Supplementary Data and Materials

Resolution and characterization of the structural polymorphism of a single quadruplex-forming sequence.

Magdalena M. Dailey¹, M. Clarke Miller², Paula J. Bates², Andrew N. Lane²*, John O. Trent²*.

1 Department of Chemistry,
2 Department of Medicine,
Clinical Translational Research Building,
University of Louisville,
505 Hancock St.
Louisville, KY 40202, USA

Contact Information*
John O. Trent, PhD
Phone 502-852-2194
Fax 502-852-7979
Email: john.trent@louisville.edu

Andrew N. Lane
Phone 502-852-3067
Fax 502-852-7979
Email: andrew.lane@louisville.edu

Department of Medicine,
Clinical Translational Research Building,
University of Louisville,
505 Hancock St.
Louisville, KY 40202, USA
S1. Polyacrylamide electrophoresis gel of AS1411, AS1411-3’-A and AS1411 fractions 1-5. The AS1411 and the AS1411-3’-A (3’A) sequence migrates as a single band under these conditions, with an apparent mobility of ca 22 bp duplex DNA (M=13 kDa). Fractions 1-5 show a very similar mobility to that of the parent mixture. Fraction 1 is a mixture of monomer and dimer species, and appears as a poorly resolved doublet by gel electrophoresis, with a slower band migrating at ca. 25 bp duplex DNA (M=15 kDa).
S2. Sedimentation Velocity Curves
Sedimentation velocity profiles of annealed, unfracionated AS1411 recorded at three different loading concentrations. The sedimentation velocity profile, c(s) of the mixture shows two species with sedimentation coefficients of 2.36 and 3.56 S at 293 K. The molecular weights were estimated as 8900±500 Da and 17600±1400 Da. The former (major) species has a mass close to the formula weight of a single strand (Mr=8272), and thus is a monomer. The larger mass corresponds to a dimer. We conclude therefore that the oligonucleotide forms a mixture of stable monomeric and dimeric G-quadruplexes.
Figure S3. Properties of the Na\textsuperscript{+} form of AS1411.

In the presence of Na\textsuperscript{+} a different quadruplex forms, as shown by the CD (A) and NMR spectra (B). The sodium form appears to be mainly a single, monomeric species, that has a considerably lower thermodynamic stability (T\textsubscript{m}= 314 K at 100 mM Na\textsuperscript{+}) than the potassium forms. The spectrum was recorded at 18.8 T, 20 °C using Watergate with an acquisition time of 1.5 s and a recycle time of 4 s.

A

![CD spectrum of AS1411 Na\textsuperscript{+}](chart1.png)

B.

![NMR spectrum of AS1411 Na\textsuperscript{+}](chart2.png)
Table S1. Fractional intensities from $^{13}\text{C}-^{1}\text{H}$ HSQC of AS1411

G8 was uniformly $^{13}\text{C},^{15}\text{N}$ labeled. Peak volumes in the $^{1}\text{H}({^{13}\text{C}})$HSQC were measured for several groups of peaks corresponding to C8-H8, C1'-H1' and C3'H1'. The volumes were summed to normalize individual intensities. The largest 6 resonances are tabulated in rank order.

| C8-H8 | C1'-H1' | C3'-H3' |
|-------|--------|--------|
| 0.21  | 0.37   | 0.15   |
| 0.18  | 0.15   | 0.14   |
| 0.18  | 0.17   | 0.11   |
| 0.09  | 0.08   | 0.11   |
| 0.09  | 0.07   | 0.09   |
| 0.07  | 0.07   | 0.08   |

$\Delta \Delta G = -RT \ln [n_i/n_j]$

Range ca. 0-1.7 RT
Simulations of kinetics and thermodynamics of folding and melting with multiple states.

(i) Exchange rates between states

Consider the mechanism:

\[
\begin{align*}
&k_1 & & k_2 \\
&N_1 \Leftrightarrow U \Leftrightarrow N_2 \\
&k_1' & & k_2'
\end{align*}
\] (S1A)

N1, N2 are stable folded species, and U is the unfolded form. At low T, the amount of U is insignificant. Assume that at equilibrium, N1 and N2 are equally populated.

