



HAL
open science

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

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

► **To cite this version:**

Yu Luo, Anton Granzhan, Daniela Verga, Jean-Louis Mergny. FRET-MC: a fluorescence melting competition assay for studying G4 structures in vitro. *Biopolymers*, 2021, 112 (4), pp.e23415. 10.1002/bip.23415 . hal-03294475

HAL Id: hal-03294475

<https://hal.science/hal-03294475v1>

Submitted on 21 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

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

Yu Luo^{1,2}, Anton Granzhan¹, 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

31 1. Introduction

32 G-quadruplexes (G4) are four-stranded nucleic acid structures adopted by G-rich DNA and
33 RNA sequences. G4 result from the stacking of two or more G-quartets (also called G-tetrads),
34 which are formed by four guanine bases interacting through Hoogsteen hydrogen bonds¹. G4
35 structures have been widely used to design biosensors to detect specific small molecules²⁻⁵ and
36 to control the assembly of supramolecular DNA complexes⁶⁻⁹. G4 structures also exist *in vivo*
37 and play important roles in cells^{10,11}. For example, G4s can be formed at human telomeres and
38 G4 ligands may interfere with telomeric functions, leading to telomere shortening and/or
39 uncapping¹²⁻¹⁴. G4s are also found in the promoter regions of genes critical in cancer, including
40 *KRAS*, *BCL2* and *VEGF*¹⁵⁻¹⁷: *BCL2* plays an essential role in cell survival; *VEGF* is a key
41 angiogenic growth factor which contributes to angiogenesis and tumor progression, and *KRAS*
42 is one of the most frequently mutated oncogenes in many signal transduction pathways, relevant
43 for different types of human carcinomas¹⁶.

44 Given the importance of G4 structures in biology and nanotechnology¹⁸, algorithms such as
45 G4Hunter^{19,20} have been developed to predict if a specific sequence is G4-prone. However, for
46 most DNA or RNA motifs, experimental validation is required and, for this purpose, a number
47 of biophysical methods have been developed to characterize G4 structures *in vitro*. Some
48 classical methods are based on the physical properties of G4 structures, such as UV-melting at
49 295 nm²¹, nuclear magnetic resonance (NMR)²², circular dichroism (CD) spectroscopy²³,
50 isothermal difference spectroscopy (IDS) and thermal difference spectroscopy (TDS)²⁴.
51 Fluorescence light-up assays employing dyes such as Thioflavin T²⁵, *N*-methylmesoporphyrin
52 IX (NMM)²⁶, tailor-made dyes²⁷ or combinations of dyes^{28,29} are also used to evidence G4
53 formation. Although there is a wide range of choices to check if a quadruplex is formed or not,
54 this validation is not always straightforward; therefore, a combination of techniques is often
55 required to unambiguously establish whether a sequence folds into a quadruplex or not¹⁹.

56 High-affinity G4 ligands can stabilize a G4 structure and alter its biological functions³⁰. Typical
57 assays used to characterize these ligands are the Fluorescent Intercalator Displacement (FID)
58 assay³¹ and the fluorescence-based Förster Resonance Energy Transfer (FRET)-melting assay

59 ^{32, 33}. The FRET-melting assay is based on the stabilization induced by a quadruplex ligand,
60 leading to a difference in melting temperature (T_m) between the nucleic acid alone and in
61 presence of this ligand ³⁴: in the presence of the latter, the thermal stability of the structure
62 increases in a concentration-dependent manner. This FRET-melting assay has been extensively
63 used to estimate whether a compound is a good quadruplex ligand or not ³⁴, with some biases
64 when ranking ligands potency using melting experiments ³⁵. More recently, this assay was
65 adapted to assess G4 ligands in near-physiological conditions ³⁶.

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

78

79 2. Materials and Methods

80 2.1 Samples

81 Oligonucleotides were purchased from Eurogentec, Belgium, as dried samples: unmodified
82 oligonucleotides were purified by RP cartridge while a dual-labeled F21T was purified by RP-
83 HPLC. Stock solutions were prepared at 100 μM strand concentration for the unlabeled
84 oligonucleotide and at 200 μM strand concentration for F21T in ddH₂O. 50 μM oligonucleotide
85 solutions were annealed in the corresponding buffer, kept at 95 °C for 5 min and slowly cooled
86 to room temperature before measurement. The FRET buffer contains 10 mM KCl, 90 mM LiCl,
87 10 mM lithium cacodylate, pH 7.2. The K-100 buffer contains 100 mM KCl, 10 mM lithium
88 cacodylate, pH 7.2.

89 2.2 FRET-melting competition assay

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

99 ΔT_m is determined as the difference in T_m with the sample containing F21T in the absence of
100 PhenDC3. The T_m of an oligonucleotide is defined as the temperature at which 50% of the
101 oligonucleotide is unfolded. The most common method to obtain T_m values is approximated
102 as corresponding to half of the height at the normalized melting curve⁴⁰.

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

106 are analyzed one by one. The DoseResp Function in Origin Pro package allows collect Tm in
107 batch. As shown in **Fig. S1A**, Log X₀ in fitting curve is the Tm. This method can only be used
108 in curve (i) (see results) as non-linear fitting would fail if there is no high plateau at row curve.
109 In general, the difference of Tm calculated by these two methods is very small: when using
110 both the ‘1/2 height’ and non-linear fitting methods to calculate Tm of F21T alone, the results
111 are 59.8°C and 59.4°C, respectively (**Fig. S1B**). In a few rare cases that should be noted (**Fig.**
112 **S1C**), curves are irregular, and the minimum of X-axis (usually 25°C in FRET-melting assay)
113 does not correspond to y=0 for the normalized curve. In this case, using ‘1/2 height’ is more
114 accurate; Tm calculated by DoseResp always referenced the X value corresponding to Y=0.5,
115 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,
116 1/2 height at Y-axis is 0.525, Tm calculated by ‘1/2 height’ is 53.0°C, to be compared with 51.8°C
117 by DoseResp.

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

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

124 2.3 UV-melting assay

125 UV-melting curves of 5 μM oligonucleotides in FRET buffer were recorded with a Cary 300
126 (Agilent Technologies, France) spectrophotometer. Heating runs were performed between
127 10°C and 95 °C, the temperature was increased by 0.2 °C per minute, and absorbance was
128 recorded at 260 nm and 295 nm. Tm was determined as the temperature corresponding to half
129 of the height of the normalized melting curve.

130 2.4 Circular dichroism

131 3 μM SP-PGQ-1 was kept in 1000 μL FRET buffer. 3 μM Oligonucleotides of testing set were
132 kept in 1000 μL K-100 buffer. CD spectra were recorded on a JASCO J-1500 (France)
133 spectropolarimeter at room temperature, using a scan range of 400–230 nm, a scan rate of 200
134 nm/min and averaging four accumulations.

135 2.5 Thermal difference spectra

136 3 μM SP-PGQ-1 was kept in FRET buffer. 3 μM Oligonucleotides of testing set were kept in
137 1000 μL K-100 buffer. Absorbance spectra were recorded on a Cary 300 (Agilent Technologies,
138 France) spectrophotometer at 25 $^{\circ}\text{C}$ (scan range: 500-200 nm; scan rate: 600 nm/min; automatic
139 baseline correction). After recording the first spectra (folded), temperature was increased to
140 95 $^{\circ}\text{C}$, and the second UV-absorbance spectra was recorded after 15 min of equilibration at
141 high temperature. TDS corresponds to the arithmetic difference between the initial (folded;
142 25 $^{\circ}\text{C}$) and second (unfolded; 95 $^{\circ}\text{C}$) spectra.

143 2.6 Isothermal difference spectra

144 3 μM Oligonucleotides of testing set were kept in 900 μL K-100 buffer. Absorbance spectra
145 were recorded on a Cary 300 (Agilent Technologies, France) spectrophotometer at 25 $^{\circ}\text{C}$ (scan
146 range: 500-200 nm; scan rate: 600 nm/min; automatic baseline correction). 100 μL of 1M KCl
147 was added after recording the first spectrum, and the second UV-absorbance spectrum was
148 recorded after 15 min of equilibration. IDS correspond to the arithmetic difference between the
149 initial (unfolded) and second (folded, thanks to the addition of K^+) spectra, after correction for
150 dilution.

151

152 3. Results and Discussion

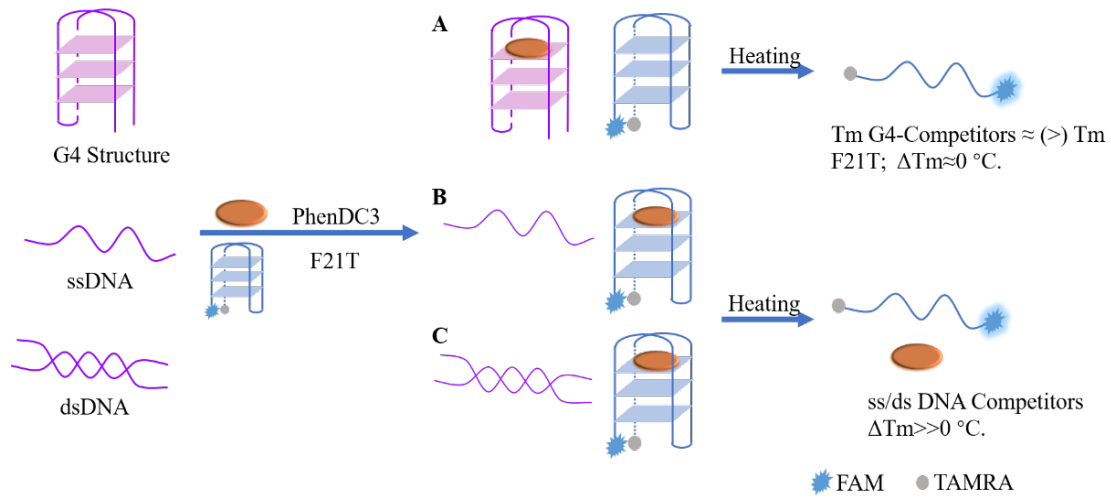
153 3.1 Principle of the FRET-MC assay

154 In a K^+ -containing buffer, F21T forms a stable G4 structure which can be used as a FRET-
155 melting probe thanks to the fluorophores attached at both ends: a 5'-appended donor
156 (Fluorescein) and a 3'-appended acceptor (TAMRA), allowing efficient energy transfer
157 between the donor and acceptor dyes when the oligonucleotide is folded (**Fig. 1**). Thermal
158 unfolding leads to the disruption of the G-quadruplex structure and a decrease in FRET
159 efficiency, as the 5' and 3' ends become distant when the sequence is single-stranded. In the
160 presence of a G4 ligand, the melting temperature (T_m) of F21T increases, as the ligand makes
161 the structure of F21T more stable. F21T is typically used in a FRET-melting assay, in which
162 the specificity of a ligand is tested by adding various specific (G4-forming) and unspecific
163 (duplexes or single-stranded) competitors³³. In this report, we are radically changing our
164 viewpoint: rather than testing ligands of unknown specificity against known competitors, we
165 are investigating one of the best-characterized G4-ligand, PhenDC3, against a variety of
166 oligonucleotide competitors. Adding a large excess of an unlabeled oligonucleotide may lead
167 to two possible scenarios:

- 168 (i) The competitor is unable to trap the quadruplex ligand. In this case, T_m of the (F21T
169 + PhenDC3) system is not affected by the competing oligonucleotide (in other words,
170 ΔT_m remains high). This is the expected outcome for a single-strand or a DNA or RNA
171 duplex.
- 172 (ii) If the unlabeled competitor has a high affinity for PhenDC3, it will sequester a
173 significant fraction of the compound, which will be no longer available for F21T
174 stabilization, leading to a decrease in T_m . In this case, if the competition is very
175 efficient, the T_m should fall back close to the value obtained without quadruplex ligand,
176 meaning for F21T alone (in other words, $\Delta T_m \approx 0$).

