 20] have been developed to predict if a specific sequence is G4-prone. However, for most DNA or RNA motifs, experimental validation is required and, for this purpose, a number of biophysical methods have been developed to characterize G4 structures *in vitro*. Some classical methods are based on the physical properties of G4 structures, such as UV-melting at 295 nm [21], nuclear magnetic resonance (NMR) [22], circular dichroism (CD) spectroscopy [23], isothermal difference spectroscopy (IDS) and thermal difference spectroscopy (TDS) [24]. Fluorescence light-up assays employing dyes such as Thioflavin T [25], *N*-methylmesoporphyrin IX (NMM) [26], tailor-made dyes [27] or combinations of dyes [28, 29] are also used to evidence G4 formation. Although there is a wide range of choices to check if a quadruplex is formed or not, this validation is not always straightforward; therefore, a combination of techniques is often required to unambiguously establish whether a sequence folds into a quadruplex or not [19].

High-affinity G4 ligands can stabilize a G4 structure and alter its biological functions [30]. Typical assays used to characterize these ligands are the Fluorescent Intercalator Displacement (FID) assay [31] and the fluorescence-based Förster Resonance Energy Transfer (FRET)-melting assay

[32, 33]. The FRET-melting assay is based on the stabilization induced by a quadruplex ligand, leading to a difference in melting temperature (Tm) between the nucleic acid alone and in presence of this ligand [34]: in the presence of the latter, the thermal stability of the structure increases in a concentration-dependent manner. This FRET-melting assay has been extensively used to estimate whether a compound is a good quadruplex ligand or not [34], with some biases when ranking ligands potency using melting experiments [35]. More recently, this assay was adapted to assess G4 ligands in near-physiological conditions [36].

In this report, instead of testing unknown compounds, we make use of one of the most-characterized G4 ligands, PhenDC3 [37], to determine if an unknown DNA sequence forms a G-quadruplex structure. PhenDC3 is a high affinity G4 ligand capable of binding to a variety of G4 structures, but with a low affinity for other conformations [38, 39]. We took advantage of these observations to design a novel FRET-melting competition assay, termed FRET-MC, in which the interaction between a fluorescent G4-forming oligonucleotide and PhenDC3 is challenged by the (unlabeled) sequence of interest added in excess. Sixty-five sequences with a known structure were tested to validate this FRET-melting competition assay, which allowed us to determine whether a sequence forms a stable quadruplex or not. Finally, eight novel sequences were used to determine if this assay was accurate: the conclusions reached by this technique were supported by other biophysical methods (CD, IDS and TDS). Advantages and drawbacks of the FRET-MC method are discussed.

2. Materials and Methods

2.1 Samples

Oligonucleotides were purchased from Eurogentec, Belgium, as dried samples: unmodified oligonucleotides were purified by RP cartridge while a dual-labeled F21T was purified by RP-HPLC. Stock solutions were prepared at 100 µM strand concentration for the unlabeled oligonucleotide and at 200 µM strand concentration for F21T in ddH_2O. 50 µM oligonucleotide solutions were annealed in the corresponding buffer, kept at 95 ℃ for 5 min and slowly cooled to room temperature before measurement. The FRET buffer contains 10 mM KCl, 90 mM LiCl, 10 mM lithium cacodylate, pH 7.2. The K-100 buffer contains 100 mM KCl, 10 mM lithium cacodylate, pH 7.2.

2.2 FRET-melting competition assay

FRET melting experiments were performed in 96-well plates using a HT7900 RT-PCR instrument (Applied BioSystem). Each well contained competitors either at a single concentration (3 µM), or at 6 different concentrations, ranging from 0.2 to 3 µM. 0.2 µM of the fluorescent oligonucleotide F21T was incubated with or without 0.4 µM G4 ligand (PhenDC3 for most experiments; TMPyP4 was used in a control experiment shown in supplementary information) in FRET buffer in a final volume of 25 µL. The FAM channel was used to collect the fluorescence signal. Samples were kept at 25 °C for 5 min, then the temperature was increased by 0.5 °C per minute until 95 °C. Each experimental condition was tested in duplicate on at least two separate plates.

ΔTm is determined as the difference in Tm with the sample containing F21T in the absence of PhenDC3. The Tm of an oligonucleotide is defined as the temperature at which 50% of the oligonucleotide is unfolded. The most common method to obtain Tm values is approximated as corresponding to half of the height at the normalized melting curve [40].

The FRET-melting assay uses a 96-well plate as a sample holder, which allows to process 48 sequences simultaneously. The traditional ‘midpoint’ determination requires a manual analysis [41]. Given the number of profiles to be analyzed, this process is time consuming as the curves

are analyzed one by one. The DoseResp Function in Origin Pro package allows collect Tm in batch. As shown in **Fig. S1A**, Log X_0 in fitting curve is the Tm. This method can only be used in curve *(i)* (see results) as non-linear fitting would fail if there is no high plateau at row curve. In general, the difference of Tm calculated by these two methods is very small: when using both the '1/2 height' and non-linear fitting methods to calculate Tm of F21T alone, the results are 59.8℃ and 59.4℃, respectively (**Fig. S1B**). In a few rare cases that should be noted (**Fig. S1C**), curves are irregular, and the minimum of X-axis (usually 25℃ in FRET-melting assay) does not correspond to y=0 for the normalized curve. In this case, using '1/2 height' is more accurate; Tm calculated by DoseResp always referenced the X value corresponding to Y=0.5, while in some curves the 1/2 height ($Y_{x=95}$ - $Y_{x=25}$) at Y-axis is not always 0.5. In the instance, 1/2 height at Y-axis is 0.525, Tm calculated by '1/2 height' is 53.0℃, to be compared with 51.8℃ by DoseResp.

DoseResp Function: $$Y = A_1 + \frac{A_2 - A_1}{1 + 10^{(\log x_0 - x)p}}$$

In brief, DoseResp is only truly accurate for curves with appropriate upper and lower baselines, while the '1/2 height' approach can be applied in all cases, provided that normalization is accurate. If possible, we suggest to use the same method to calculate Tm for experiments performed in parallel, although differences of Tm determined by these two approaches is small in general.

2.3 UV-melting assay

UV-melting curves of 5 μM oligonucleotides in FRET buffer were recorded with a Cary 300 (Agilent Technologies, France) spectrophotometer. Heating runs were performed between 10°C and 95 °C, the temperature was increased by 0.2 °C per minute, and absorbance was recorded at 260 nm and 295 nm. Tm was determined as the temperature corresponding to half of the height of the normalized melting curve.

2.4 Circular dichroism

3 µM SP-PGQ-1 was kept in 1000 µL FRET buffer. 3 µM Oligonucleotides of testing set were kept in 1000 µL K-100 buffer. CD spectra were recorded on a JASCO J-1500 (France) spectropolarimeter at room temperature, using a scan range of 400–230 nm, a scan rate of 200 nm/min and averaging four accumulations.

2.5 Thermal difference spectra

3 µM SP-PGQ-1 was kept in FRET buffer. 3 µM Oligonucleotides of testing set were kept in 1000 µL K-100 buffer. Absorbance spectra were recorded on a Cary 300 (Agilent Technologies, France) spectrophotometer at 25 °C (scan range: 500-200 nm; scan rate: 600 nm/min; automatic baseline correction). After recording the first spectra (folded), temperature was increased to 95 °C, and the second UV-absorbance spectra was recorded after 15 min of equilibration at high temperature. TDS corresponds to the arithmetic difference between the initial (folded; 25 °C) and second (unfolded; 95 °C) spectra.

2.6 Isothermal difference spectra

3 µM Oligonucleotides of testing set were kept in 900 µL K-100 buffer. Absorbance spectra were recorded on a Cary 300 (Agilent Technologies, France) spectrophotometer at 25 °C (scan range: 500-200 nm; scan rate: 600 nm/min; automatic baseline correction). 100 µL of 1M KCl was added after recording the first spectrum, and the second UV-absorbance spectrum was recorded after 15 min of equilibration. IDS correspond to the arithmetic difference between the initial (unfolded) and second (folded, thanks to the addition of K^+) spectra, after correction for dilution.

3. Results and Discussion

3.1 Principle of the FRET-MC assay

In a K^+-containing buffer, F21T forms a stable G4 structure which can be used as a FRET-melting probe thanks to the fluorophores attached at both ends: a 5'-appended donor (Fluorescein) and a 3'-appended acceptor (TAMRA), allowing efficient energy transfer between the donor and acceptor dyes when the oligonucleotide is folded **(Fig. 1)**. Thermal unfolding leads to the disruption of the G-quadruplex structure and a decrease in FRET efficiency, as the 5' and 3' ends become distant when the sequence is single-stranded. In the presence of a G4 ligand, the melting temperature (T_m) of F21T increases, as the ligand makes the structure of F21T more stable. F21T is typically used in a FRET-melting assay, in which the specificity of a ligand is tested by adding various specific (G4-forming) and unspecific (duplexes or single-stranded) competitors [33]. In this report, we are radically changing our viewpoint: rather than testing ligands of unknown specificity against known competitors, we are investigating one of the best-characterized G4-ligand, PhenDC3, against a variety of oligonucleotide competitors. Adding a large excess of an unlabeled oligonucleotide may lead to two possible scenarios:

(i) The competitor is unable to trap the quadruplex ligand. In this case, Tm of the (F21T + PhenDC3) system is not affected by the competing oligonucleotide (in other words, ΔTm remains high). This is the expected outcome for a single-strand or a DNA or RNA duplex.

(ii) If the unlabeled competitor has a high affinity for PhenDC3, it will sequester a significant fraction of the compound, which will be no longer available for F21T stabilization, leading to a decrease in Tm. In this case, if the competition is very efficient, the Tm should fall back close to the value obtained without quadruplex ligand, meaning for F21T alone (in other words, $\Delta Tm \approx 0$).

To proceed, we selected a variety of competitors for which the structure was previously investigated and characterized [42]. This collection of over 60 sequences includes a variety of

quadruplex-forming motifs (with various topologies) as well as single- and double-stranded DNAs (sequences shown in **Table S1**).

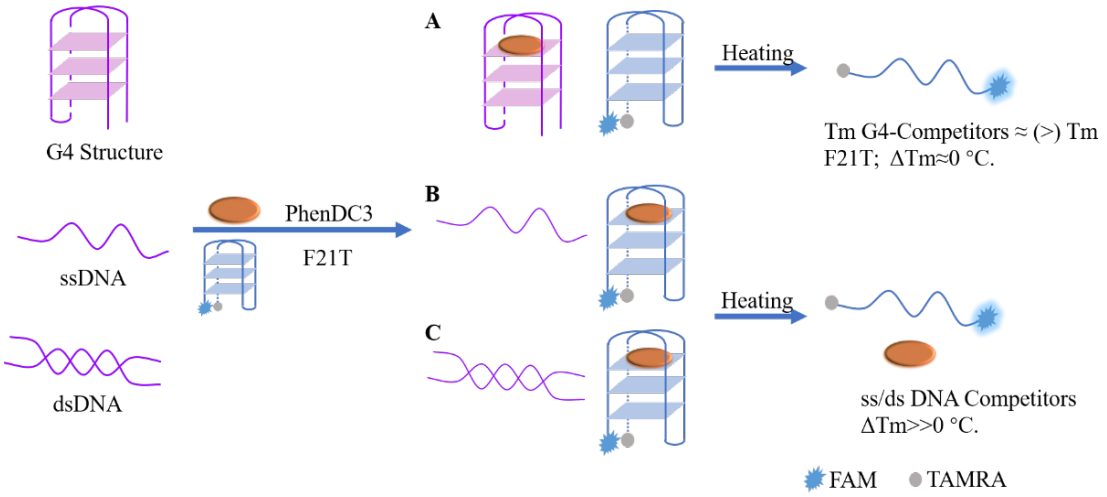

Fig. 1. Principle of the FRET- MC assay. In panel **A**, the competitor forms a quadruplex and traps PhenDC3 (shown as an orange oval). In panels **B** and **C**, the competitor does not form a quadruplex and has no affinity for PhenDC3, which remains bound to F21T (in light blue). Competitor sequences are shown in purple, PhenDC3 is represented as an orange oval, and F21T is represented in blue.

3.2 Validation of the FRET-melting competition assay with a set of 65 sequences.

A trivial, but important, control was first performed by checking that the competitors do not directly interact with F21T. To that aim, Tm of F21T was measured alone or in the presence of each competitor, in the absence of PhenDC3. As expected, most sequences tested had negligible, if any, effect on F21T melting (**Fig. S2**; normalized FRET-melting curves are shown in **Fig. S3)**. A few motifs (46AG, T95-2T, T2B-1, AT11, LWDLN1, AND1, RND3, RND6 and AT26) led to a significant decrease in Tm (ΔTm > 5 °C).

We next investigated the impact of the competitors on the stabilization effect (ΔTm) induced by PhenDC3 on F21T. **Fig. 2** presents examples of FRET-melting profiles for F21T alone (**Fig.**

2A), F21T in competition with a stable G4 (**Fig. 2B**), a single-strand (**Fig. 2C**) or a duplex DNA (**Fig. 2D**), respectively.