Suppose N1 is purified, and the rate of exchange back to N2 through U is slow (the conditions for separation), and direct exchange is insignificant. How fast does N2 re-equilibrate? The formation of N2 is given by the biexponential form:

\[
n_2(t) = \langle n_2 \rangle + \langle n_2 \rangle/\left(\lambda_1 - \lambda_2\right)[\lambda_2 \exp \lambda_1 t - \lambda_1 \exp \lambda_2 t]
\] (S1B)

and

\[
2 \lambda_+ = -(k_1+k_2+k_1'+k_2') \pm \left[(k_1+k_2+k_1'+k_2')^2-4[k_1'k_2+k_2'(k_1+k_1')]\right]^{0.5}
\] (S1C)

In this case, \(\langle n_2 \rangle=0.5\).

For AS1411 at 303 K, the dissociation rate constants (\(k_1, k_2\)) are in the range 0.01 to 0.25 h\(^{-1}\), so the refolding rate constants must be much larger at this temperature (say 1000 to 10\(^6\) fold).

Simulations with \(k_1=k_2\) and \(k_1'=k_2\) in h\(^{-1}\) gave half times follows:

| number | \(k_1=k_2\) | \(k_1'=k_2\) | \(t_{1/2}\) app (h) | \(t_{1/2}\) \(\Rightarrow U\) $^\$ |
|--------|--------------|--------------|-----------------|-----------------|
| 1      | 0.0036       | 3.6          | 99              | 0.19            |
| 2      | 0.025        | 25           | >200            | 0.027           |
| 3      | 0.25         | 25           | 27.8            | 0.027           |
| 4      | 0.01         | 10           | 69.5            | 0.069           |
| 5      | 0.01         | 100          | 69.5            | 0.0069          |

* time for \(n_2\) to reach half of its equilibrium value

$^\$ half life to unfold to U in the absence of N2

For a single species equilibrating with U, the half-life is just \(0.693/(k_1+k_1')\) etc.

In effect, once N1 unfolds to U, it is rapidly removed by refolding into both N1 and N2, which maintains the reverse reaction rate slow; in essence, the rate is determined by the unidirectional rate constant. If there are several competing species for U, then the rate
becomes determined entirely by the unidirectional unfolding rate constant from N1, and largely independent of the folding rate constants (Fig. S4A). This is why the isolated species are in fact apparently kinetically very stable: the unfolding rate constant has to be slow on the time scale of days (or rate constants < $10^{-5}$ s$^{-1}$). When $k_{\text{fold}}>k_{\text{unfold}}$, $\lambda_1$ and $\lambda_2$ approach $k_-$ and $k_+$, so the formation kinetics reduce to:

$$n_2(t) = <n_2>[1-\exp(-k_\cdot t)]$$

**Figure S4A.** Time course of N1:N2 exchange according to model S1A. The fraction of the total that is N2 was calculated as a function of time using the rate parameters given in the table.

| T/K  | k h$^{-1}$ | Half life/min |
|------|------------|---------------|
| 293  | .000019    | 2.2E6         |
| 303  | 0.01       | 4158          |
| 313  | 0.4        | 104           |
| 323  | 12.8       | 3.25          |
| 333  | 331        | 0.13          |
| 343  | 7084       | .00059        |

Suppose k$-$ = 0.01 h$^{-1}$ at 303 K, then for $E_a$=70 kcal/mole, k$1$ varies with T as follows:
(ii) Simulation for \( N_1 \Leftrightarrow U \Leftrightarrow 3N_2 \) \hspace{1cm} (S2)

Here it is assumed that at equilibrium, the N2 form is three times as much as the N1 form.

At thermal equilibrium, \( n_2 = 3n_1 \) (low T)

Let \( k^- = 1 \times 10^{-6} \text{ s}^{-1} \) at 303 K, \( E_a(\cdot) = 70 \text{ kcal/mol} \)

Then in the limit that the reaction can be described as single steps, \( k^+ = 0.01 \text{ s}^{-1} \) at 303 K, \( T_m = 340 \text{ K} \). This further implies \( E_a(\cdot) = 18.711 \text{ kcal/mol} \)

Thus at low temperature, the equilibration rates between N1 and other forms is slow, but at high temperatures the equilibration becomes fast, indicating an asymmetric unfolding process. The aim is to simulate the behavior of such as system under typical heating rates, in this case 0.5 °C min

Approximate calculations of the concentrations of N2, N1, and U starting from 100% N1, and the equilibrium profile allowing the rate constants to vary with temperatures according to the Arrhenius relationships given above.

**Figure S4B. Unfolding of an isolated species under non-equilibrium conditions.**

The appropriate rate equations for model S2 were integrated over time and mapped to the temperature-time profile. \(<u>, <n_1> \text{ and } <n_2>\) are the distributions of U, N1 and N2 under equilibrium conditions. \( u, n_1 \text{ and } n_2(t) \) are the kinetically determined populations of U, N1 and N2.