177 To proceed, we selected a variety of competitors for which the structure was previously
178 investigated and characterized⁴². This collection of over 60 sequences includes a variety of

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



181

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

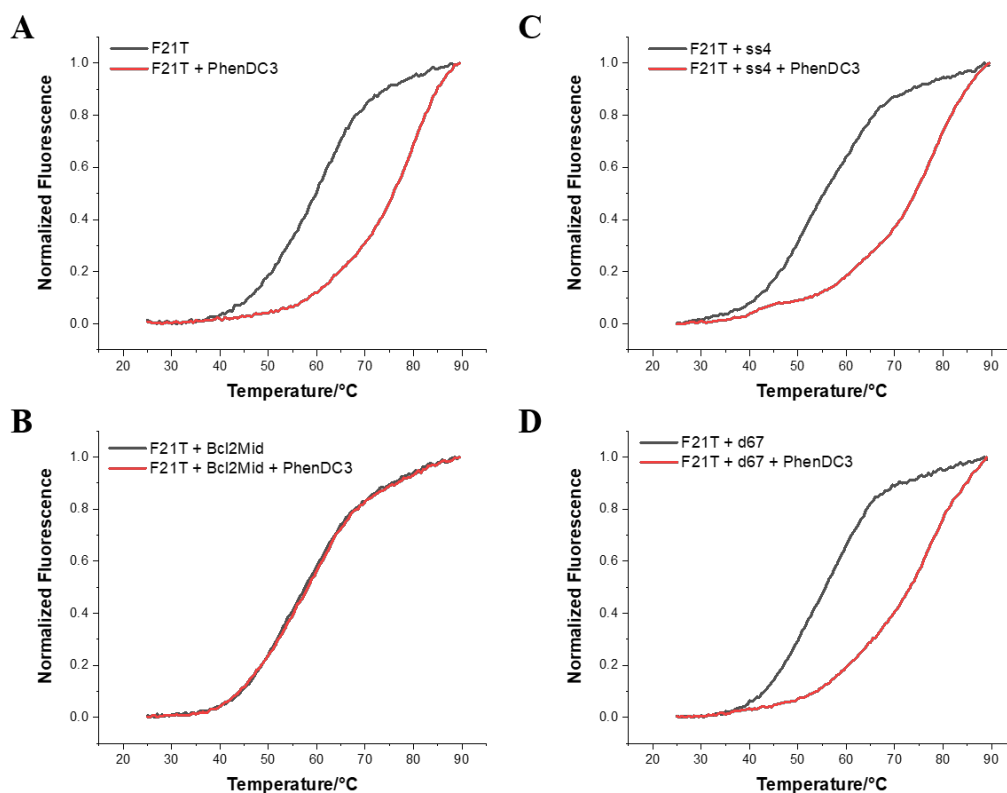
187

188 3.2 Validation of the FRET-melting competition assay with a set of 65 sequences.

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

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

197 **2A**), F21T in competition with a stable G4 (**Fig. 2B**), a single-strand (**Fig. 2C**) or a duplex DNA
198 (**Fig. 2D**), respectively.



199

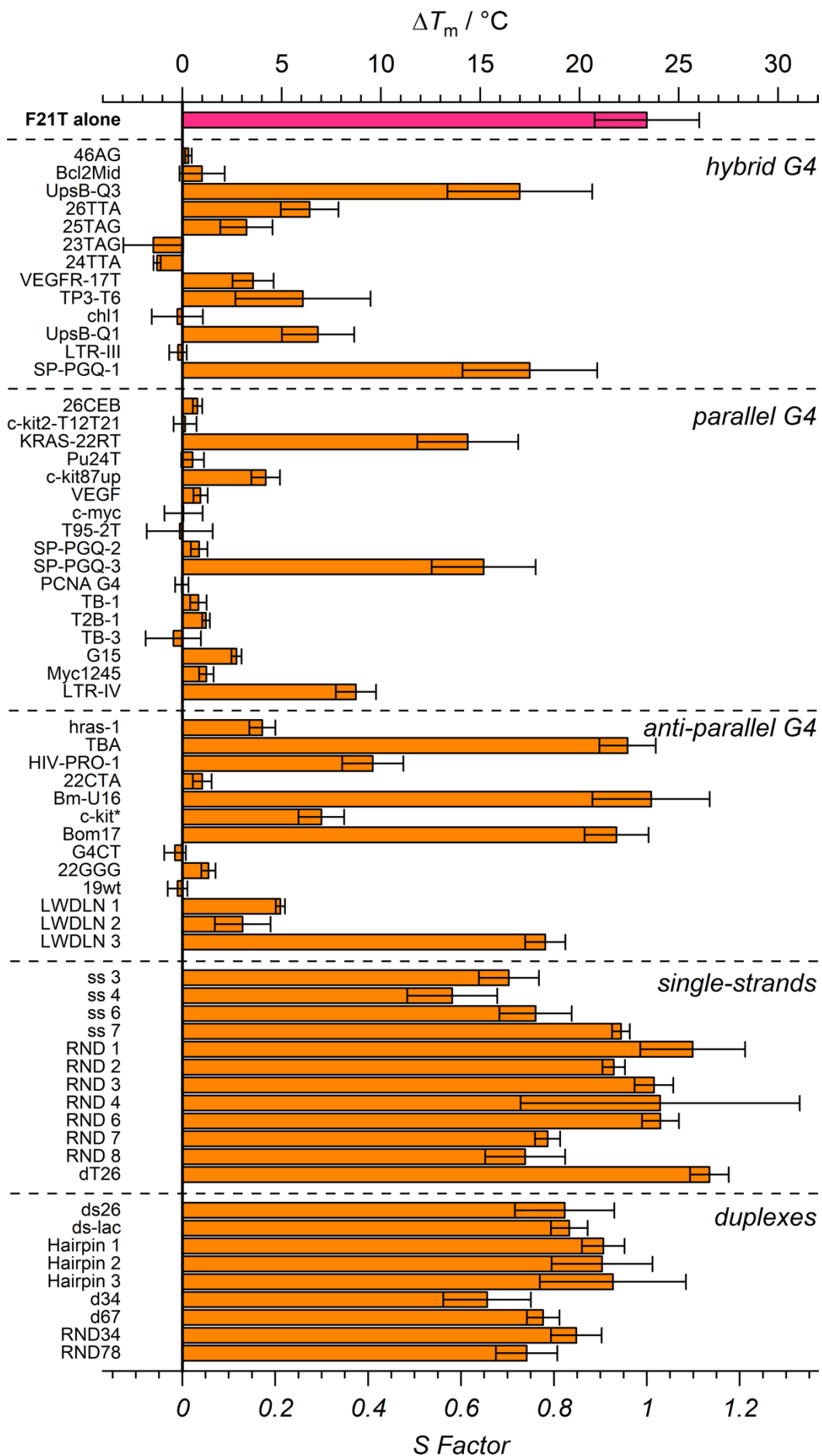
200 **Fig. 2. FRET-melting profiles of 0.2 μ M F21T alone or in the presence of 3 μ M competitors,**
201 **and with (red) or without (black) 0.4 μ M PhenDC3. (A) F21T alone and in the presence of**
202 **PhenDC3, and with the addition of the competitor: (B) G4 structure, (C) single-strand DNA**
203 **and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

205 In the absence of any competitor, 2 eqv. of PhenDC3 induces a ΔT_m of 23.4°C, in agreement
206 with previous results⁴³. As expected, none of the competitors induced a further significant
207 increase in T_m , as compared to F21T + PhenDC3. **Fig. 3** summarizes the ΔT_m results obtained
208 for F21T in the presence of PhenDC3 and in the presence or absence of oligonucleotide
209 competitor (normalized FRET-melting curves shown in **Fig. S3**). Many, but not all, of the
210 sequences known to form G4 structures led to a significant drop in ΔT_m values (upper half of
211 the figure), showing that they acted as efficient competitors. In contrast, single-stranded and

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

$$216 \quad S \text{ Factor} = \frac{\Delta T_m \text{ of F21T with competitors}}{\Delta T_m \text{ of F21T alone}}$$

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



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

225 As expected, the majority of G4-forming sequences led to S values close to 0. However, several
226 known G4 structures (UpsB-Q3, SP-PGQ-1, KRAS-22RT, SP-PGQ-3, TBA, Bm-U16, Bom17
227 and LWDLN3) were not efficient competitors ($S > 0.6$). In order to understand these results,
228 we performed UV-melting experiments for these 8 sequences and collected T_m values in **Table**
229 **S2** (detailed UV-melting curves are shown in **Fig. S4**). The thermal stability of all these
230 sequences (except for SP-PGQ-1) was relatively low, indicating they form unstable G4
231 structures, which are likely to be unfolded in the temperature range where F21T starts to melt:
232 they are then “seen” as non-specific single-strands rather than true G-quadruplexes. SP-PGQ-
233 1 behaved differently: UV-absorbance at 295 nm of a quadruplex should decrease upon heating
234 due to the unfolding of the quadruplex structure. Although SP-PGQ-1 was reported to form a
235 hybrid G4 structure⁴⁵, our results show an unexpected *increase* in absorbance at 295 nm upon
236 UV-melting, incompatible with the unfolding of a quadruplex, and rather suggesting the
237 formation of a another structure (*e.g.*, a mismatched duplex) at low temperatures. CD
238 spectroscopy and TDS confirmed this hypothesis (**Fig. S5**). In contrast, the T_m of 10 different
239 G4 sequences acting as effective competitors ($S < 0.3$ for 46AG, Bcl2Mid, 25TGA, Chl, LTR-
240 III, c-kit-T12T2, Pu24T, c-kit87up, VEGF and T95-2T) were always higher than the T_m of the
241 false negative sequences, as shown in **Table S2** (UV-melting curves showed in **Fig. S4**).
242 Although some of the T_m are still lower than the T_m of F21T, their presence in large excess
243 (15-fold molar excess as compared to F21T) may compensate for a partial denaturation.