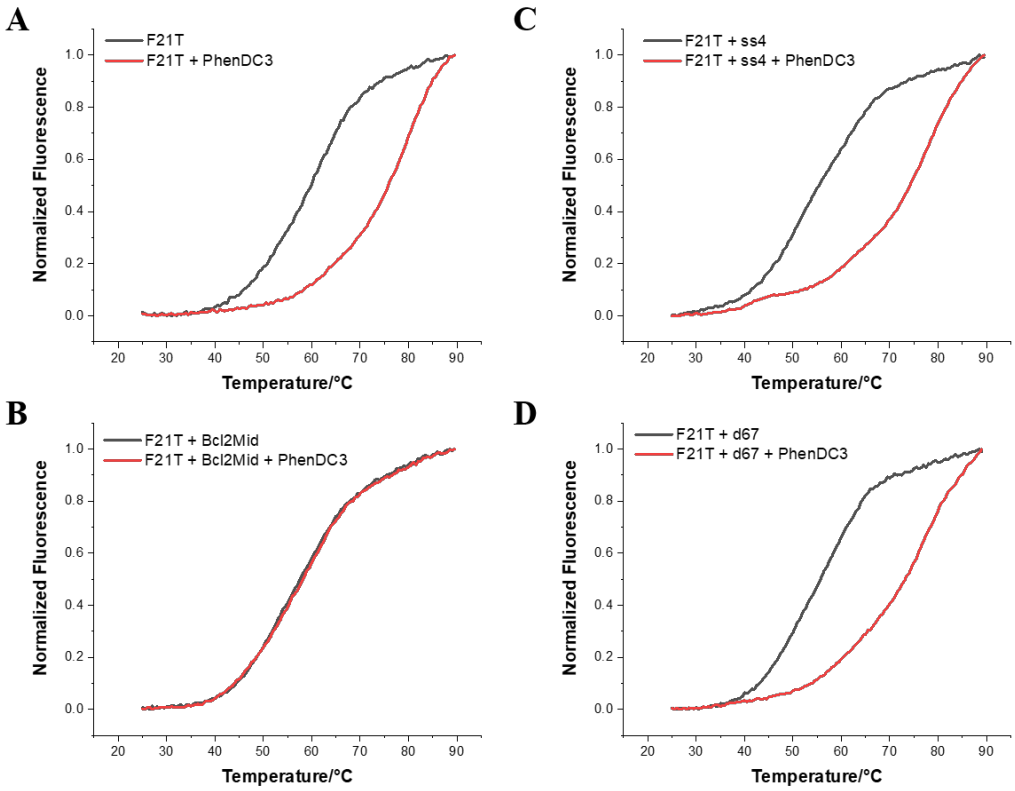

Fig. 2. FRET-melting profiles of 0.2 µM F21T alone or in the presence of 3 µM competitors, and with (red) or without (black) 0.4 µM PhenDC3. (A) F21T alone and in the presence of PhenDC3, and with the addition of the competitor: **(B)** G4 structure, **(C)** single-strand DNA and **(D)** duplex DNA. Samples were annealed and measured in 10 mM KCl, 90 mM LiCl, 10 mM lithium cacodylate pH 7.2 buffer.

In the absence of any competitor, 2 eqv. of PhenDC3 induces a ΔTm of 23.4°C, in agreement with previous results [43]. As expected, none of the competitors induced a further significant increase in Tm, as compared to F21T + PhenDC3. **Fig. 3** summarizes the ΔTm results obtained for F21T in the presence of PhenDC3 and in the presence or absence of oligonucleotide competitor (normalized FRET-melting curves shown in **Fig. S3**). Many, but not all, of the sequences known to form G4 structures led to a significant drop in ΔTm values (upper half of the figure), showing that they acted as efficient competitors. In contrast, single-stranded and

duplex DNAs had little impact, as ΔTm values remained high, close to the value found for F21T + PhenDC3 with no competitor. To quantitate this competition effect, we defined the *S Factor*, as originally described in [44], which corresponds to the relative PhenDC3 stabilization remaining in the presence of the competitor:

$$S\ Factor = \frac{\Delta\text{Tm of F21T with competitors}}{\Delta\text{Tm of F21T alone}}$$

Based on *S Factor*, the competitors can be divided into two categories: *(i)* ineffective competitors, for which *S* remains ≈ 1, meaning that competition is nearly completely unproductive, as expected for a structure for which PhenDC3 has no affinity, and conversely *(ii)* potent competitors (*i.e.,* stable quadruplexes) would give a *S* Factor close to 0 (**Fig. 3**).

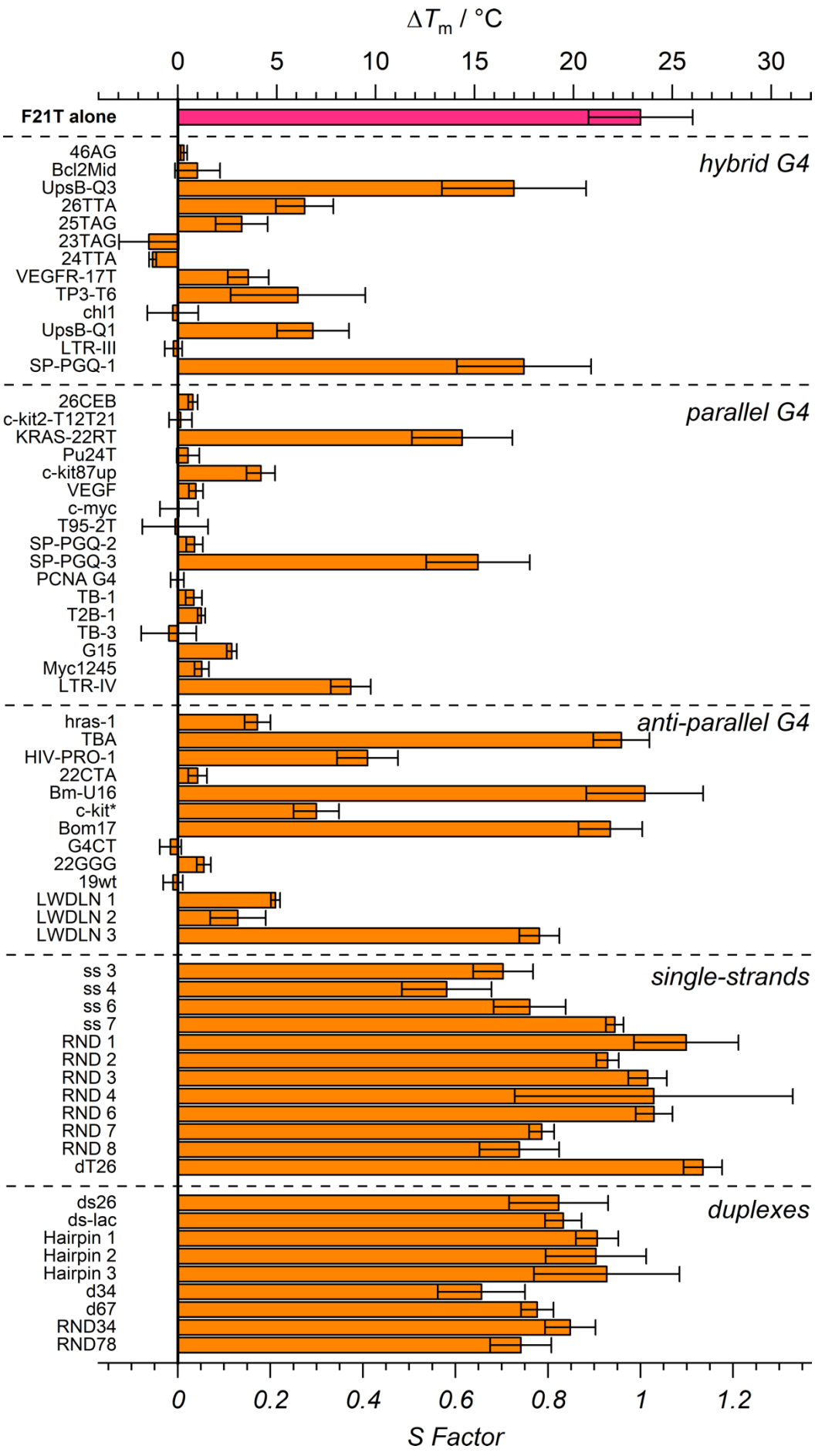
ΔTm / °C
0
5
10
15
20
25
30
F21T alone
hybrid G4
46AG
Bcl2Mid
UpsB-Q3
26TTA
25TAG
23TAG
24TTA
VEGFR-17T
TP3-T6
chl1
UpsB-Q1
LTR-III
SP-PGQ-1
parallel G4
26CEB
c-kit2-T12T21
KRAS-22RT
Pu24T
c-kit87up
VEGF
c-myc
T95-2T
SP-PGQ-2
SP-PGQ-3
PCNA G4
TB-1
T2B-1
TB-3
G15
Myc1245
LTR-IV
anti-parallel G4
hras-1
TBA
HIV-PRO-1
22CTA
Bm-U16
c-kit*
Bom17
G4CT
22GGG
19wt
LWDLN 1
LWDLN 2
LWDLN 3
single-strands
ss 3
ss 4
ss 6
ss 7
RND 1
RND 2
RND 3
RND 4
RND 6
RND 7
RND 8
dT26
duplexes
ds26
ds-lac
Hairpin 1
Hairpin 2
Hairpin 3
d34
d67
RND34
RND78
0
0.2
0.4
0.6
0.8
1
1.2
S Factor

Fig. 3. ΔTm induced by 0.4 µM PhenDC3 on 0.2 µM F21T, alone or in the presence of 3 µM competitors. The *S* Factor (bottom X-axis) provides a normalized value. Samples were annealed and measured in 10 mM KCl, 90 mM LiCl, 10 mM lithium cacodylate pH 7.2 buffer.

As expected, the majority of G4-forming sequences led to *S* values close to 0. However, several known G4 structures (UpsB-Q3, SP-PGQ-1, KRAS-22RT, SP-PGQ-3, TBA, Bm-U16, Bom17 and LWDLN3) were not efficient competitors ($S > 0.6$). In order to understand these results, we performed UV-melting experiments for these 8 sequences and collected Tm values in **Table S2** (detailed UV-melting curves are shown in **Fig. S4)**. The thermal stability of all these sequences (except for SP-PGQ-1) was relatively low, indicating they form unstable G4 structures, which are likely to be unfolded in the temperature range where F21T starts to melt: they are then "seen" as non-specific single-strands rather than true G-quadruplexes. SP-PGQ-1 behaved differently: UV-absorbance at 295 nm of a quadruplex should decrease upon heating due to the unfolding of the quadruplex structure. Although SP-PGQ-1 was reported to form a hybrid G4 structure [45], our results show an unexpected *increase* in absorbance at 295 nm upon UV-melting, incompatible with the unfolding of a quadruplex, and rather suggesting the formation of a another structure (*e.g.*, a mismatched duplex) at low temperatures. CD spectroscopy and TDS confirmed this hypothesis **(Fig. S5)**. In contrast, the Tm of 10 different G4 sequences acting as effective competitors ($S < 0.3$ for 46AG, Bcl2Mid, 25TGA, Chl, LTR-III, c-kit-T12T2, Pu24T, c-kit87up, VEGF and T95-2T) were always higher than the Tm of the false negative sequences, as shown in **Table S2** (UV-melting curves showed in **Fig. S4**). Although some of the Tm are still lower than the Tm of F21T, their presence in large excess (15-fold molar excess as compared to F21T) may compensate for a partial denaturation.

We then investigated whether one could substitute PhenDC3 by a different G4 ligand, TMPyP4, a cationic porphyrin which also has a high affinity for G-quadruplexes, but is much less selective. **Fig. S6** presents ΔTm / *S* values for 0.4 µM TMPyP4 on 0.2 µM F21T, alone or in the presence of various competitors at 3 µM strand concentration. Experiments were done in a buffer identical to the one used for PhenDC3. We tested 5 representative sequences for each structural type considered here (25 different competitors in total). Although parallel

quadruplexes were the most efficient competitors, some single-strands and most duplexes were also competing, with *S* value around 0.5, lower than the *S* values found for most anti-parallel and two hybrid quadruplexes with TMPyP4. This experiment demonstrates that, for this method to be reliable, a truly specific G4 ligand has no preferential binding to any G4 topology must be chosen. While other compounds than PhenDC3 may fit the bill, moderately selective compounds will not.

A critical factor for the competition efficiency (measured by the *S* value) in the FRET-MC assay should be the affinity and number of binding sites present on the competitor sequence for the quadruplex ligand (PhenDC3 in all further experiments). We wanted to investigate if the thermal stability of the structure itself would also contribute. All duplex sequences tested here had a significantly higher Tm value than F21T (as shown in **Table S2**); they were still poor competitors, even when added in large excess, as PhenDC3 is unable to bind to duplexes.

For the quadruplexes, in order to investigate how *S* value correlates with Tm, we tried a range of 6 different competitor concentrations (1x to 30x molar excess, as compared to F21T) for 9 different G4-forming sequences. As shown in **Fig. S7,** the Tm for each quadruplex (detailed UV-melting curves are shown in **Fig. S4**) was experimentally determined under identical conditions, and is shown in red below; sequences were ranked from left (lowest Tm) to right (highest Tm). Overall, there is indeed a correlation between Tm and *S* values: the competitors with a high thermal stability lead to low *S* values when the competitor is not in huge excess. Since *S* values are low for nearly all sequences at 3 or 6 μM, independently of Tm, we did not consider these two concentrations in **Fig. S8**, which provides a different view / representation of these *S* values, in which sequences are clustered in three categories according to Tm. The most striking difference is found between the high stability group and the others. **Table S3** summarizes average *S* values (from 1x to 10x molar excess) for these 9 sequences. Generally, average *S* values decreased with increasing Tm but the correlation is far from perfect. The average *S* values for high stability sequences are 0.25 or below, while *S* values for the others are above 0.35.

3.3 Experimental validation on novel sequences.

To test this method on a set of novel sequences, we used eight genomic DNA sequences of unknown structure. These motifs are found close to the transcription start site of different human promoters, and their biological relevance is currently being investigated. According to G4Hunter analysis [19], four of them (UN1–4) were likely to form a quadruplex structure, as their G4Hunter score was above 1.2, while the remaining four (UN5–8) were unlikely to adopt a G4 conformation, with a G4Hunter score < 1.0. Detailed sequences with location information (Human hg19) and G4Hunter scores are shown in **Table S4.**

The ΔTm values and *S*-factors for the tested sequences are shown on **Fig. 4**. The first four oligonucleotides (UN1-4) with high G4-hunter scores gave low *S-factor* values (< 0.1), consistent with G4 formation. In contrast, the remaining four samples (UN5-8) were poor competitors (S > 0.8) and unlikely to form quadruplexes, in agreement with their low G4Hunter scores. These conclusions were confirmed by three independent techniques (CD, IDS and TDS) measured in 100 mM KCl (**Fig S9**), all giving results consistent with the formation of G4 structures by UN1–4 and the absence of these structures in the case of UN5–8.

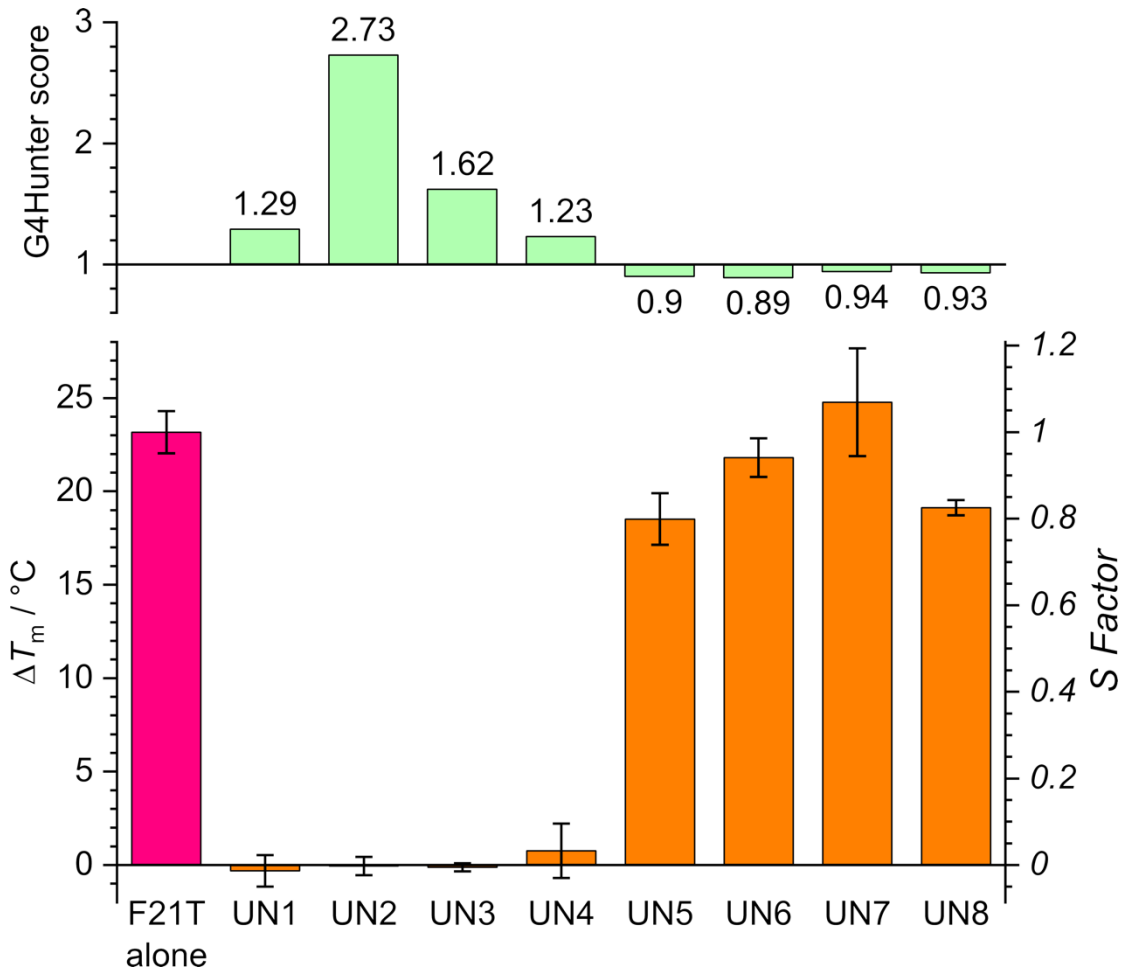

Fig. 4. Testing 8 different sequences (UN1-UN8) with the FRET-MC assay. The G4Hunter score for each competitor is indicated on the upper part of the figure (green bars). The ΔTm induced by 0.4 µM PhenDC3 on 0.2 µM F21T, alone or in the presence of 3 µM competitors (UN1-8) is plotted on the lower part of the graph. The *S* Factor (right Y-axis) provides a normalized value. Samples were annealed and measured in 10 mM KCl, 90 mM LiCl, 10 mM lithium cacodylate pH 7.2 buffer.

4. Conclusion

The FRET-MC assay described here is a fast and inexpensive characterization method to determine if a sequence is forming a stable quadruplex or not. It offers several advantages:

- It is relatively *inexpensive*: while the F21T dual-labeled oligonucleotide is relatively expensive, minimal amounts are used for each point as the volume is reduced (25 µL)

and its concentration is only 0.2 µM, and can even be further reduced if necessary, provided a sensitive RT-PCR instrument is available. Conversely, the assayed sequences are unmodified oligonucleotides and do not require extensive purification; only 25 picomoles are needed per point.

- It is *fast*: FRET-melting takes 1-2 hours while UV-melting requires several hours (14 h with the temperature gradient used here) for each experiment.
- It is *simple* to set up: all reagents are commercially available (PhenDC3 included) and the FRET melting assay is now routinely used in a number of labs.
- It allows testing *multiple samples in parallel*. While a classical UV spectrophotometer can only read up to 6 or 9 samples, FRET-melting uses a 96-well plate as a sample holder, and it is able to process 48 sequences in duplicate. It may even be transposed to 384-well format.

At the same time, this method has several limitations:

- The main assumptions for this technique are that PhenDC3 *i)* indiscriminately binds to all G-quadruplexes, and *ii)* does not bind to other structures. In other words, for this method to work, we need a perfect, general G4 ligand with high structure specificity. Previously published studies [31, 37] have shown that PhenDC3 does indeed bind to all G4 tested so far, and has excellent specificity. We cannot exclude, however, that PhenDC3 would also recognize other unusual motifs such as G-triplexes [46]. Additional experiments are therefore required to reach a clear conclusion for a given sequence: as previously stated [19], we advocate the use of several independent techniques to assess G4 formation. On the other hand, a high-affinity but poorly selective quadruplex ligand such as TMPyP4 proved to be ineffective for this application.

- The sequence to be tested should not be complementary to F21T, as it would interfere with the structure of the fluorescent probe. C-rich sequences, and especially repetitions of the CCCTAA hexanucleotide motif should be avoided. The training set chosen for this study did not involve any i-motif sequence, resulting from the folding of C-rich oligonucleotides. These sequences would then be partially complementarity to the fluorescent G-rich oligonucleotide F21T: duplex formation would “kill” the assay by interfering with G4 formation. On the other

hand, the method itself may later be adapted to the analysis of i-motif sequences. This assay would involve a "complementary" sequence for i-motif formation, in which the fluorescently labeled oligonucleotide is C-rich, not G-rich, and forms an i-motif itself. But this would require a "perfect" i-motif ligand as well, *i.e.* a compound that would bind reasonably well to all i-motifs while having no affinity for any other structure. The i-motif ligands we have tested do not meet these criteria, and cannot be considered as equivalent to PhenDC3 for this purpose.

- The main limitation of this FRET-melting competition assay is its inability to detect unstable quadruplexes which behave as single-strands (**Fig. 3**) at the temperature where F21T starts to melt. This assay should therefore be employed to identify moderately stable or highly stable G4 structures. A possible way to circumvent this limitation would be to replace or complement F21T by a quadruplex probe with lower stability [47]. A two-quartet quadruplex such as the thrombin binding aptamer (TBA) could be proposed, keeping in mind that these G4 are often weaker binders for G4 ligands such as PhenDC3. In any case, unstable quadruplexes are less likely to be biologically relevant [48] and unlikely to be identified by genome-wide methods such as G4-seq [49] as extension is performed at 60 °C during Illumina sequencing.

- Finally, the *S* value cannot be used as a proxy for the (thermal) stability of the tested quadruplex. While there is some correlation between thermal stability and competition efficiency (stable G4 tend to give lower *S* values), other factors contribute to the competition efficiency, such as the affinity of the PhenDC3 ligand for this topology, and the number of binding sites available.

Overall, despites the shortcomings listed above, the FRET-melting competition assay should constitute an interesting addition to the *in vitro* "G4 characterization toolbox".

Data availability

Raw data / melting profiles may be downloaded at https://data.mendeley.com/datasets/gspc9r73r5/1.

Acknowledgments

We thank both reviewers for excellent suggestions, Laurent Lacroix (ENS, Paris) for helpful discussions, and Corinne Landras Guetta and Marie-Paule Teulade-Fichou (Institut Curie, Orsay) for a sample of PhenDC3. This manuscript is dedicated to the memory of Prof. Michael J. Waring, with whom J.L.M. had interesting lively discussions about DNA ligands during his sabbatical in France.

Conflict of interest

None

References

1. Gellert, M.; Lipsett, M. N.; Davies, D. R., Helix formation by guanylic acid. *Proc Natl Acad Sci U S A* **1962,** *48* (12), 2013-8.

2. Guo, Y.; Xu, L.; Hong, S.; Sun, Q.; Yao, W.; Pei, R., Label-free DNA-based biosensors using structure-selective light-up dyes. *Analyst* **2016,** *141* (24), 6481-6489.

3. He, H. Z.; Chan, D. S.; Leung, C. H.; Ma, D. L., G-quadruplexes for luminescent sensing and logic gates. *Nucleic Acids Res* **2013,** *41* (8), 4345-59.

4. Ren, J.; Wang, T.; Wang, E.; Wang, J., Versatile G-quadruplex-mediated strategies in label-free biosensors and logic systems. *Analyst* **2015,** *140* (8), 2556-72.

5. Ruttkay-Nedecky, B.; Kudr, J.; Nejdl, L.; Maskova, D.; Kizek, R.; Adam, V., G-quadruplexes as sensing probes. *Molecules* **2013,** *18* (12), 14760-79.

6. Lustgarten, O.; Carmieli, R.; Motiei, L.; Margulies, D., A Molecular Secret Sharing Scheme. *Angew Chem Int Ed Engl* **2019,** *58* (1), 184-188.

7. Mergny, J. L.; Sen, D., Correction to DNA Quadruple Helices in Nanotechnology. *Chem Rev* **2020**.

8. Stefan, L.; Monchaud, D., Applications of guanine quartets in nanotechnology and chemical biology. *Nature Reviews Chemistry* **2019,** *3* (11), 650-668.

9. Cao, Y.; Gao, S.; Yan, Y.; Bruist, M. F.; Wang, B.; Guo, X., Assembly of supramolecular DNA complexes containing both G-quadruplexes and i-motifs by enhancing the G-repeat-bearing capacity of i-motifs. *Nucleic Acids Res* **2017,** *45* (1), 26-38.

10. Spiegel, J.; Adhikari, S.; Balasubramanian, S., The Structure and Function of DNA G-Quadruplexes. *Trends Chem* **2020,** *2* (2), 123-136.

11. Varshney, D.; Spiegel, J.; Zyner, K.; Tannahill, D.; Balasubramanian, S., The regulation and functions of DNA and RNA G-quadruplexes. *Nat Rev Mol Cell Biol* **2020,** *21* (8), 459-474.

12. Burger, A. M.; Dai, F.; Schultes, C. M.; Reszka, A. P.; Moore, M. J.; Double, J. A.; Neidle, S., The G-Quadruplex-Interactive Molecule BRACO-19 Inhibits Tumor Growth, Consistent with Telomere Targeting and Interference with Telomerase Function. *Cancer Research* **2005,** *65* (4), 1489.