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

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

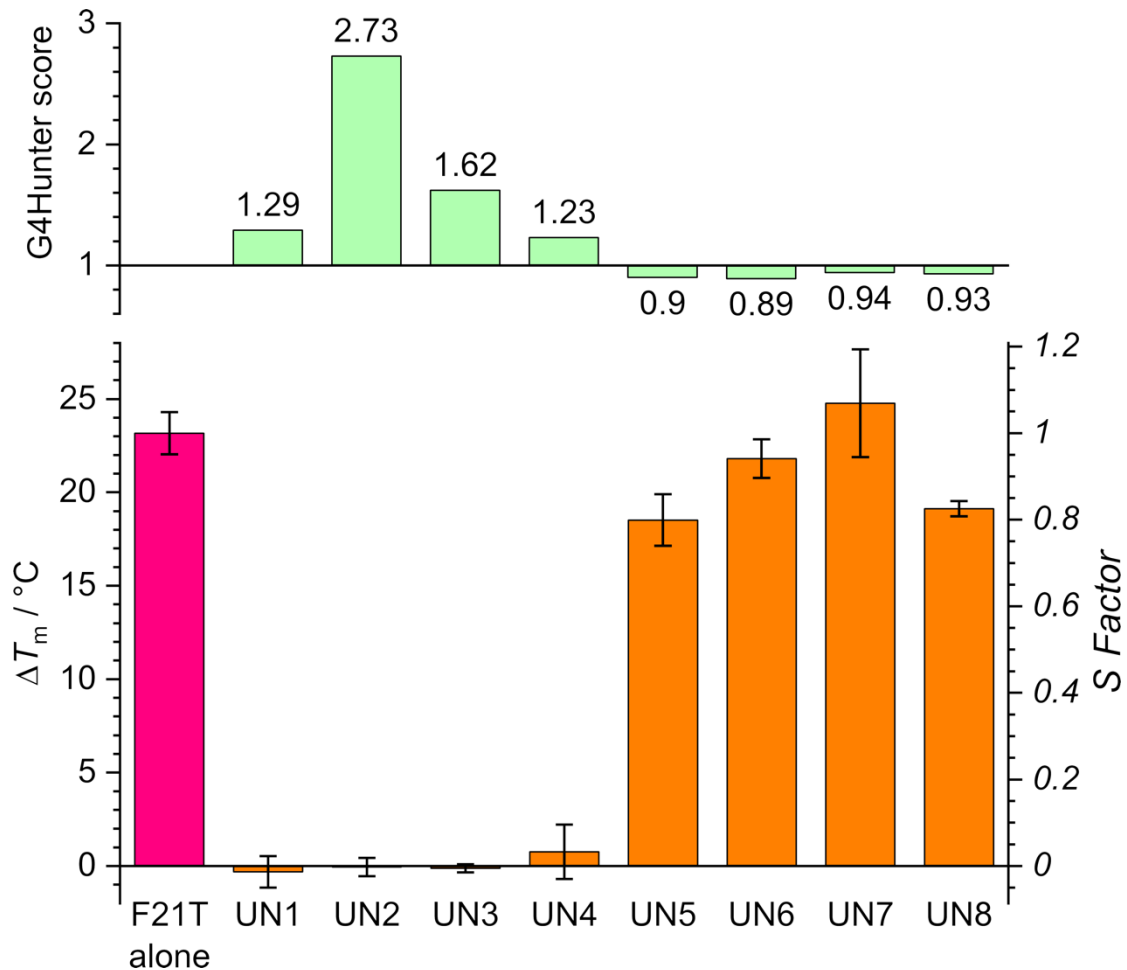
256 A critical factor for the competition efficiency (measured by the S value) in the FRET-MC
257 assay should be the affinity and number of binding sites present on the competitor sequence for
258 the quadruplex ligand (PhenDC3 in all further experiments). We wanted to investigate if the
259 thermal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

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

277 3.3 Experimental validation on novel sequences.

278 To test this method on a set of novel sequences, we used eight genomic DNA sequences of
279 unknown structure. These motifs are found close to the transcription start site of different
280 human promoters, and their biological relevance is currently being investigated. According to
281 G4Hunter analysis¹⁹, four of them (UN1–4) were likely to form a quadruplex structure, as their
282 G4Hunter score was above 1.2, while the remaining four (UN5–8) were unlikely to adopt a G4
283 conformation, with a G4Hunter score < 1.0. Detailed sequences with location information
284 (Human hg19) and G4Hunter scores are shown in **Table S4**.

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



292

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

299

300 4. Conclusion

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

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

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

309 - It is *fast*: FRET-melting takes 1-2 hours while UV-melting requires several hours (14
310 h with the temperature gradient used here) for each experiment.

311 - It is *simple* to set up: all reagents are commercially available (PhenDC3 included) and
312 the FRET melting assay is now routinely used in a number of labs.

313 - It allows testing *multiple samples in parallel*. While a classical UV spectrophotometer
314 can only read up to 6 or 9 samples, FRET-melting uses a 96-well plate as a sample
315 holder, and it is able to process 48 sequences in duplicate. It may even be transposed
316 to 384-well format.

317 At the same time, this method has several limitations:

318 - The main assumptions for this technique are that PhenDC3 *i)* indiscriminately binds to all G-
319 quadruplexes, and *ii)* does not bind to other structures. In other words, for this method to work,
320 we need a perfect, general G4 ligand with high structure specificity. Previously published
321 studies^{31,37} have shown that PhenDC3 does indeed bind to all G4 tested so far, and has excellent
322 specificity. We cannot exclude, however, that PhenDC3 would also recognize other unusual
323 motifs such as G-triplexes⁴⁶. Additional experiments are therefore required to reach a clear
324 conclusion for a given sequence: as previously stated¹⁹, we advocate the use of several
325 independent techniques to assess G4 formation. On the other hand, a high-affinity but poorly
326 selective quadruplex ligand such as TMPyP4 proved to be ineffective for this application.

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

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

339 - The main limitation of this FRET-melting competition assay is its inability to detect unstable
340 quadruplexes which behave as single-strands (**Fig. 3**) at the temperature where F21T starts to
341 melt. This assay should therefore be employed to identify moderately stable or highly stable
342 G4 structures. A possible way to circumvent this limitation would be to replace or complement
343 F21T by a quadruplex probe with lower stability⁴⁷. A two-quartet quadruplex such as the
344 thrombin binding aptamer (TBA) could be proposed, keeping in mind that these G4 are often
345 weaker binders for G4 ligands such as PhenDC3. In any case, unstable quadruplexes are less
346 likely to be biologically relevant⁴⁸ and unlikely to be identified by genome-wide methods such
347 as G4-seq⁴⁹ as extension is performed at 60 °C during Illumina sequencing.

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

353 Overall, despite the shortcomings listed above, the FRET-melting competition assay should
354 constitute an interesting addition to the *in vitro* “G4 characterization toolbox”.

355 Data availability

356 Raw data / melting profiles may be downloaded at

357 <https://data.mendeley.com/datasets/gspc9r73r5/1>.

358

359 Acknowledgments

360 We thank both reviewers for excellent suggestions, Laurent Lacroix (ENS, Paris) for helpful

361 discussions, and Corinne Landras Guetta and Marie-Paule Teulade-Fichou (Institut Curie,

362 Orsay) for a sample of PhenDC3. This manuscript is dedicated to the memory of Prof. Michael

363 J. Waring, with whom J.L.M. had interesting lively discussions about DNA ligands during his

364 sabbatical in France.

365

366 Conflict of interest

367 None

368

369 **References**

- 370 1. Gellert, M.; Lipsett, M. N.; Davies, D. R., Helix formation by guanylic acid. *Proc Natl*
371 *Acad Sci U S A* **1962**, *48* (12), 2013-8.
- 372 2. Guo, Y.; Xu, L.; Hong, S.; Sun, Q.; Yao, W.; Pei, R., Label-free DNA-based
373 biosensors using structure-selective light-up dyes. *Analyst* **2016**, *141* (24), 6481-6489.
- 374 3. He, H. Z.; Chan, D. S.; Leung, C. H.; Ma, D. L., G-quadruplexes for luminescent
375 sensing and logic gates. *Nucleic Acids Res* **2013**, *41* (8), 4345-59.
- 376 4. Ren, J.; Wang, T.; Wang, E.; Wang, J., Versatile G-quadruplex-mediated strategies in
377 label-free biosensors and logic systems. *Analyst* **2015**, *140* (8), 2556-72.
- 378 5. Ruttkay-Nedecky, B.; Kudr, J.; Nejdil, L.; Maskova, D.; Kizek, R.; Adam, V., G-
379 quadruplexes as sensing probes. *Molecules* **2013**, *18* (12), 14760-79.
- 380 6. Lustgarten, O.; Carmieli, R.; Motiei, L.; Margulies, D., A Molecular Secret Sharing
381 Scheme. *Angew Chem Int Ed Engl* **2019**, *58* (1), 184-188.
- 382 7. Mergny, J. L.; Sen, D., Correction to DNA Quadruple Helices in Nanotechnology. *Chem*
383 *Rev* **2020**.
- 384 8. Stefan, L.; Monchaud, D., Applications of guanine quartets in nanotechnology and
385 chemical biology. *Nature Reviews Chemistry* **2019**, *3* (11), 650-668.
- 386 9. Cao, Y.; Gao, S.; Yan, Y.; Bruist, M. F.; Wang, B.; Guo, X., Assembly of
387 supramolecular DNA complexes containing both G-quadruplexes and i-motifs by enhancing
388 the G-repeat-bearing capacity of i-motifs. *Nucleic Acids Res* **2017**, *45* (1), 26-38.
- 389 10. Spiegel, J.; Adhikari, S.; Balasubramanian, S., The Structure and Function of DNA G-
390 Quadruplexes. *Trends Chem* **2020**, *2* (2), 123-136.
- 391 11. Varshney, D.; Spiegel, J.; Zyner, K.; Tannahill, D.; Balasubramanian, S., The
392 regulation and functions of DNA and RNA G-quadruplexes. *Nat Rev Mol Cell Biol* **2020**, *21*
393 (8), 459-474.
- 394 12. Burger, A. M.; Dai, F.; Schultes, C. M.; Reszka, A. P.; Moore, M. J.; Double, J.
395 A.; Neidle, S., The G-Quadruplex-Interactive Molecule BRACO-19 Inhibits Tumor Growth,
396 Consistent with Telomere Targeting and Interference with Telomerase Function. *Cancer*
397 *Research* **2005**, *65* (4), 1489.
- 398 13. Read, M.; Harrison, R. J.; Romagnoli, B.; Tanius, F. A.; Gowan, S. H.; Reszka,
399 A. P.; Wilson, W. D.; Kelland, L. R.; Neidle, S., Structure-based design of selective and
400 potent G quadruplex-mediated telomerase inhibitors. *Proceedings of the National Academy of*
401 *Sciences* **2001**, *98* (9), 4844.
- 402 14. Riou, J. F.; Guittat, L.; Mailliet, P.; Laoui, A.; Renou, E.; Petitgenet, O.;
403 Mègnin-Chanet, F.; Hélène, C.; Mergny, J. L., Cell senescence and telomere shortening
404 induced by a new series of specific G-quadruplex DNA ligands. *Proceedings of the National*
405 *Academy of Sciences USA* **2002**, *99* (5), 2672.