13. Read, M.; Harrison, R. J.; Romagnoli, B.; Tanious, F. A.; Gowan, S. H.; Reszka, A. P.; Wilson, W. D.; Kelland, L. R.; Neidle, S., Structure-based design of selective and potent G quadruplex-mediated telomerase inhibitors. *Proceedings of the National Academy of Sciences* **2001,** *98* (9), 4844.

14. Riou, J. F.; Guittat, L.; Mailliet, P.; Laoui, A.; Renou, E.; Petitgenet, O.; Mégnin-Chanet, F.; Hélène, C.; Mergny, J. L., Cell senescence and telomere shortening induced by a new series of specific G-quadruplex DNA ligands. *Proceedings of the National Academy of Sciences USA* **2002,** *99* (5), 2672.

15. Brooks, T. A.; Kendrick, S.; Hurley, L., Making sense of G-quadruplex and i-motif functions in oncogene promoters. *FEBS J* **2010,** *277* (17), 3459-69.

16. Jana, J.; Mondal, S.; Bhattacharjee, P.; Sengupta, P.; Roychowdhury, T.; Saha, P.; Kundu, P.; Chatterjee, S., Chelerythrine down regulates expression of VEGFA, BCL2 and KRAS by arresting G-Quadruplex structures at their promoter regions. *Sci Rep* **2017,** *7*, 40706.

17. Marquevielle, J.; Robert, C.; Lagrabette, O.; Wahid, M.; Bourdoncle, A.; Xodo, L. E.; Mergny, J. L.; Salgado, G. F., Structure of two G-quadruplexes in equilibrium in the KRAS promoter. *Nucleic Acids Res* **2020,** *48* (16), 9336-9345.

18. Spiegel, J.; Adhikari, S.; Balasubramanian, S., The Structure and Function of DNA G-Quadruplexes. *Trends in Chemistry* **2020,** *2* (2), 123-136.

19. Bedrat, A.; Lacroix, L.; Mergny, J. L., Re-evaluation of G-quadruplex propensity with G4Hunter. *Nucleic Acids Res* **2016,** *44* (4), 1746-59.

20. Brazda, V.; Kolomaznik, J.; Lysek, J.; Bartas, M.; Fojta, M.; Stastny, J.; Mergny, J. L., G4Hunter web application: a web server for G-quadruplex prediction. *Bioinformatics* **2019,** *35* (18), 3493-3495.

21. Alba, J. J.; Sadurni, A.; Gargallo, R., Nucleic Acid i-Motif Structures in Analytical Chemistry. *Crit Rev Anal Chem* **2016,** *46* (5), 443-54.

22. Adrian, M.; Heddi, B.; Phan, A. T., NMR spectroscopy of G-quadruplexes. *Methods* **2012,** *57* (1), 11-24.

23. Kypr, J.; Kejnovska, I.; Renciuk, D.; Vorlickova, M., Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res* **2009,** *37* (6), 1713-25.

24. Mergny, J. L.; Li, J.; Lacroix, L.; Amrane, S.; Chaires, J. B., Thermal difference spectra: a specific signature for nucleic acid structures. *Nucleic Acids Res* **2005,** *33* (16), e138.

25. Renaud de la Faverie, A.; Guedin, A.; Bedrat, A.; Yatsunyk, L. A.; Mergny, J. L., Thioflavin T as a fluorescence light-up probe for G4 formation. *Nucleic Acids Res* **2014,** *42* (8), e65.

26. Sabharwal, N. C.; Savikhin, V.; Turek-Herman, J. R.; Nicoludis, J. M.; Szalai, V. A.; Yatsunyk, L. A., N-methylmesoporphyrin IX fluorescence as a reporter of strand orientation in guanine quadruplexes. *FEBS J* **2014,** *281* (7), 1726-37.

27. Xie, X.; Renvoisé, A.; Granzhan, A.; Teulade-Fichou, M.-P., Aggregating distyrylpyridinium dye as a bimodal structural probe for G-quadruplex DNA. *New Journal of Chemistry* **2015,** *39* (8), 5931-5935.

28. Kreig, A.; Calvert, J.; Sanoica, J.; Cullum, E.; Tipanna, R.; Myong, S., G-quadruplex formation in double strand DNA probed by NMM and CV fluorescence. *Nucleic Acids Res* **2015,** *43* (16), 7961-70.

29. Zuffo, M.; Xie, X.; Granzhan, A., Strength in Numbers: Development of a Fluorescence Sensor Array for Secondary Structures of DNA. *Chemistry* **2019,** *25* (7), 1812-1818.

30. O'Hagan, M. P.; Morales, J. C.; Galan, M. C., Binding and Beyond: What Else Can G-Quadruplex Ligands Do? *European Journal of Organic Chemistry* **2019,** *2019* (31-32), 4995-5017.

31. Tran, P. L.; Largy, E.; Hamon, F.; Teulade-Fichou, M. P.; Mergny, J. L., Fluorescence intercalator displacement assay for screening G4 ligands towards a variety of G-quadruplex structures. *Biochimie* **2011,** *93* (8), 1288-96.

32. De Cian, A.; Guittat, L.; Kaiser, M.; Sacca, B.; Amrane, S.; Bourdoncle, A.; Alberti, P.; Teulade-Fichou, M. P.; Lacroix, L.; Mergny, J. L., Fluorescence-based melting assays for studying quadruplex ligands. *Methods* **2007,** *42* (2), 183-95.

33. Mergny, J.-L.; Maurizot, J.-C., Fluorescence Resonance Energy Transfer as a Probe for G-Quartet Formation by a Telomeric Repeat. *ChemBioChem* **2001,** *2* (2), 124-132.

34. Renciuk, D.; Zhou, J.; Beaurepaire, L.; Guedin, A.; Bourdoncle, A.; Mergny, J. L., A FRET-based screening assay for nucleic acid ligands. *Methods* **2012,** *57* (1), 122-8.

35. Marchand, A.; Rosu, F.; Zenobi, R.; Gabelica, V., Thermal Denaturation of DNA G-Quadruplexes and Their Complexes with Ligands: Thermodynamic Analysis of the Multiple States Revealed by Mass Spectrometry. *J Am Chem Soc* **2018,** *140* (39), 12553-12565.

36. Morgan, R. K.; Psaras, A. M.; Lassiter, Q.; Raymer, K.; Brooks, T. A., G-quadruplex deconvolution with physiological mimicry enhances primary screening: Optimizing the FRET Melt(2) assay. *Biochim Biophys Acta Gene Regul Mech* **2020,** *1863* (1), 194478.

37. De Cian, A.; DeLemos, E.; Mergny, J.-L.; Teulade-Fichou, M.-P.; Monchaud, D., Highly Efficient G-Quadruplex Recognition by Bisquinolinium Compounds. *Journal of the American Chemical Society* **2007,** *129* (7), 1856-1857.

38. Marchand, A.; Granzhan, A.; Iida, K.; Tsushima, Y.; Ma, Y.; Nagasawa, K.; Teulade-Fichou, M. P.; Gabelica, V., Ligand-induced conformational changes with cation ejection upon binding to human telomeric DNA G-quadruplexes. *J Am Chem Soc* **2015,** *137* (2), 750-6.

39. Ruggiero, E.; Richter, S. N., G-quadruplexes and G-quadruplex ligands: targets and tools in antiviral therapy. *Nucleic Acids Res* **2018,** *46* (7), 3270-3283.

40. Mergny, J.-L.; Lacroix, L., UV Melting of G-Quadruplexes. *Current Protocols in Nucleic Acid Chemistry* **2009,** *37* (1), 17.1.1-17.1.15.

41. Mergny, J.-L.; Lacroix, L., Analysis of Thermal Melting Curves. *Oligonucleotides* **2003,** *13* (6), 515-537.

42. Zuffo, M.; Gandolfini, A.; Heddi, B.; Granzhan, A., Harnessing intrinsic fluorescence for typing of secondary structures of DNA. *Nucleic Acids Res* **2020,** *48* (11), e61.

43. Gueddouda, N. M.; Hurtado, M. R.; Moreau, S.; Ronga, L.; Das, R. N.; Savrimoutou, S.; Rubio, S.; Marchand, A.; Mendoza, O.; Marchivie, M.; Elmi, L.; Chansavang, A.; Desplat, V.; Gabelica, V.; Bourdoncle, A.; Mergny, J. L.; Guillon, J., Design, Synthesis, and Evaluation of 2,9-Bis[(substituted-aminomethyl)phenyl]-1,10-phenanthroline Derivatives as G-Quadruplex Ligands. *ChemMedChem* **2017,** *12* (2), 146-160.

482 44. Monchaud, D.; Allain, C.; Bertrand, H.; Smargiasso, N.; Rosu, F.; Gabelica, V.;
483 De Cian, A.; Mergny, J. L.; Teulade-Fichou, M. P., Ligands playing musical chairs with G-
484 quadruplex DNA: a rapid and simple displacement assay for identifying selective G-quadruplex
485 binders. *Biochimie* **2008,** *90* (8), 1207-23.

486 45. Mishra, S. K.; Jain, N.; Shankar, U.; Tawani, A.; Sharma, T. K.; Kumar, A.,
487 Characterization of highly conserved G-quadruplex motifs as potential drug targets in
488 Streptococcus pneumoniae. *Sci Rep* **2019,** *9* (1), 1791.

489 46. Bonnat, L.; Dautriche, M.; Saidi, T.; Revol-Cavalier, J.; Dejeu, J.; Defrancq, E.;
490 Lavergne, T., Scaffold stabilization of a G-triplex and study of its interactions with G-
491 quadruplex targeting ligands. *Organic & Biomolecular Chemistry* **2019,** *17* (38), 8726-8736.

492 47. De Rache, A.; Mergny, J. L., Assessment of selectivity of G-quadruplex ligands via an
493 optimised FRET melting assay. *Biochimie* **2015,** *115*, 194-202.

494 48. Piazza, A.; Adrian, M.; Samazan, F.; Heddi, B.; Hamon, F.; Serero, A.; Lopes,
495 J.; Teulade-Fichou, M. P.; Phan, A. T.; Nicolas, A., Short loop length and high thermal
496 stability determine genomic instability induced by G-quadruplex-forming minisatellites.
497 *EMBO J* **2015,** *34* (12), 1718-34.

498 49. Chambers, V. S.; Marsico, G.; Boutell, J. M.; Di Antonio, M.; Smith, G. P.;
499 Balasubramanian, S., High-throughput sequencing of DNA G-quadruplex structures in the
500 human genome. *Nat Biotechnol* **2015,** *33* (8), 877-81.

501

FRET-MC: a fluorescence melting competition assay for studying G4 structures *in vitro*

Yu Luo[1,2], Anton Granzhan[1], Daniela Verga[1*] & Jean-Louis Mergny[2*]

1. Université Paris Saclay, CNRS UMR9187, INSERM U1196, Institut Curie, 91400 Orsay, France.
2. Laboratoire d'Optique et Biosciences, Ecole Polytechnique, CNRS, Inserm, Institut Polytechnique de Paris, 91128 Palaiseau, France.

* Authors to whom correspondence may be addressed: daniela.verga@curie.fr; jean-louis.mergny@inserm.fr

Supplementary information

Tables S1-S4

Figures S1-S9

Table S1. Training set of DNA sequences.

Name	Sequence (5′-3′)	Reported conformation	PDB entry
F21T	FAM-GGGTTAGGGTTAGGGTTAGGG-TAMRA	-	-
46AG	AGGGTTAGGGTTAGGGTTAGGGTTAGGGTTA GGGTTAGGG	hybrid G4	-
Bcl2Mid	GGGCGCGGGAGGAATTGGGCGGG	hybrid G4	2F8U
UpsB-Q3	CAGGGTTAAGGGTATACATTTAGGGGTTAGGGTT	hybrid G4	-
26TTA	TTAGGGTTAGGGTTAGGGTTAGGGTT	hybrid G4	2JPZ
25TGA	TAGGGTTAGGGTTAGGGTTAGGGTT	hybrid G4	2JSL
23TAG	TAGGGTTAGGGTTAGGGTTAGGG	hybrid G4	2JSK
24TTA	TTAGGGTTAGGGTTAGGGTTAGGG	hybrid G4	2JSL
VEGFR-17T	GGGTACCCGGGTGAGGTGCGGGGT	hybrid G4	5ZEV
TP3-T6	TGGGGTCCGAGGCGGGGCTTGGG	hybrid G4	6AC7
chl1	GGGTGGGGAAGGGGTGGGT	hybrid G4	2KPR
UpsB-Q1	CAGGGTTAAGGGTATAACTTTAGGGGTTAGGGTT	hybrid G4	5MTA
LTR-III	GGGAGGCGTGGCCTGGGCGGGACTGGGG	hybrid G4	6H1K
SP-PGQ-1	GGGCAACTTGGCTGGGGTCTAGTTCCACGGGACGGG	hybrid G4	-
26CEB	AAGGGTGGGTGTAAGTGTGGGTGGGT	parallel G4	2LPW
c-kit2-T12T21	CGGGCGGGCGCTAGGGAGGGT	parallel G4	2KYP
KRAS-22RT	AGGGCGGTGTGGGAATAGGGAA	parallel G4	5I2V
Pu24T	TGAGGGTGGTGAGGGTGGGGAAGG	parallel G4	2A5P
c-kit87up	AGGGAGGGCGCTGGGAGGAGGG	parallel G4	2O3M
VEGF	CGGGGCGGGCCTTGGGCGGGGT	parallel G4	2M27
c-myc	TGAGGGTGGGTAGGGTGGGTAA	parallel G4	1XAV
T95-2T	TTGGGTGGGTGGGTGGGT	parallel G4	2LK7
SP-PGQ-2	GGGCTAGTGGGGGGAGGGGG	parallel G4	-
SP-PGQ-3	GGGCTAATAGGGAGAGCAGGGACGGGG	parallel G4	-
PCNA G4	CAGGGCGACGGGGGCGGGGCGGGGCG	parallel G4	-
TB-1	TTGTGGTGGGTGGGTGGGT	parallel G4	2M4P
T2B-1	TTGTTGGTGGGTGGGTGGGT	parallel G4	-
TB-3	TTGGGTGTGGTGGGTGGGT	parallel G4	-
G15	TTGGGGGGGGGGGGGGGT	parallel G4	2MB2
Myc1245	TTGGGGAGGGTTTTAAGGGTGGGGAAT	parallel G4	6NEB
AT11	TGGTGGTGGTTGTTGTGGTGGTGGTGGT	parallel G4	2N3M
LTR-IV	CTGGGCGGGACTGGGGAGTGGT	parallel G4	2N4Y
hras-1	TCGGGTTGCGGGCGCAGGGCACGGGCG	anti-parallel G4	-
TBA	GGTTGGTGTGGTTGG	anti-parallel G4	148D

HIV-PRO-1	TGGCCTGGGCGGGACTGGG	anti-parallel G4	-
22CTA	AGGGCTAGGGCTAGGGCTAGGG	anti-parallel G4	-
Bm-U16	TAGGTTAGGTTAGGTUAGG	anti-parallel G4	-
c-kit*	GGCGAGGAGGGGCGTGGCCGGC	anti-parallel G4	6GH0
Bom17	GGTTAGGTTAGGTTAGG	anti-parallel G4	-
G4CT	GGGGCTGGGGCTGGGGCTGGGG	anti-parallel G4	-
22GGG	GGGTTAGGGTTAGGGTTAGGGT	anti-parallel G4	2KF8
19wt	GGGGGAGGGGTACAGGGGTACAGGGG	anti-parallel G4	6FTU
LWDLN 1	GGGTTTGGGTTTTGGGAGGG	anti-parallel G4	5J05
LWDLN 2	GGGGTTGGGGTTTTGGGGAAGGGG	anti-parallel G4	2M6W
LWDLN 3	GGTTTGGTTTTGGTTGG	anti-parallel G4	5J4W
ss 3	GTCGCCGGGCCAGTCGTCCATAC	single strand	-
ss 4	GTATGGACGACTGGCCCGGCGAC	single strand	-
ss 6	GACGTGTCGAAAGAGCTCCGATTA	single strand	-
ss 7	TAATCGGAGCTCTTTCGACACGTC	single strand	-
RND1	CTATACGAAAACCTTTTGTATCATT	single strand	-
RND2	AATGATACAAAAGGTTTTCGTATAG	single strand	-
RND3	TAACGTTTATAATGTAGTCTCATTA	single strand	-
RND4	TAATGAGACTACATTATAAACGTTA	single strand	-
RND6	GTTGTCATTGCCCCCGAATAATTCT	single strand	-
RND7	GCCTTGCGGAGGCATGCGTCATGCT	single strand	-
RND8	AGCATGACGCATGCCTCCGCAAGGC	single strand	-
dT26	TTTTTTTTTTTTTTTTTTTTTTTTT	single strand	-
ds26	CAATCGGATCGAATTCGATCCGATTG	duplex	-
ds-lac	GAATTGTGAGCGCTCACAATTC	duplex	-
Hairpin 1	GGATTCTTGGATTTTCCAAGAATCC	duplex	-
Hairpin 2	TCGGTATTGTGTTTCACAATACCGA	duplex	-
Hairpin 3	AGGACGGTGTATTTTACACCGTCCT	duplex	-
d34	GTCGCCGGGCCAGTCGTCCATAC		-
	GTATGGACGACTGGCCCGGCGAC	duplex	-
d67	GACGTGTCGAAAGAGCTCCGATTA		-
	TAATCGGAGCTCTTTCGACACGTC	duplex	-
RND34	TAACGTTTATAATGTAGTCTCATTA		-
	TAATGAGACTACATTATAAACGTTA	duplex	-
RND78	AGAATTATTCGGGGGCAATGACAAC		-
	GTTGTCATTGCCCCCGAATAATTCT	duplex	-

Table S2. Tm of some competitors collected by UV-melting[a]

Name	Tm/°C	Note	Reported conformation
F21T	58.1 [a]	Probe	-
UpsB-Q3	39.4	False negative	hybrid G4
SP-PGQ-1[b]	46.8	Not a G4; rather a duplex	hybrid G4
KRAS-22RT	32.8	False negative	parallel G4
SP-PGQ-3	34.7	False negative	parallel G4
TBA	42.1	False negative	anti-parallel G4
Bm-U16	26.8	False negative	anti-parallel G4
LWDLN3	33.0	False negative	anti-parallel G4
Bom17	30.3	False negative	anti-parallel G4
46AG	43.3	Positive control [c]	hybrid G4
Bcl2Mid	52.0	Positive control [c]	hybrid G4
25TGA	47.3	Positive control [c]	hybrid G4
Chl1	62.7	Positive control [c]	hybrid G4
LTR-III	46.5	Positive control [c]	hybrid G4
c-kit-T12T2	46.2	Positive control [c]	parallel G4
Pu24T	74.1	Positive control [c]	parallel G4
c-kit87up	48.7	Positive control [c]	parallel G4
VEGF	66.5	Positive control [c]	parallel G4
T95-2T	81.4	Positive control [c]	parallel G4
22CTA	45.5	-	anti-parallel G4
G4CT	69.5	-	anti-parallel G4
22GGG	54.4	-	anti-parallel G4
19wt	80.9	-	anti-parallel G4
LWDLN1	45.1	-	anti-parallel G4
ds26	76.3	-	duplex
ds-lac	70.1	-	duplex
Hairpin1	71.4	-	duplex
Hairpin2	74.9	-	duplex
d34	78.8	-	duplex

[a] Tm of F21T was measured by FRET-melting, Tm of reported G4 sequences and duplex were collected at 295 nm and 260 nm, respectively.

[b] Tm of SP-PGQ-1 calculated by UV-melting curve at 260 nm.

[c] Positive controls are G4-forming sequences which give an S value close to 0.

Table S3. Tm and *S* Factor average of some sequences in the training set

Name	Tm/°C[a]	*S* Factor average [b]
46AG	43.3	0.49
c-kit2-T12T12	46.2	0.40
25TAG	47.3	0.60
Bcl2Mid	52.0	0.62
22GGG	54.4	0.50
G4CT	69.5	0.35
Pu24T	74.1	0.23
19wt	80.9	0.22
T95-2T	81.4	0.26

[a] Tm were collected at 295 nm.

[b] *S* Factor averages were determined by the *S* of 0.2 / 0.6 / 1 / 2 μM (1x to 10x molar excess, as compared to F21T) competitor concentrations.

Table S4. Testing set of DNA sequences.

Name	Sequence (5′-3′)	G4-Hunter score	Location (Human hg19)
F21T	FAM-GGGTTAGGGTTAGGGTTAGGG-TAMRA	-	-
UN1	CGGGCAGGGAGGGCGGCTGTGCGGGGC	1.59	chr3: 196045150-196045176
UN2	TGGGGCGGGGGAAGAGGGGGCGGGGT	2.73	chr8: 57124113-57124138
UN3	CGGGAAGGGGCGGGCGCAATGGGC	1.62	chr17: 62915412-62915435
UN4	TGGGAGGCGGAGGTGGGCAGGTTGCT	1.23	chr17: 16258771-16258791
UN5	GTGCTGGGGCGCCCACTTCGGGGTGGTGC	0.90	chr11: 364605-364633
UN6	AGTTGGTAGGCTGAGGCGGGAGGATTGC	0.89	chr5: 64905453-64905472
UN7	AGGGCCGGGAGAGGGATCCGCCATATTGGAGCTGGGGC	0.94	chr14: 36278253-36278290
UN8	AGGAAGCTGGGGTAGGAGAATTGCTTGA	0.93	chr12: 57876780-57876802

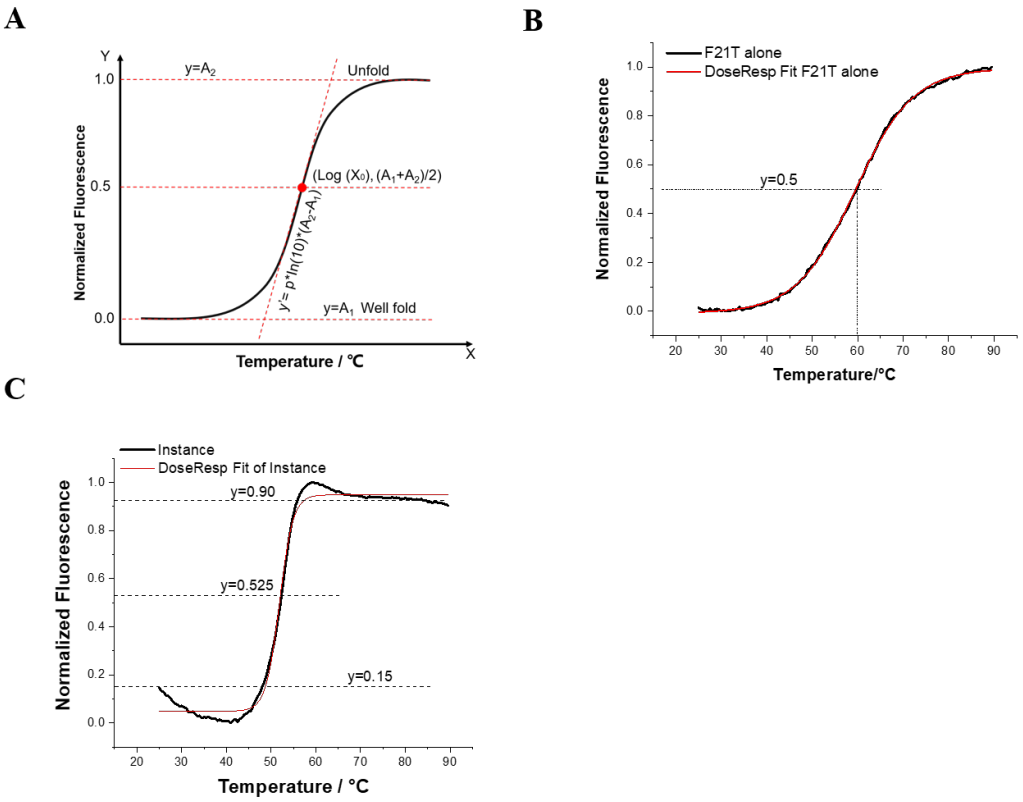

Fig. S1. Different methods to calculate Tm. (A) Sample curve of DoseResp. **(B)** Two methods to calculate Tm of F21T alone. **(C)** Example of an atypical FRET-melting curve, for which accurate Tm determination is more difficult.