- 406 15. Brooks, T. A.; Kendrick, S.; Hurley, L., Making sense of G-quadruplex and i-motif
407 functions in oncogene promoters. *FEBS J* **2010**, *277* (17), 3459-69.
- 408 16. Jana, J.; Mondal, S.; Bhattacharjee, P.; Sengupta, P.; Roychowdhury, T.; Saha,
409 P.; Kundu, P.; Chatterjee, S., Chelerythrine down regulates expression of VEGFA, BCL2 and
410 KRAS by arresting G-Quadruplex structures at their promoter regions. *Sci Rep* **2017**, *7*, 40706.
- 411 17. Marquevielle, J.; Robert, C.; Lagrabette, O.; Wahid, M.; Bourdoncle, A.; Xodo,
412 L. E.; Mergny, J. L.; Salgado, G. F., Structure of two G-quadruplexes in equilibrium in the
413 KRAS promoter. *Nucleic Acids Res* **2020**, *48* (16), 9336-9345.
- 414 18. Spiegel, J.; Adhikari, S.; Balasubramanian, S., The Structure and Function of DNA G-
415 Quadruplexes. *Trends in Chemistry* **2020**, *2* (2), 123-136.
- 416 19. Bedrat, A.; Lacroix, L.; Mergny, J. L., Re-evaluation of G-quadruplex propensity with
417 G4Hunter. *Nucleic Acids Res* **2016**, *44* (4), 1746-59.
- 418 20. Brazda, V.; Kolomaznik, J.; Lysek, J.; Bartas, M.; Fojta, M.; Stastny, J.; Mergny,
419 J. L., G4Hunter web application: a web server for G-quadruplex prediction. *Bioinformatics*
420 **2019**, *35* (18), 3493-3495.
- 421 21. Alba, J. J.; Sadurni, A.; Gargallo, R., Nucleic Acid i-Motif Structures in Analytical
422 Chemistry. *Crit Rev Anal Chem* **2016**, *46* (5), 443-54.
- 423 22. Adrian, M.; Heddi, B.; Phan, A. T., NMR spectroscopy of G-quadruplexes. *Methods*
424 **2012**, *57* (1), 11-24.
- 425 23. Kypr, J.; Kejnovska, I.; Renciuik, D.; Vorlickova, M., Circular dichroism and
426 conformational polymorphism of DNA. *Nucleic Acids Res* **2009**, *37* (6), 1713-25.
- 427 24. Mergny, J. L.; Li, J.; Lacroix, L.; Amrane, S.; Chaires, J. B., Thermal difference
428 spectra: a specific signature for nucleic acid structures. *Nucleic Acids Res* **2005**, *33* (16), e138.
- 429 25. Renaud de la Faverie, A.; Guedin, A.; Bedrat, A.; Yatsunyk, L. A.; Mergny, J. L.,
430 Thioflavin T as a fluorescence light-up probe for G4 formation. *Nucleic Acids Res* **2014**, *42* (8),
431 e65.
- 432 26. Sabharwal, N. C.; Savikhin, V.; Turek-Herman, J. R.; Nicoludis, J. M.; Szalai, V.
433 A.; Yatsunyk, L. A., N-methylmesoporphyrin IX fluorescence as a reporter of strand orientation
434 in guanine quadruplexes. *FEBS J* **2014**, *281* (7), 1726-37.
- 435 27. Xie, X.; Renvoisé, A.; Granzhan, A.; Teulade-Fichou, M.-P., Aggregating
436 distyrylpyridinium dye as a bimodal structural probe for G-quadruplex DNA. *New Journal of*
437 *Chemistry* **2015**, *39* (8), 5931-5935.
- 438 28. Kreig, A.; Calvert, J.; Sanoica, J.; Cullum, E.; Tipanna, R.; Myong, S., G-
439 quadruplex formation in double strand DNA probed by NMM and CV fluorescence. *Nucleic*
440 *Acids Res* **2015**, *43* (16), 7961-70.
- 441 29. Zuffo, M.; Xie, X.; Granzhan, A., Strength in Numbers: Development of a Fluorescence
442 Sensor Array for Secondary Structures of DNA. *Chemistry* **2019**, *25* (7), 1812-1818.

- 443 30. O'Hagan, M. P.; Morales, J. C.; Galan, M. C., Binding and Beyond: What Else Can G-
444 Quadruplex Ligands Do? *European Journal of Organic Chemistry* **2019**, 2019 (31-32), 4995-
445 5017.
- 446 31. Tran, P. L.; Largy, E.; Hamon, F.; Teulade-Fichou, M. P.; Mergny, J. L.,
447 Fluorescence intercalator displacement assay for screening G4 ligands towards a variety of G-
448 quadruplex structures. *Biochimie* **2011**, 93 (8), 1288-96.
- 449 32. De Cian, A.; Guittat, L.; Kaiser, M.; Sacca, B.; Amrane, S.; Bourdoncle, A.;
450 Alberti, P.; Teulade-Fichou, M. P.; Lacroix, L.; Mergny, J. L., Fluorescence-based melting
451 assays for studying quadruplex ligands. *Methods* **2007**, 42 (2), 183-95.
- 452 33. Mergny, J.-L.; Maurizot, J.-C., Fluorescence Resonance Energy Transfer as a Probe for G-
453 Quartet Formation by a Telomeric Repeat. *ChemBioChem* **2001**, 2 (2), 124-132.
- 454 34. Renciuik, D.; Zhou, J.; Beaurepaire, L.; Guedin, A.; Bourdoncle, A.; Mergny, J. L.,
455 A FRET-based screening assay for nucleic acid ligands. *Methods* **2012**, 57 (1), 122-8.
- 456 35. Marchand, A.; Rosu, F.; Zenobi, R.; Gabelica, V., Thermal Denaturation of DNA G-
457 Quadruplexes and Their Complexes with Ligands: Thermodynamic Analysis of the Multiple
458 States Revealed by Mass Spectrometry. *J Am Chem Soc* **2018**, 140 (39), 12553-12565.
- 459 36. Morgan, R. K.; Psaras, A. M.; Lassiter, Q.; Raymer, K.; Brooks, T. A., G-quadruplex
460 deconvolution with physiological mimicry enhances primary screening: Optimizing the FRET
461 Melt(2) assay. *Biochim Biophys Acta Gene Regul Mech* **2020**, 1863 (1), 194478.
- 462 37. De Cian, A.; DeLemos, E.; Mergny, J.-L.; Teulade-Fichou, M.-P.; Monchaud, D.,
463 Highly Efficient G-Quadruplex Recognition by Bisquinolinium Compounds. *Journal of the*
464 *American Chemical Society* **2007**, 129 (7), 1856-1857.
- 465 38. Marchand, A.; Granzhan, A.; Iida, K.; Tsushima, Y.; Ma, Y.; Nagasawa, K.;
466 Teulade-Fichou, M. P.; Gabelica, V., Ligand-induced conformational changes with cation
467 ejection upon binding to human telomeric DNA G-quadruplexes. *J Am Chem Soc* **2015**, 137
468 (2), 750-6.
- 469 39. Ruggiero, E.; Richter, S. N., G-quadruplexes and G-quadruplex ligands: targets and tools
470 in antiviral therapy. *Nucleic Acids Res* **2018**, 46 (7), 3270-3283.
- 471 40. Mergny, J.-L.; Lacroix, L., UV Melting of G-Quadruplexes. *Current Protocols in Nucleic*
472 *Acid Chemistry* **2009**, 37 (1), 17.1.1-17.1.15.
- 473 41. Mergny, J.-L.; Lacroix, L., Analysis of Thermal Melting Curves. *Oligonucleotides* **2003**,
474 13 (6), 515-537.
- 475 42. Zuffo, M.; Gandolfini, A.; Heddi, B.; Granzhan, A., Harnessing intrinsic fluorescence
476 for typing of secondary structures of DNA. *Nucleic Acids Res* **2020**, 48 (11), e61.
- 477 43. Gueddouda, N. M.; Hurtado, M. R.; Moreau, S.; Ronga, L.; Das, R. N.;
478 Savrimoutou, S.; Rubio, S.; Marchand, A.; Mendoza, O.; Marchivie, M.; Elmi, L.;
479 Chansavang, A.; Desplat, V.; Gabelica, V.; Bourdoncle, A.; Mergny, J. L.; Guillon, J.,
480 Design, Synthesis, and Evaluation of 2,9-Bis[(substituted-aminomethyl)phenyl]-1,10-
481 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 Granzhan¹, 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	AGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTA 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	CGGGCGGGCCTTGGGCGGGGT	parallel G4	2M27
c-myc	TGAGGGTGGGTAGGGTGGGTAA	parallel G4	1XAV
T95-2T	TTGGGTGGGTGGGTGGGT	parallel G4	2LK7
SP-PGQ-2	GGGCTAGTGGGGGAGGGGG	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	TAATCGGAGCTCTTTGACACGTC	single strand	-
RND1	CTATACGAAAACCTTTTGTATCATT	single strand	-
RND2	AATGATACAAAAGGTTTTCGTATAG	single strand	-
RND3	TAACGTTTATAATGTAGTCTCATT	single strand	-
RND4	TAATGAGACTACATTATAAACGTTA	single strand	-
RND6	GTTGTCATTGCCCCGAATAATTCT	single strand	-
RND7	GCCTTGCGGAGGCATGCGTCATGCT	single strand	-
RND8	AGCATGACGCATGCCTCCGCAAGGC	single strand	-
dT26	TTTTTTTTTTTTTTTTTTTTTTTTTTT	single strand	-
ds26	CAATCGGATCGAATTCGATCCGATTG	duplex	-
ds-lac	GAATTGTGAGCGCTCACAATTC	duplex	-
Hairpin 1	GGATTCTTGGATTTTCCAAGAATCC	duplex	-
Hairpin 2	TCGGTATTGTGTTTACAATACCGA	duplex	-
Hairpin 3	AGGACGGTGTATTTTACACCGTCCT	duplex	-
d34	GTCGCCGGGCCAGTCGTCCATAC		-
	GTATGGACGACTGGCCCGGCGAC	duplex	-
d67	GACGTGTCGAAAGAGCTCCGATTA		-
	TAATCGGAGCTCTTTGACACGTC	duplex	-
RND34	TAACGTTTATAATGTAGTCTCATT		-
	TAATGAGACTACATTATAAACGTTA	duplex	-
RND78	AGAATTATTCGGGGGCAATGACAAC		-
	GTTGTCATTGCCCCGAATAATTCT	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 μ M (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	CGGGCAGGGAGGGCGGCTGTGCGGG GC	1.59	chr3: 196045150-196045176
UN2	TGGGGCGGGGAAGAGGGGCGGGG T	2.73	chr8: 57124113-57124138
UN3	CGGGAAGGGGCGGGCGCAATGGGC	1.62	chr17: 62915412-62915435
UN4	TGGGAGGCGGAGGTGGGCAGGTTGCT	1.23	chr17: 16258771-16258791
UN5	GTGCTGGGGCGCCCACTTCGGGGTGG TGC	0.90	chr11: 364605-364633
UN6	AGTTGGTAGGCTGAGGCGGGAGGATT GC	0.89	chr5: 64905453-64905472
UN7	AGGGCCGGGAGAGGGATCCGCCATAT TGGAGCTGGGGC	0.94	chr14: 36278253-36278290
UN8	AGGAAGCTGGGGTAGGAGAATTGCTTG A	0.93	chr12: 57876780-57876802

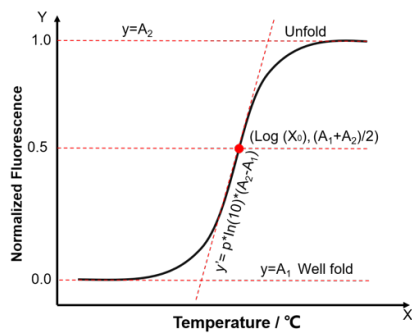
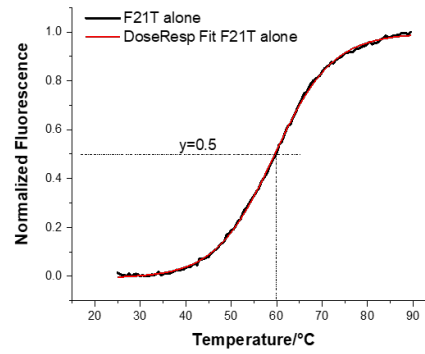
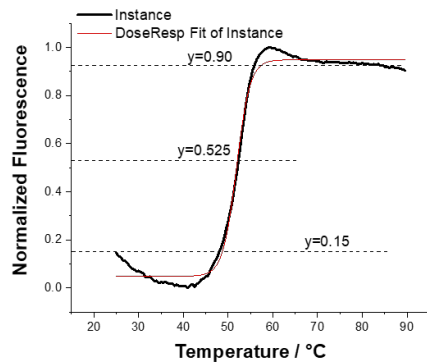
A**B****C**

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

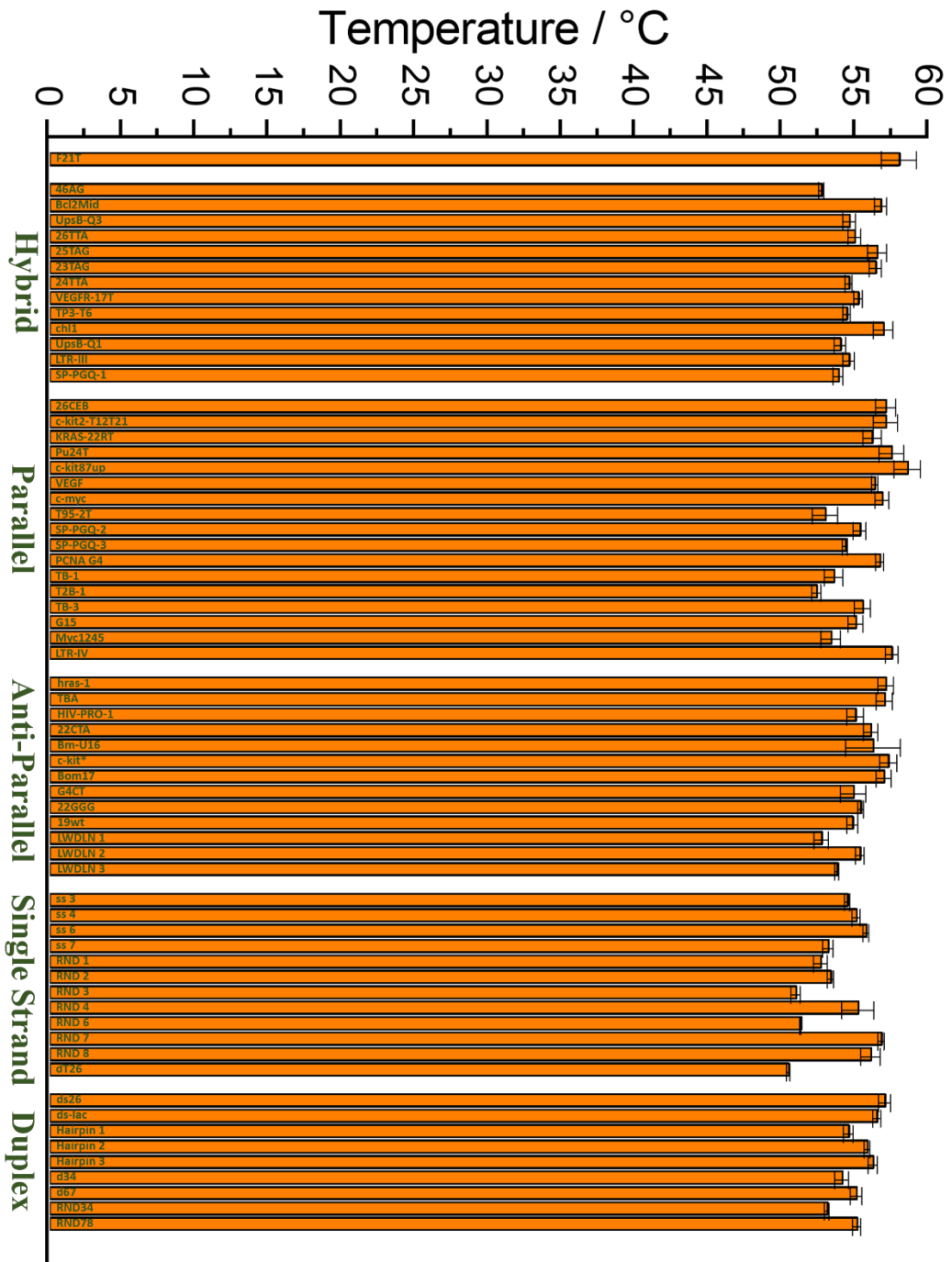


Fig. S2. FRET Tm of 0.2 μ M F21T alone (top) or in the presence of various competitors.

All competitors were tested at 3 μ M strand concentration.

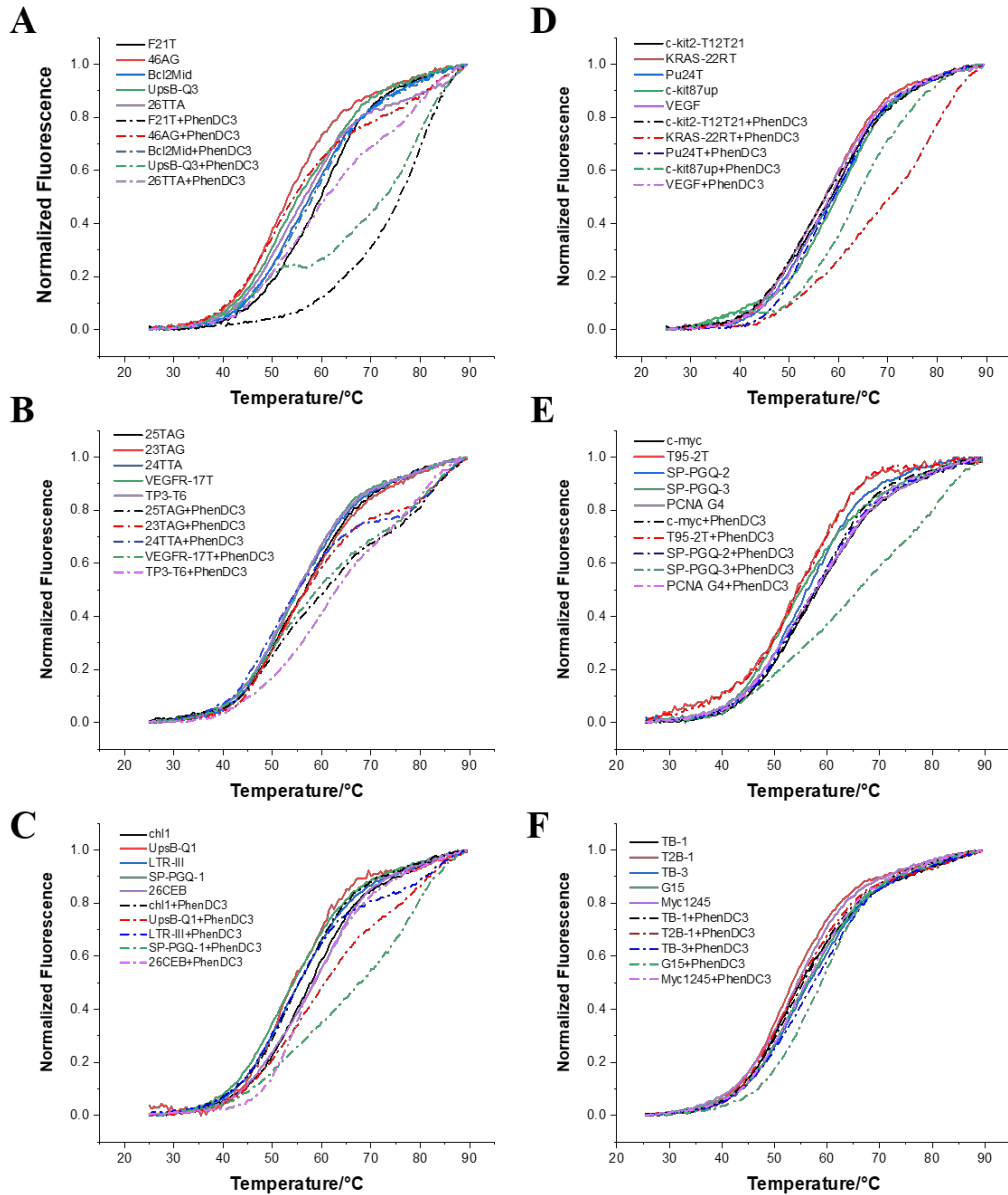


Fig. S3. Normalized FRET-melting curves of 0.2 μ M F21T in the presence of various competitors (3 μ M strand concentration) used in the training set, with or without 0.4 μ M PhenDC3. Samples were annealed and measured in 10 mM KCl, 90 mM LiCl, 10 mM lithium cacodylate pH 7.2 buffer.