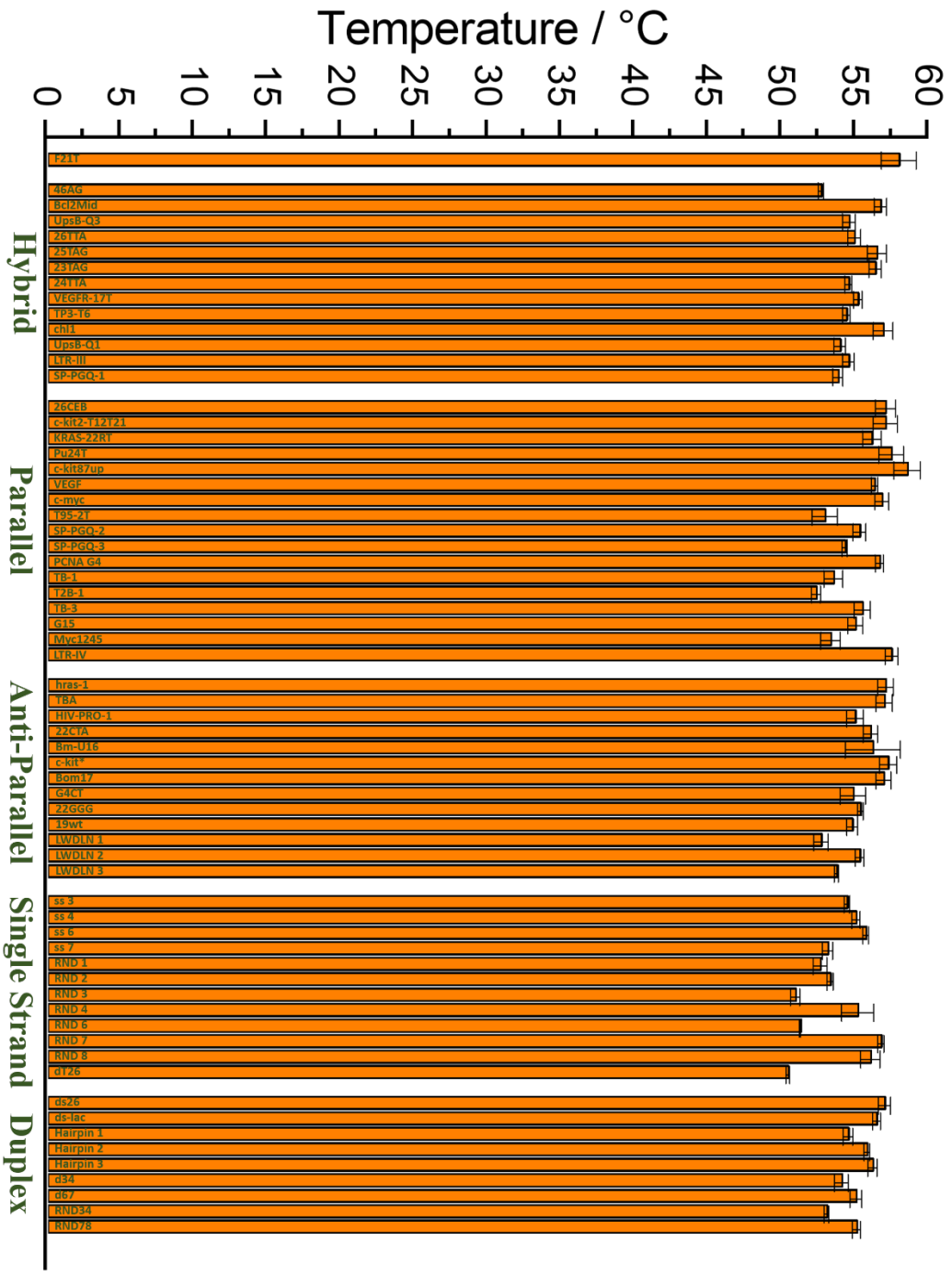

Fig. S2. FRET Tm of 0.2 µM F21T alone (top) or in the presence of various competitors.

All competitors were tested at 3 µM strand concentration.

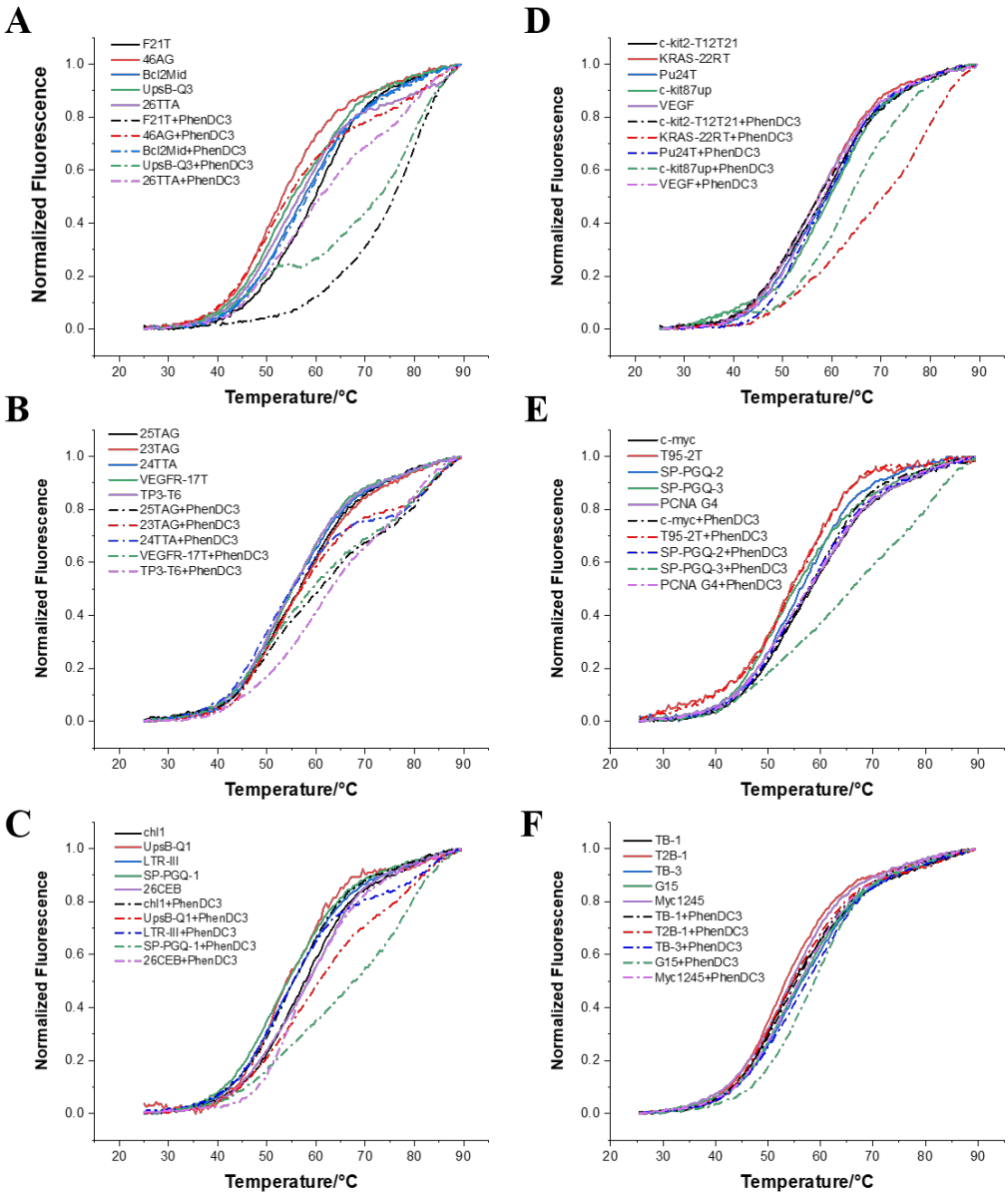

Fig. S3. Normalized FRET-melting curves of 0.2 μM F21T in the presence of various competitors (3 μM strand concentration) used in the training set, with or without 0.4 μM PhenDC3. Samples were annealed and measured in 10 mM KCl, 90 mM LiCl, 10 mM lithium cacodylate pH 7.2 buffer.

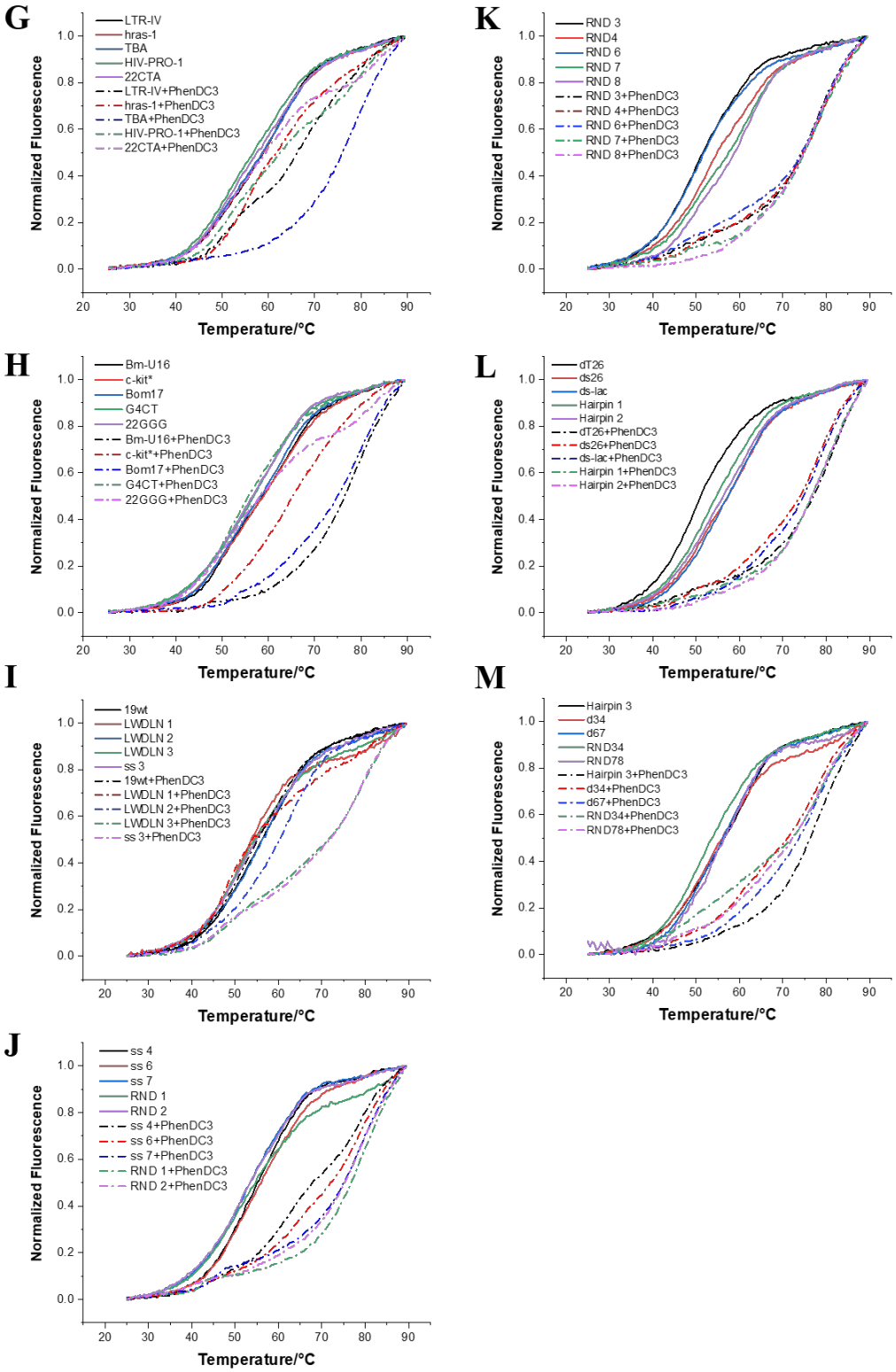

Fig. S3, *continued*.

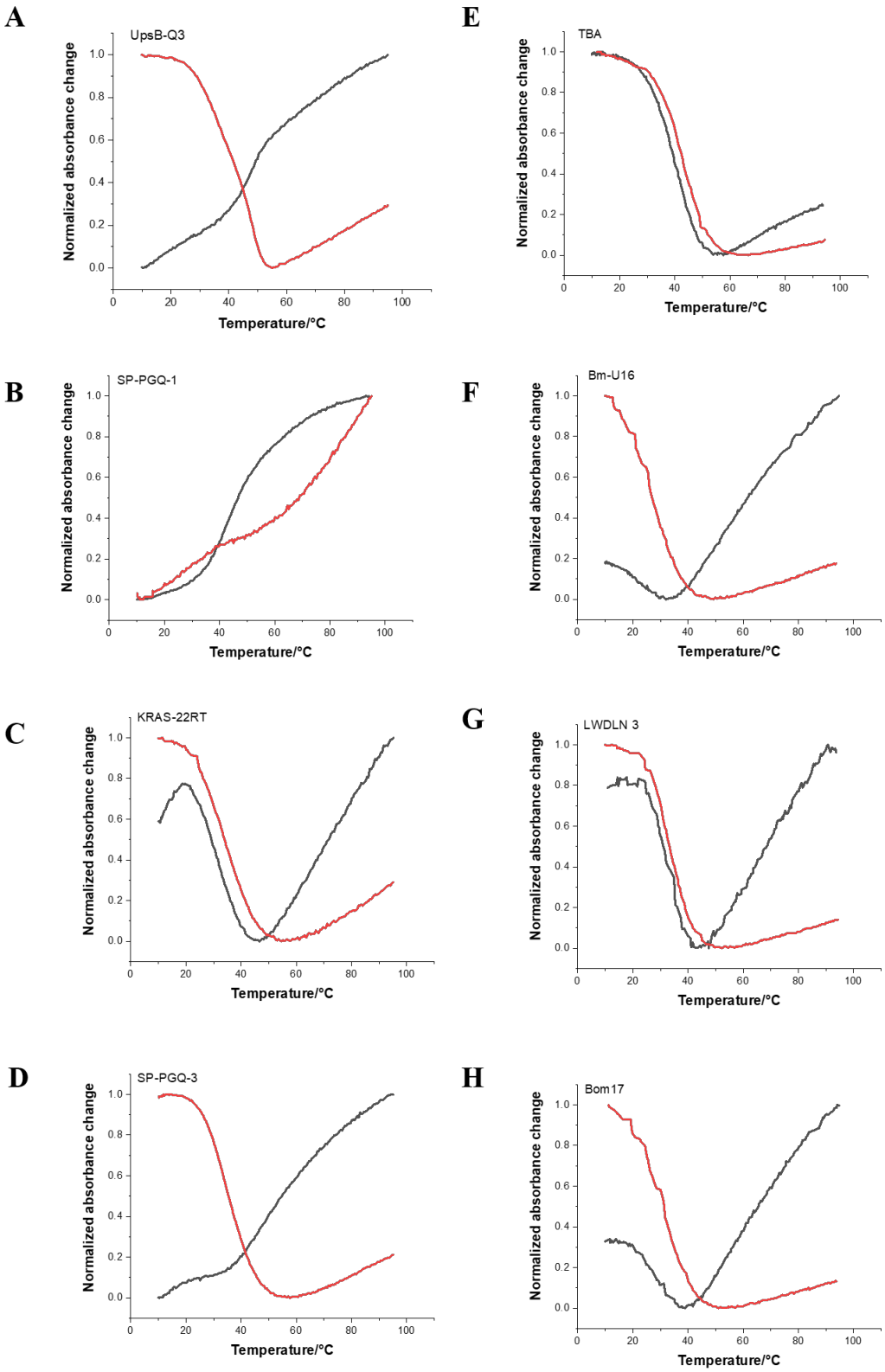

Fig. S4. Normalized UV-melting curves for some sequences of the training set. All data collected in 10-95 °C, monitored at 260 nm (Blank lines) and 295 nm (Red lines).