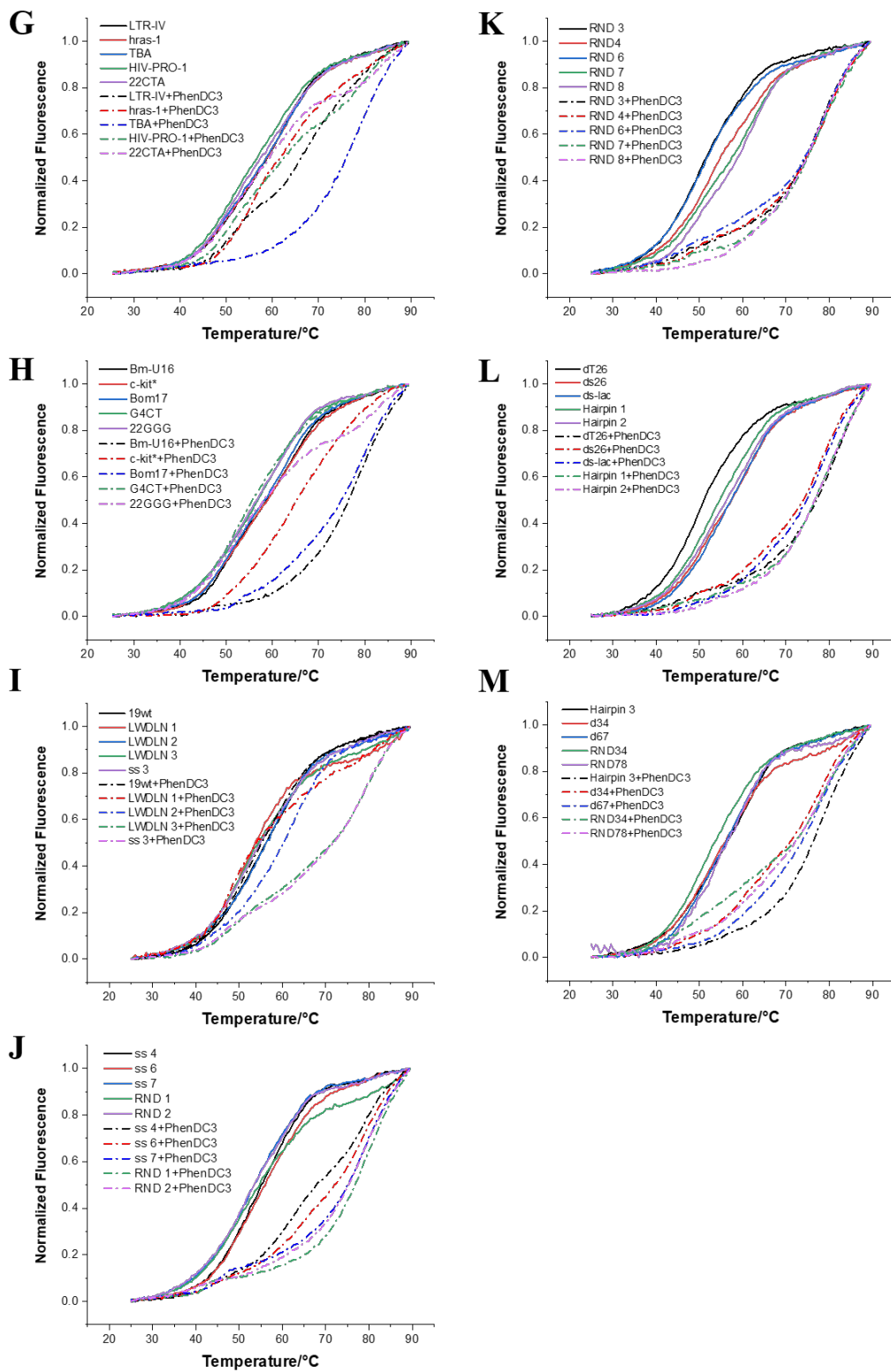


Fig. S3, continued.

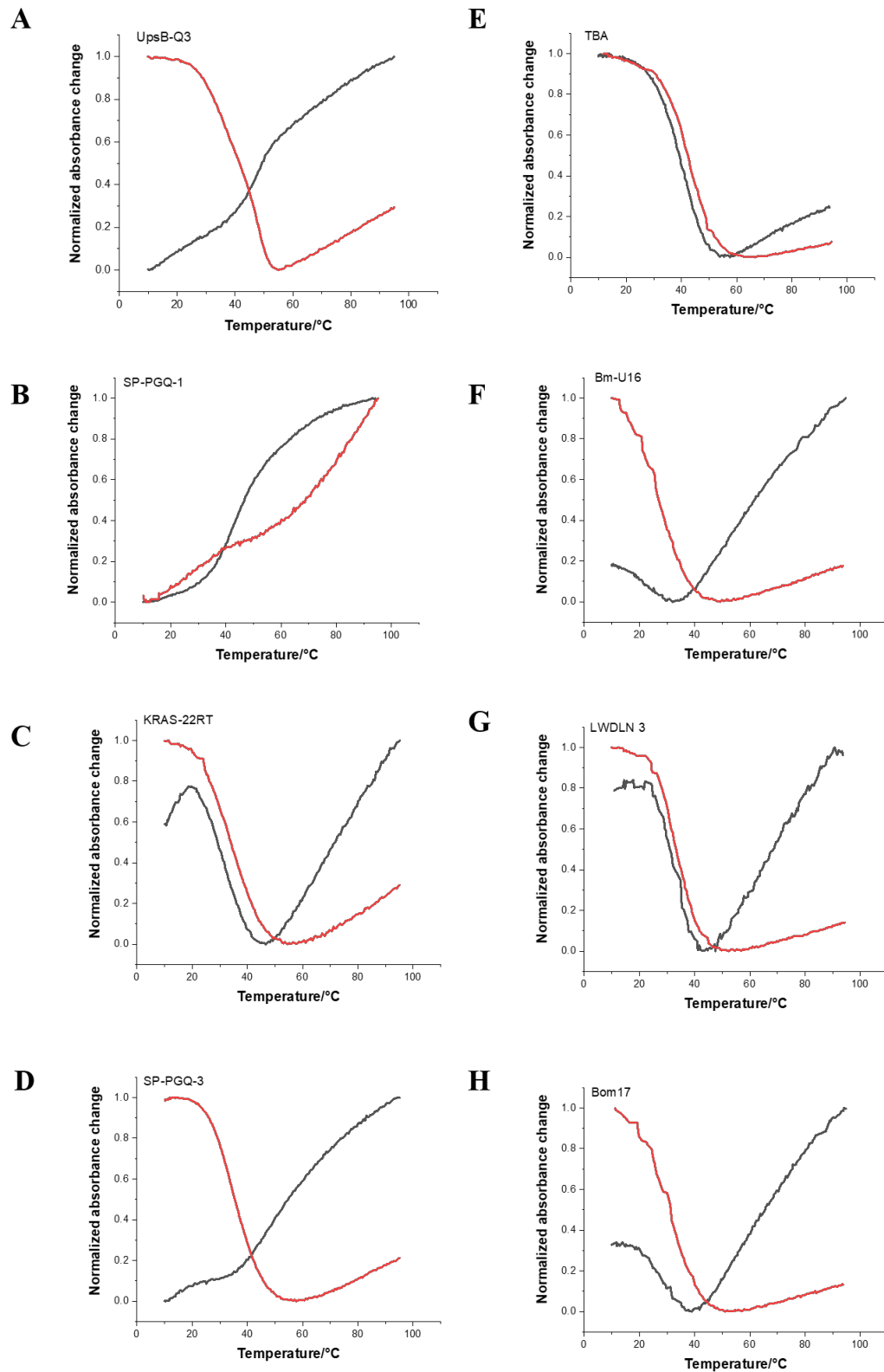


Fig. S4. Normalized UV-melting curves for some sequences of the training set. All data collected in 10-95 °C, monitored at 260 nm (Blank lines) and 295 nm (Red lines).

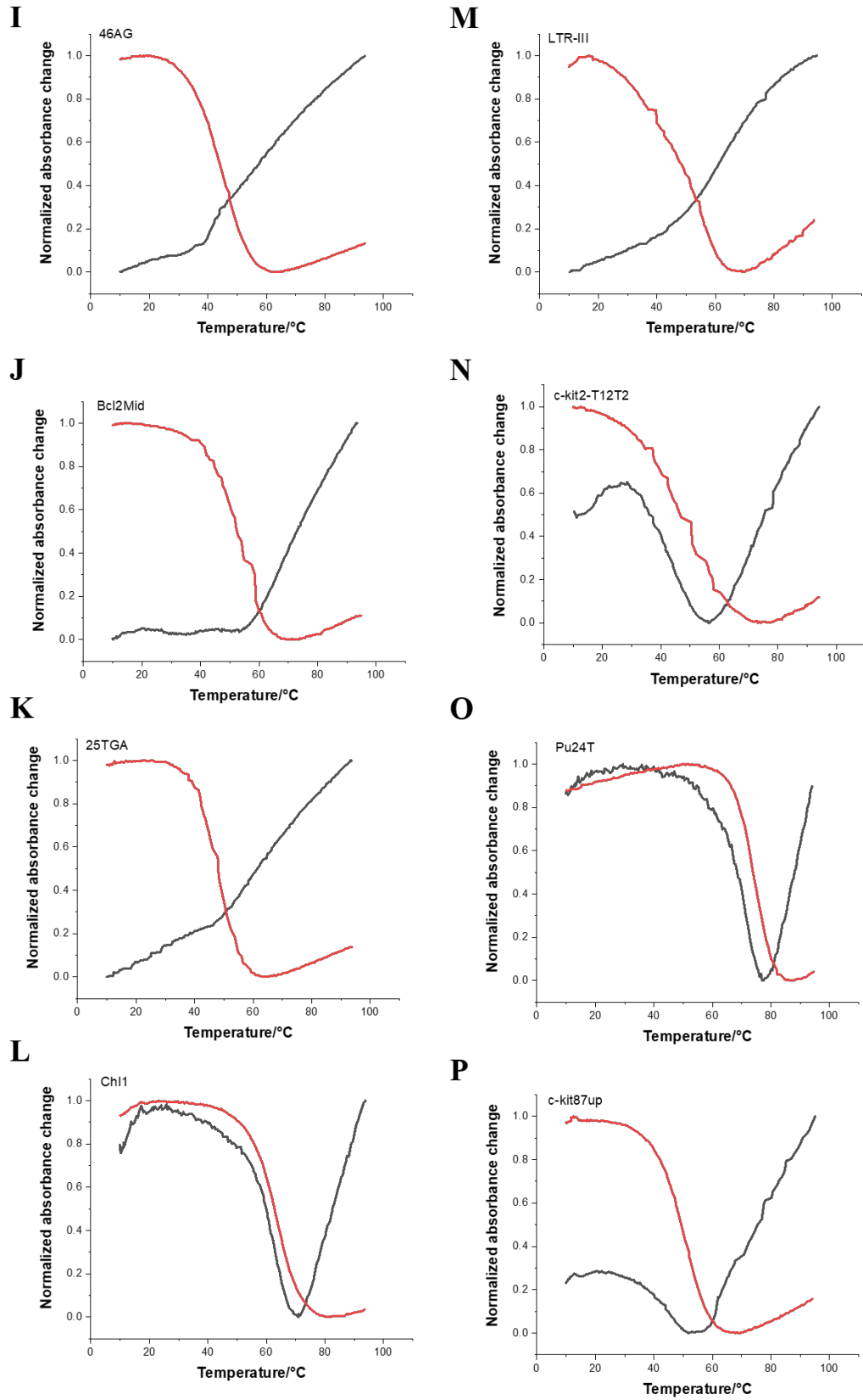


Fig. S4. *continued.*

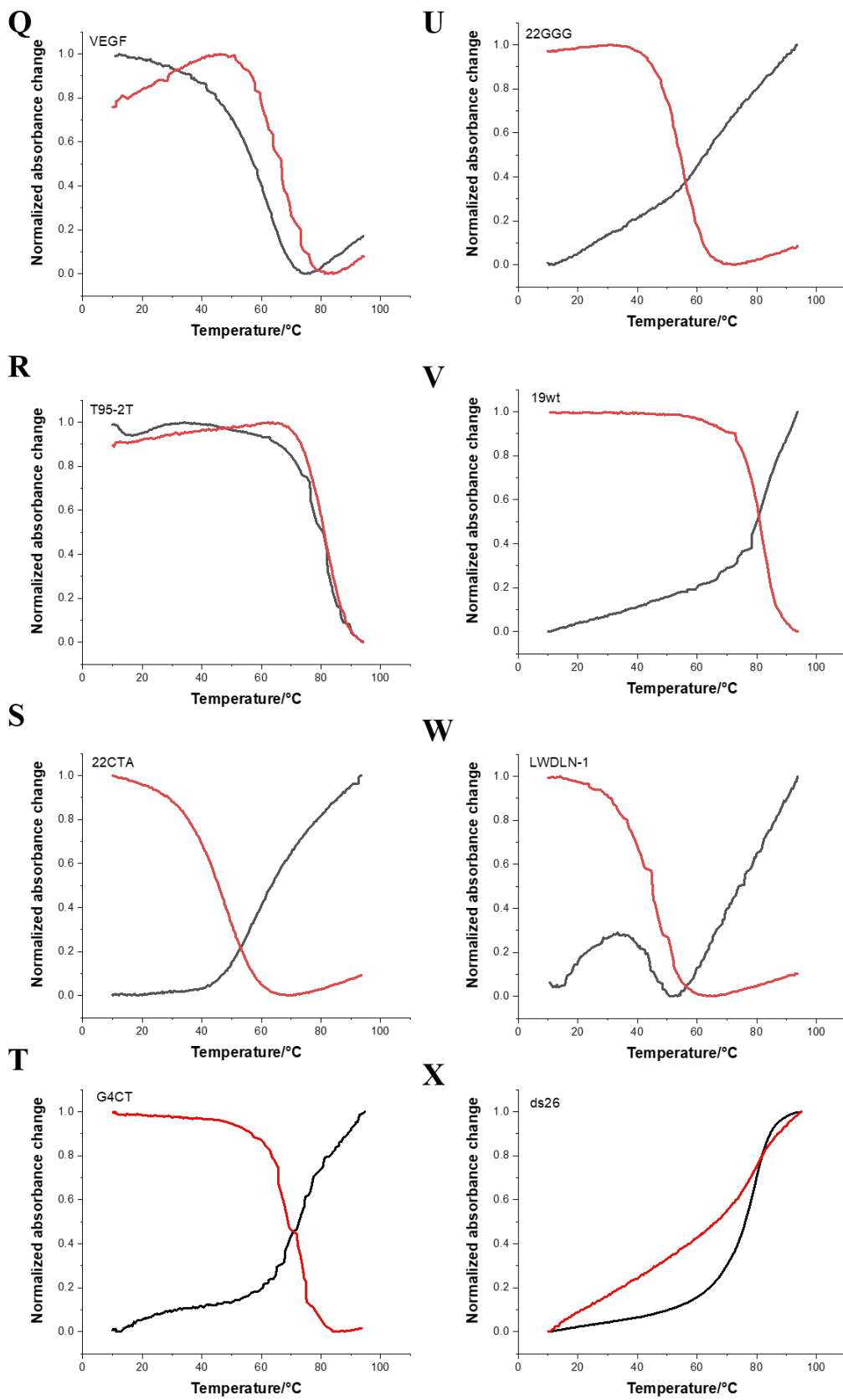


Fig. S4. *continued.*

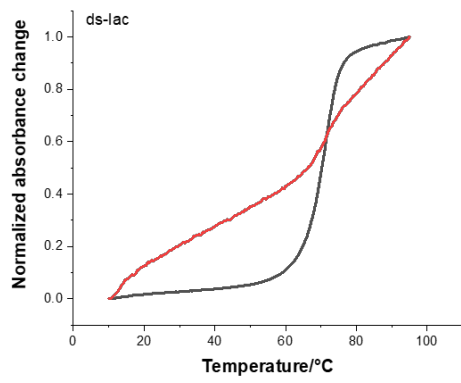
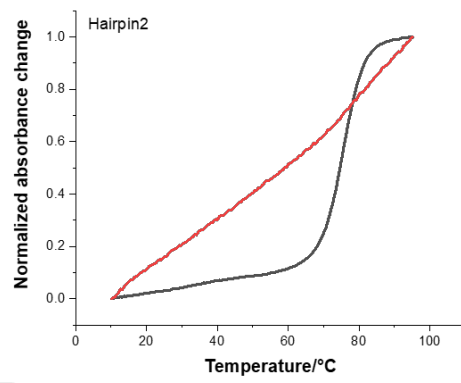
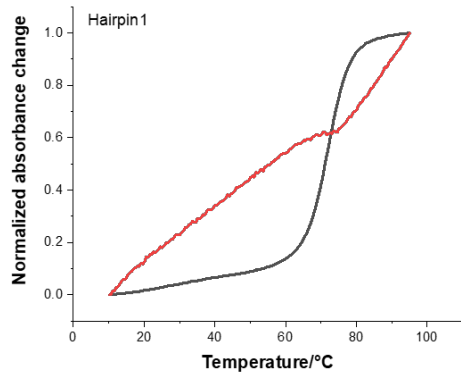
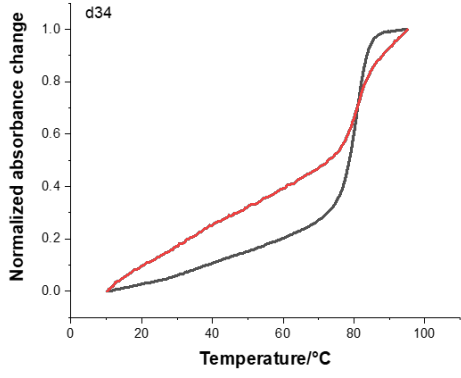
Y**AA****Z****AB**

Fig. S4. *continued.*

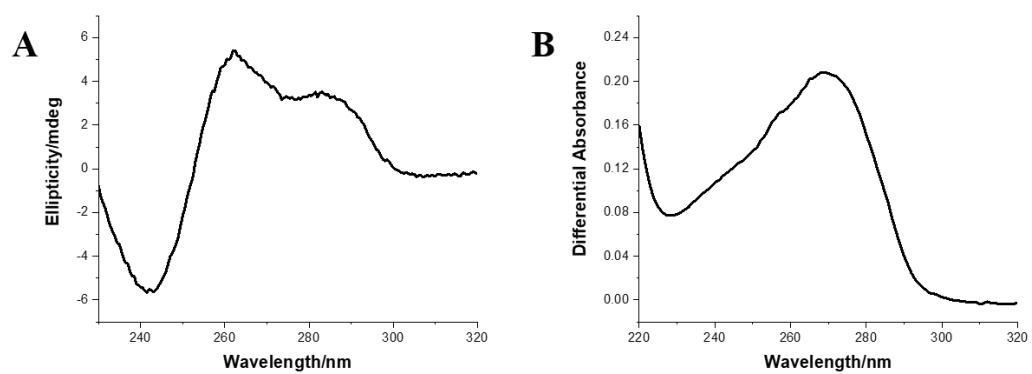


Fig. S5. CD (**A**) and thermal differential absorbance (TDS) (**B**) spectra for 3 μ M SP-PGQ-1 in the FRET buffer.

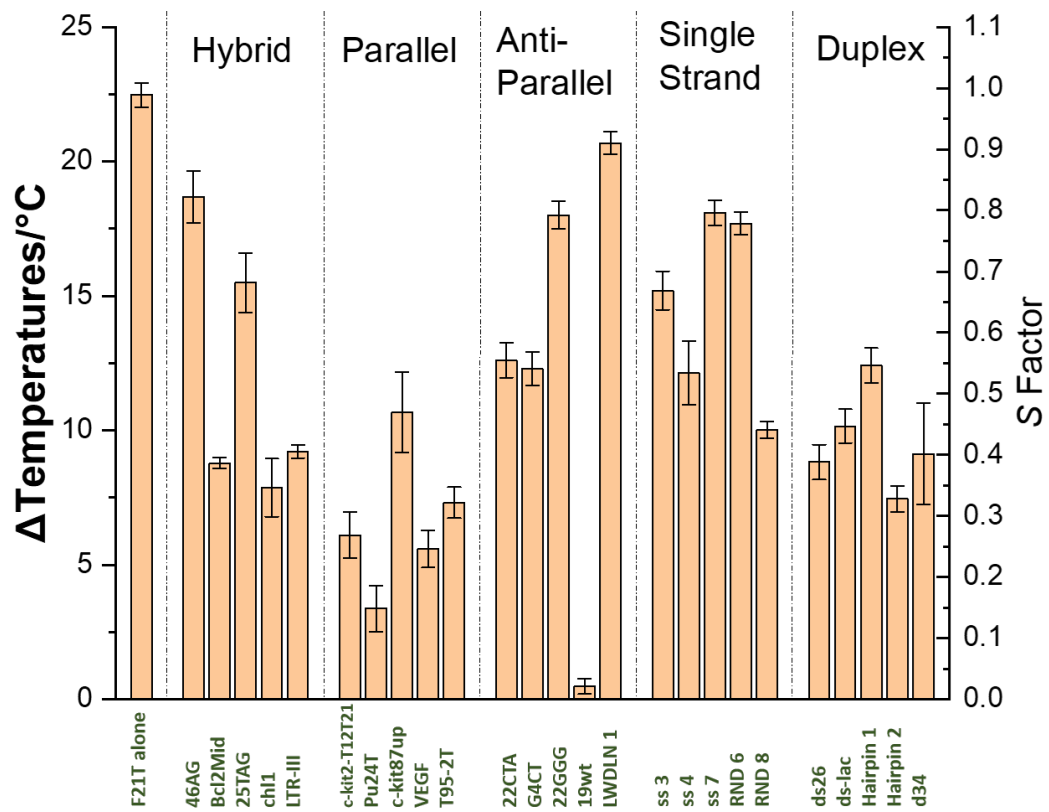


Fig. S6. ΔT_m induced by 0.4 μM TMPyP4 on 0.2 μM F21T, alone or in the presence of 3 μM competitors. The *S Factor* is also provided on the right Y-axis. Samples were annealed and measured in 10 mM KCl, 90 mM LiCl, 10 mM lithium cacodylate pH 7.2 buffer.