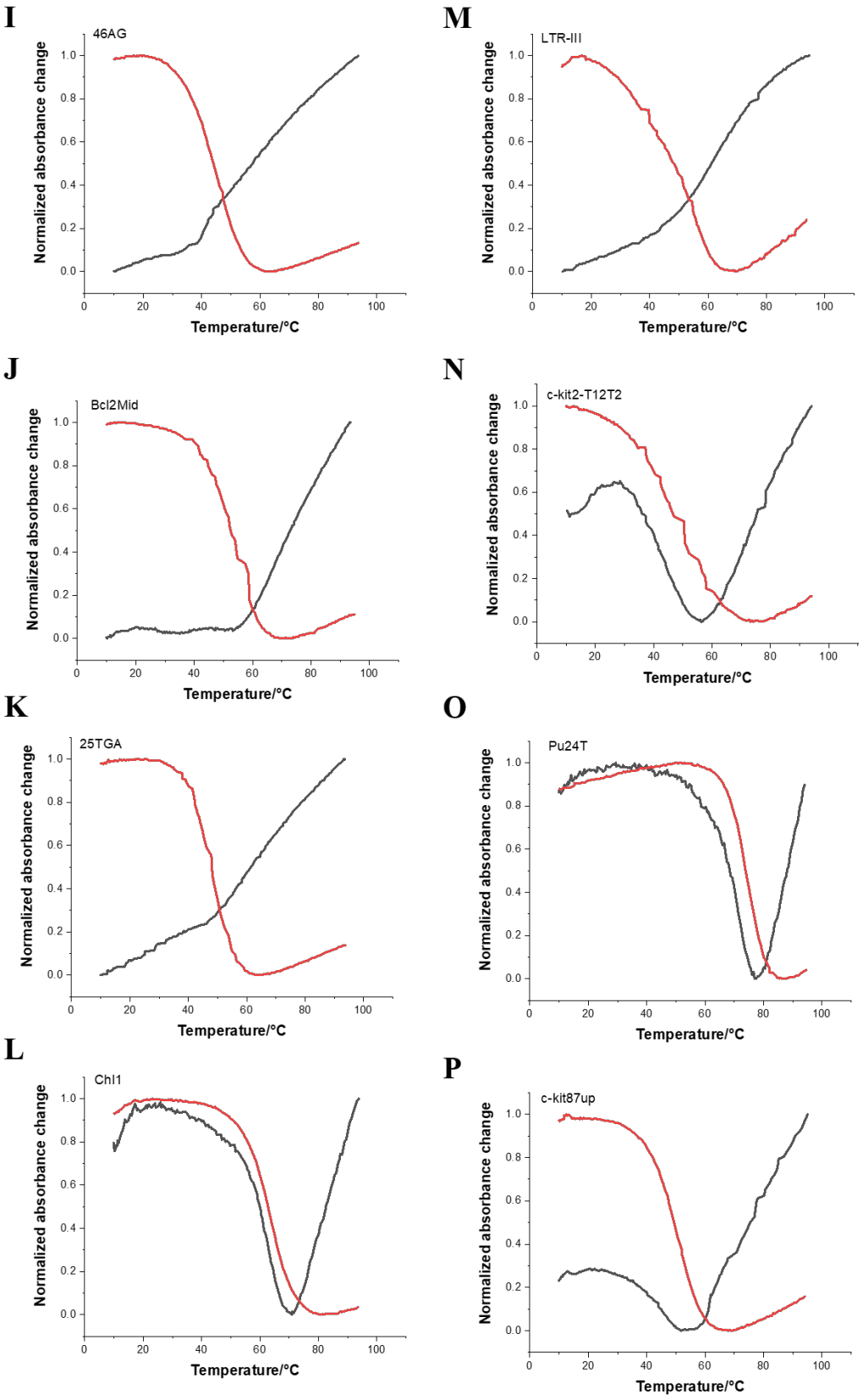

Fig. S4. *continued*.

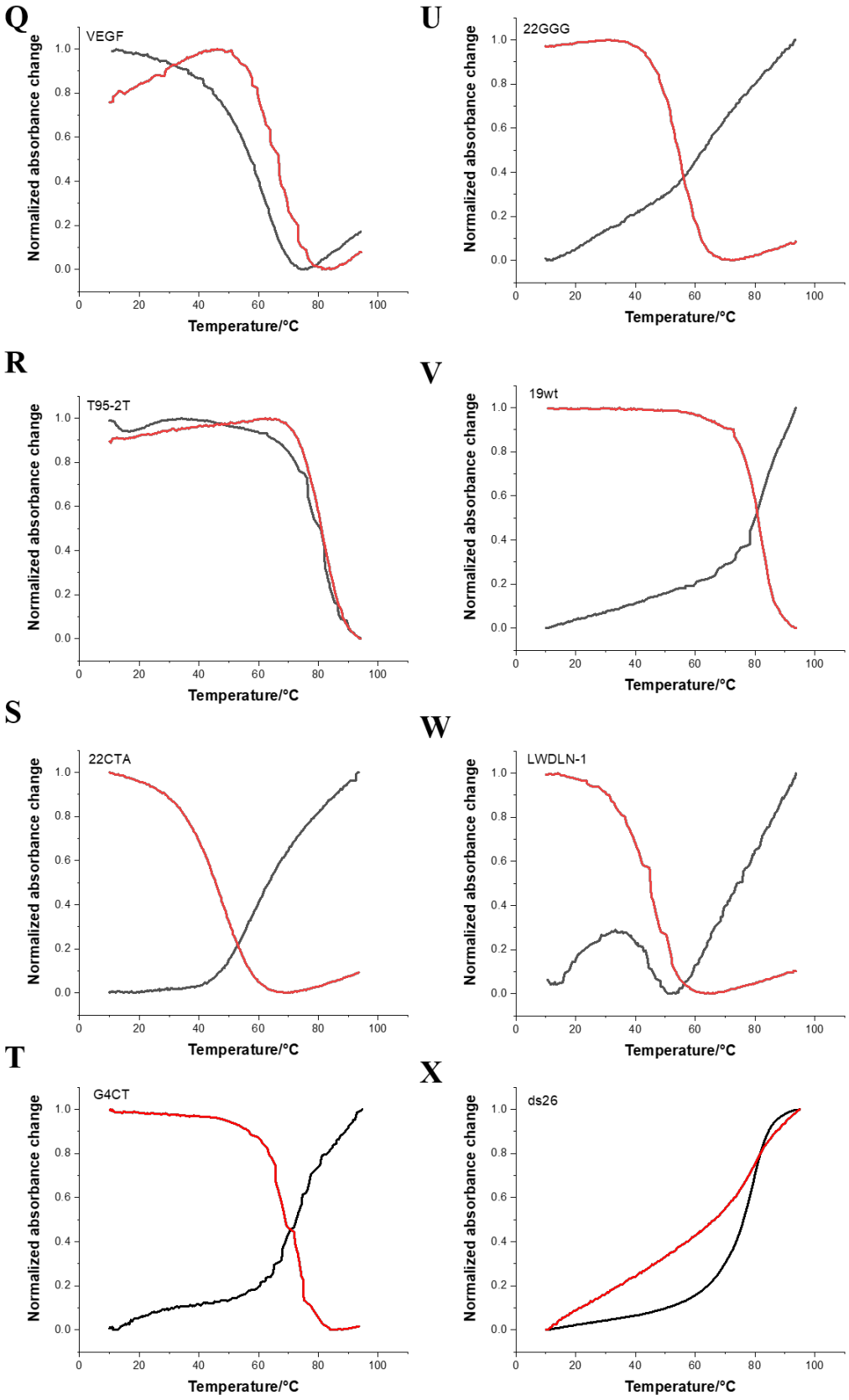

Fig. S4. *continued*.

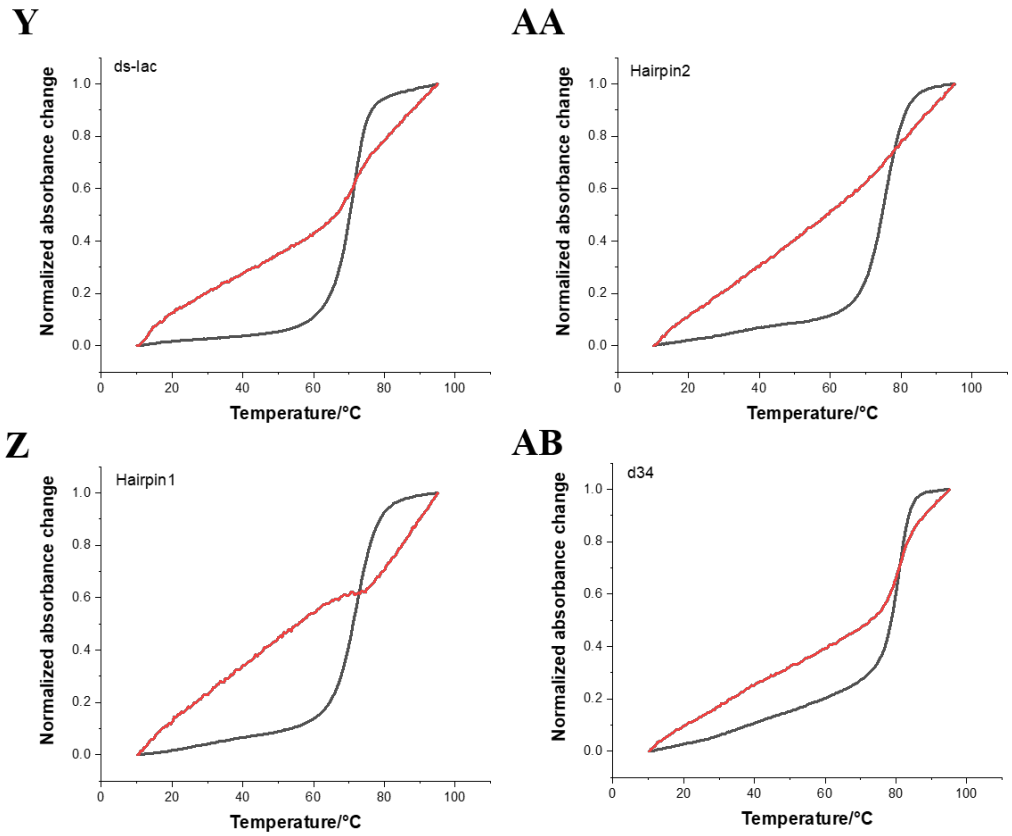

Fig. S4. *continued*.

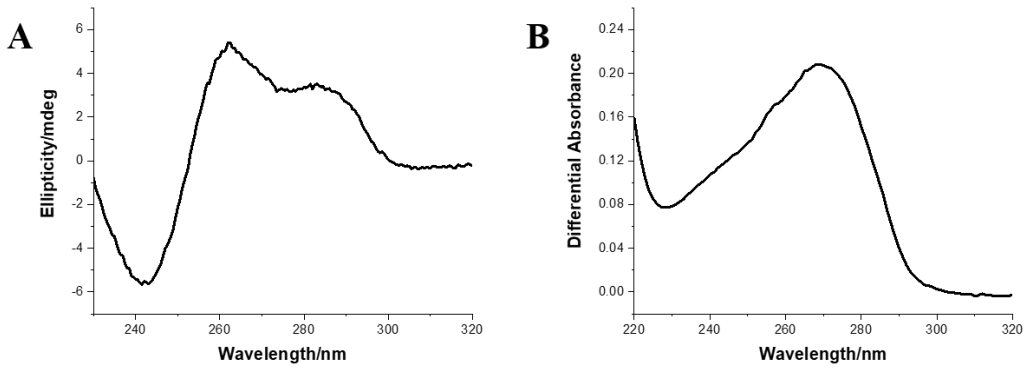

Fig. S5. CD **(A)** and thermal differential absorbance (TDS) **(B)** spectra for 3 μM SP-PGQ-1 in the FRET buffer.

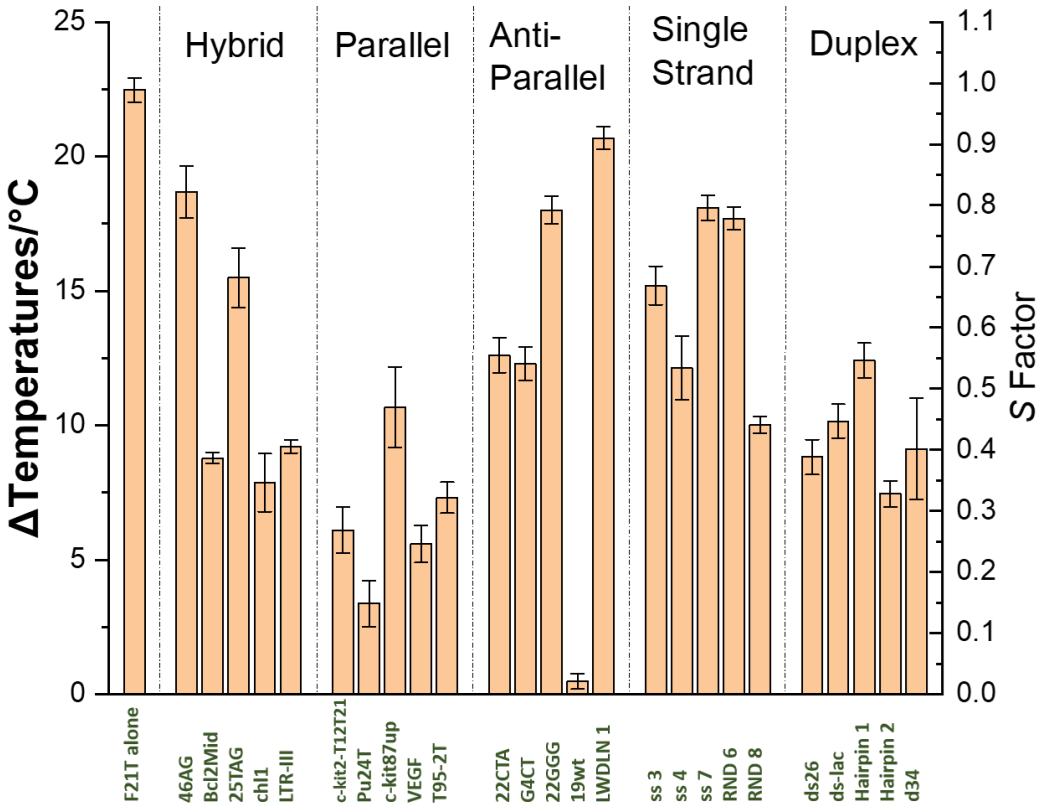

Fig. S6. ΔTm induced by 0.4 µM TMPyP4 on 0.2 µM F21T, alone or in the presence of 3 µM competitors. The *S Factor* is also provided on the right Y-axis. Samples were annealed and measured in 10 mM KCl, 90 mM LiCl, 10 mM lithium cacodylate pH 7.2 buffer.

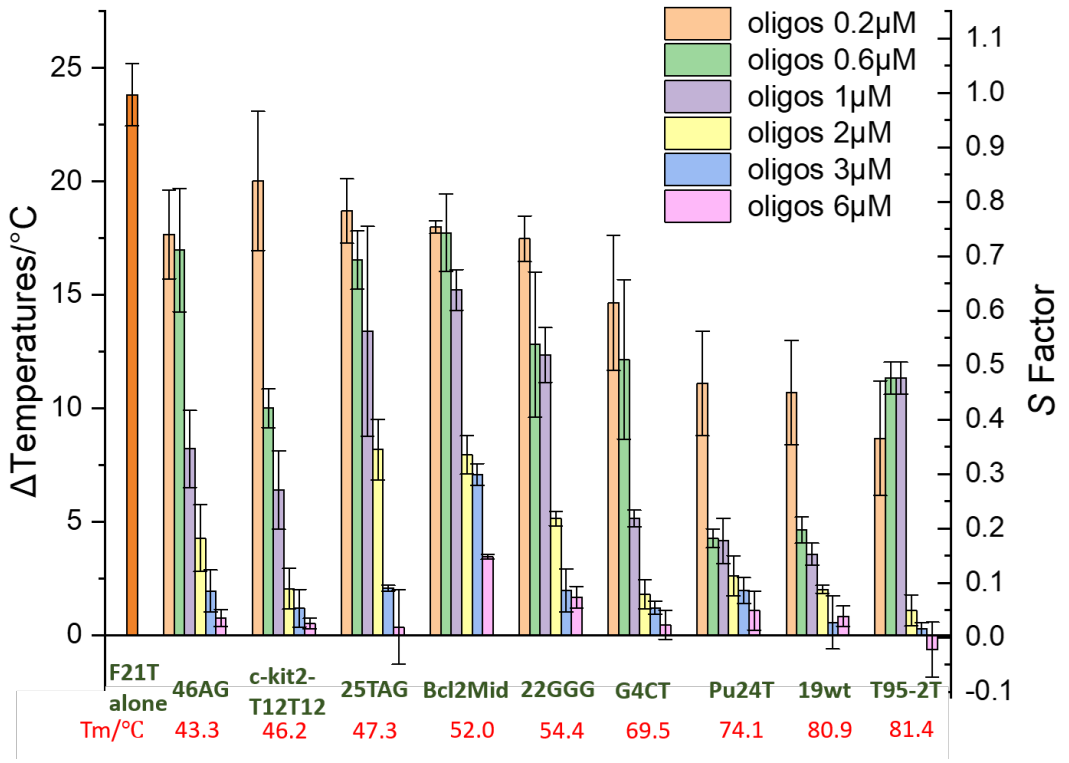

Fig. S7. ΔTm induced by 0.4 μM PhenDC3 on 0.2 μM F21T, alone or in the presence of 0.2 / 0.6 / 1 / 2 / 3 / 6 μM competitors (1x to 30x molar excess, as compared to F21T). The *S Factor* is also provided on the right Y-axis. Samples were annealed and measured in 10 mM KCl, 90 mM LiCl, 10 mM lithium cacodylate pH 7.2 buffer. The Tm for each quadruplex are shown in red below; sequences were ranked from left (lowest Tm) to right (highest Tm).

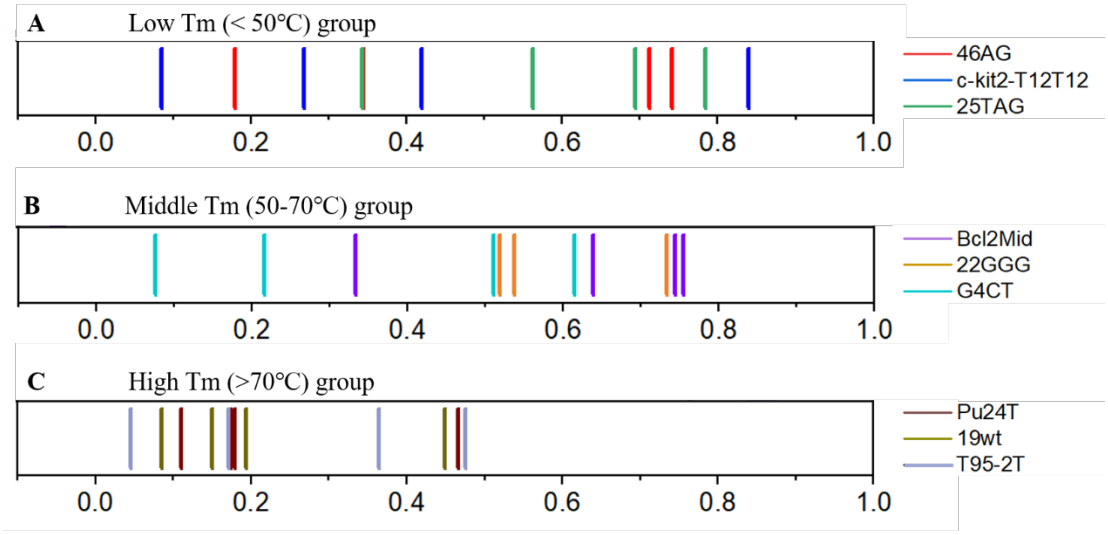

Fig. S8. Distribution of S-factor values for 0.2 / 0.6 / 1 / 2 μM competitor concentrations (1x to 10x molar excess, as compared to F21T). Oligonucleotides are grouped based on Tm: low Tm (< 50C°) **(A)**, middle Tm (50-70C°) **(B)** and high Tm (>70°C) **(C)**. The 4 *S* values shown for each competitor are depicted with the same color; the lowest *S* value (leftmost vertical bar) corresponds to the highest (2 μM) concentration while the highest *S* value (rightmost bar) corresponds to the lowest (0.2 μM) concentration

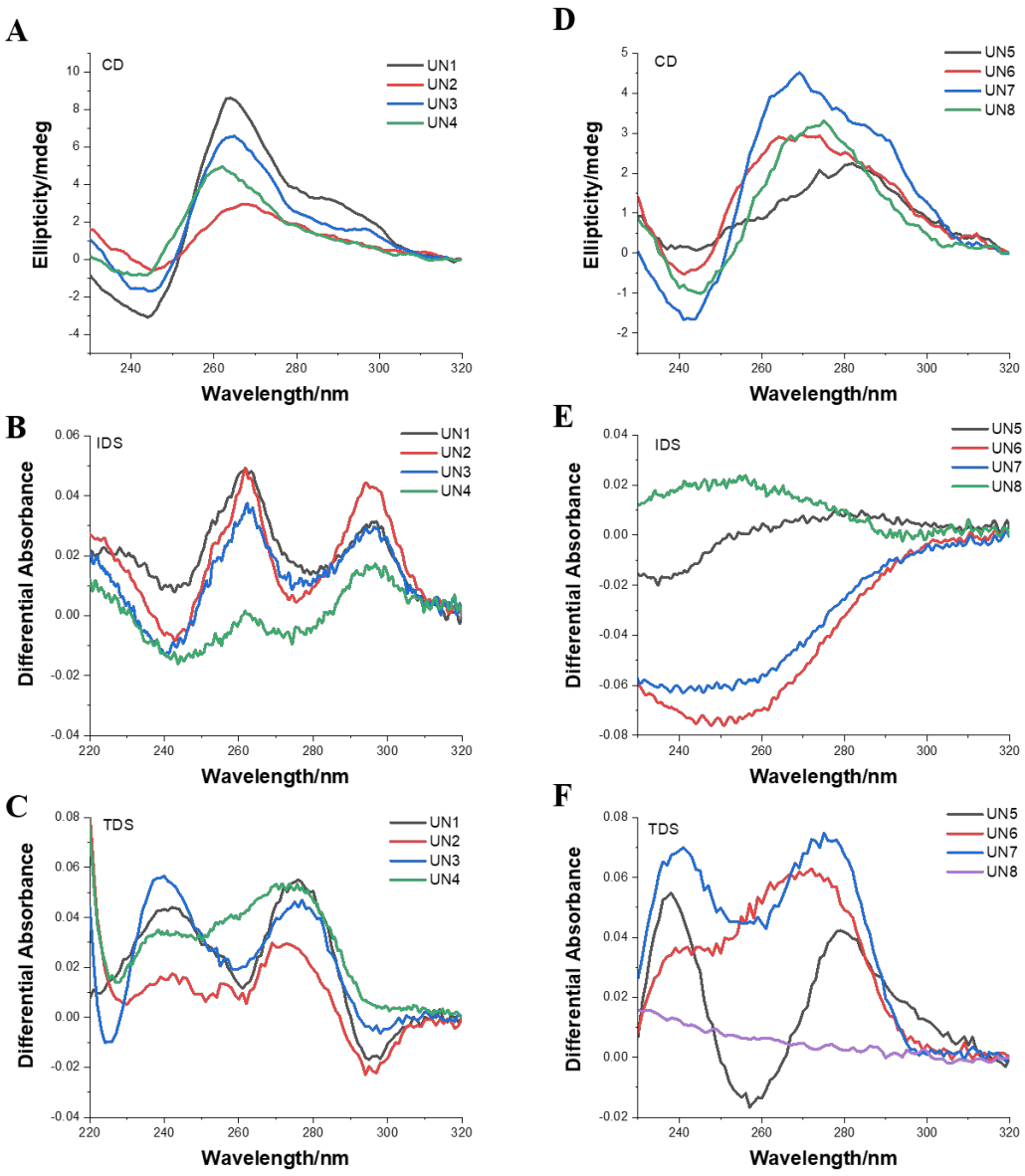

Fig. S9. CD, IDS, and TDS spectra of testing set. Panels **A, B, C** correspond to positive sequences (absolute *S* values < 0.1) and panels **D, E, F** to negative sequences, unable to compete (S > 0.8). From top to bottom: CD spectra, IDS spectra, and TDS spectra.