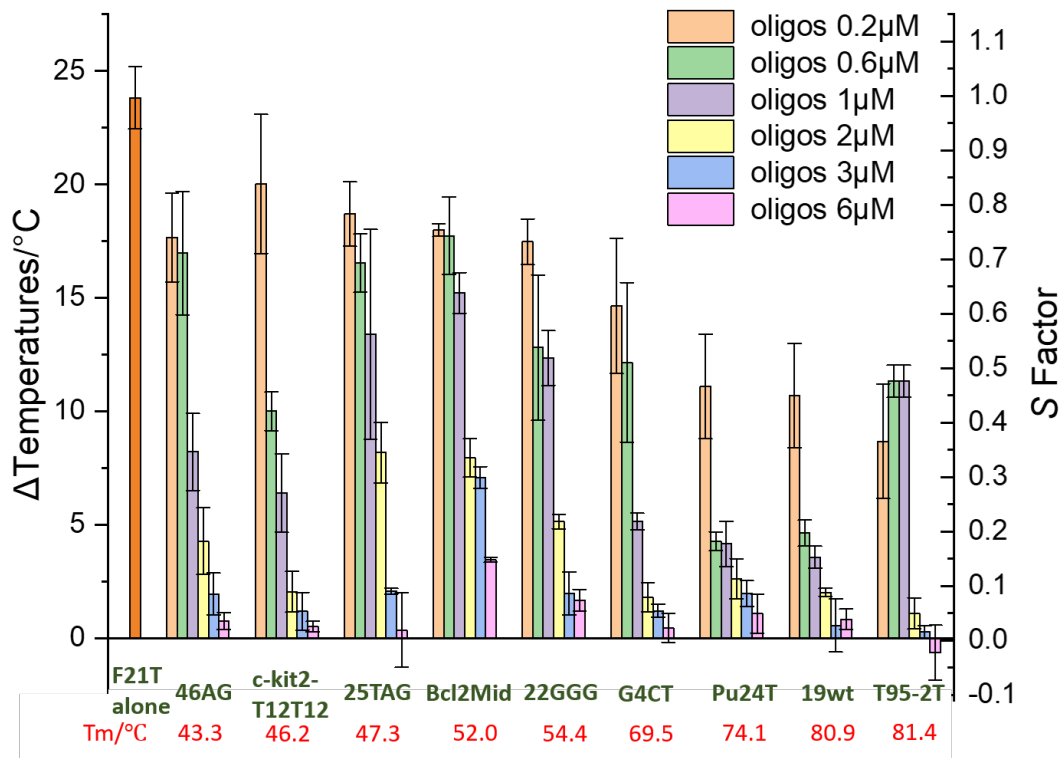
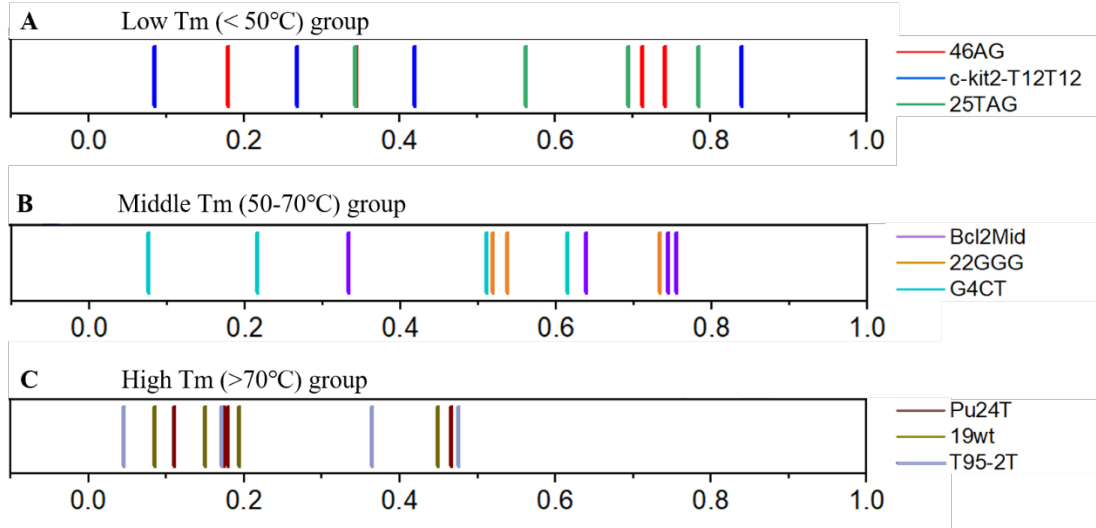


Fig. S7. ΔT_m induced by 0.4 μM PhenDC3 on 0.2 μM F21T, alone or in the presence of 0.2 / 0.6 / 1 / 2 / 3 / 6 μM competitors (1x to 30x molar excess, as compared to F21T). The *S Factor* is also provided on the right Y-axis. Samples were annealed and measured in 10 mM KCl, 90 mM LiCl, 10 mM lithium cacodylate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



Distribution of S Factor

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 T_m : low T_m ($< 50^\circ\text{C}$) (**A**), middle T_m ($50\text{-}70^\circ\text{C}$) (**B**) and high T_m ($>70^\circ\text{C}$) (**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

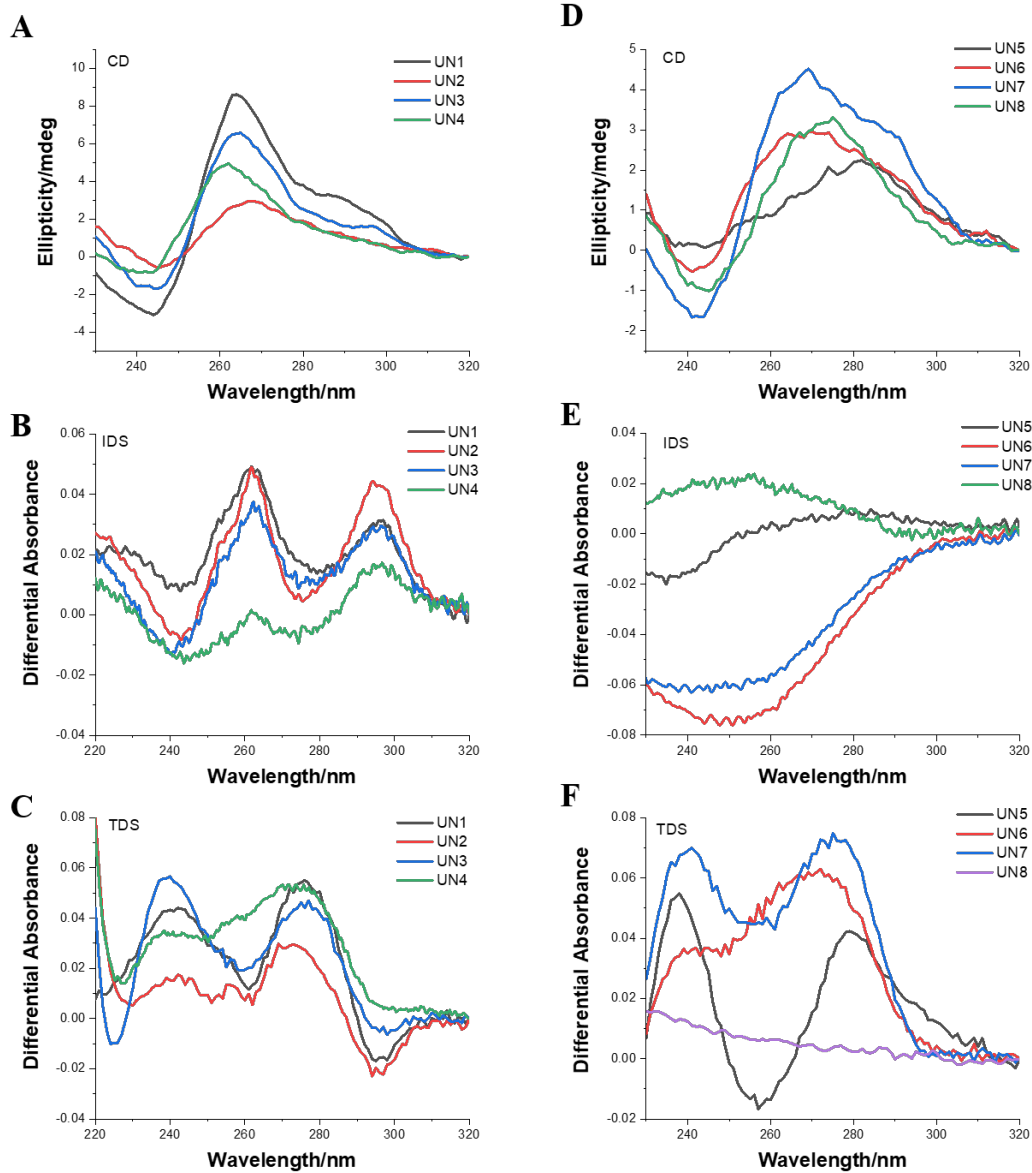


Fig. S9. CD, IDS, and TDS spectra of testing set. Panels **A, B, C** correspond to positive sequences (absolute S values < 0.1) and panels **D, E, F** to negative sequences, unable to compete ($S > 0.8$). From top to bottom: CD spectra, IDS spectra, and TDS spectra.