As shown in **Figure S4B**, at high T, the unfolding tracks the equilibrium distribution. The \( T_m \) for \( u \) is not 340 K because \( u = 50\% \) when \( K = 0.25 \) for this model. At low T, under non-equilibrium conditions and well below the global melting transition, the N1 converts to N2 once the rate constant rise to significant values for this time scale.
**Figure S5.** 3D CD melts of fractions showing the absence of isodichroic points. As the temperature is raised, a given fraction begins to unfold, and that unfolded portion will re-equilibrate with all of the other folded forms, therefore it cannot be considered a simple two state melting.
Simulations of kinetics and thermodynamics of folding and melting with multiple states.

(iii) Folding-unfolding at equilibrium: monomers only.

\[ N_1 \leftrightarrow U \leftrightarrow N_2 \]
\[ K_1 \quad K_2 \]

Assume reversible, equilibrium unfolding.

\[ K_1 = n_1/u_1; \quad K_2 = n_2/u_2 \]

\[ u_t = u + n_1 + n_2 = u(1+K_1+K_2) \]

\[ \frac{u}{u_t} = \frac{1}{1+K_1+K_2} \]

\[ \frac{n_1}{u_t} = \frac{K_1}{1+K_1+K_2} \]

\[ \frac{n_2}{u_t} = \frac{K_2}{1+K_1+K_2} \]

\[ S = \sigma_1 u + \sigma_{N_1} n_1 + \sigma_{N_2} n_2 \quad (S3) \]

\( \sigma \) is the specific signal strength and \( S \) is the observed signal.

Or \( S = \sigma_1 u_t / (1+K_1+K_2) + \sigma_{N_1} u_t K_1 / (1+K_1+K_2) + \sigma_{N_2} u_t K_2 / (1+K_1+K_2) \)

**Figure S6A.** Populations of States coupled through the unfolded form U11 is the unfolded state, N11 and N21 are folded forms. (i) populations (ii) CD signal calculated for different values of the CD intensity of N1, N2 and U, A: \( \sigma_1 = -3, \sigma_{N1}=20, \sigma_{N2}=10 \). B \( \sigma_1=-3, \sigma_{N1}=2 \sigma_{N2}=20 \)
A) Simulations were carried out using:
with \( T_{m1} = 343 \) K, \( T_{m2} = 333 \) K and \( \Delta H1=\Delta H2 = 70 \) kcal/mol.

\[
K = \exp(\Delta H/R)(1/T_{m1}/T)
\]

(S4)

Note that in this model, the ratio \( n_1/n_2 \) is independent of temperature. If \( \Delta H1 \neq \Delta H2 \), the ratio \( n_1/n_2 \) varies according to the difference in the enthalpies.

B) Vary \( T_m \) and \( \Delta H \) for \( N_1 \Leftrightarrow U \Leftrightarrow N_2 \)

The CD for an equilibrium unfolding profile was calculated as described in Fig. S5B.

The data were then fitted to 2-state transitions with linear baselines (cf Eq. 1, main text). Recovered thermodynamic parameters are given in Table S2 below.

**Figure S6B.** Populations of states coupled through the unfolded form: effect of different \( T_m \) and \( \Delta H \) values.

Vary \( T_m \) and \( \Delta H \) for \( N_1 \Leftrightarrow U \Leftrightarrow N_2 \)

1. \( T_{m1}=T_{m2}=340 \) K; \( \Delta H1=60, \Delta H2=70 \) kcal/mol
2. \( T_{m1}=330 \) K, \( T_{m2}=340 \) K, \( \Delta H1=\Delta H2=70 \) kcal/mol
3. \( T_{m1}=330 \) K, \( T_{m2}=340 \) K; \( \Delta H1=70, \Delta H2=60 \) kcal/mol.
CD with \( \sigma_0=-2, \sigma_1=10, \sigma_2=15 \)
| Model | Tm app/K | ΔH<sub>app</sub> kcal/mol | R²    |
|-------|----------|---------------------------|-------|
| 1     | 342.3    | 60.4                      | 0.9999|
| 2     | 340.1    | 70.0                      | 1.0   |
| 3     | 339.9    | 55.4                      | 0.9997|

**Table S2.** Fitted two-state values
(iv) Monomer-dimer equilibrium
Fraction 1 of AS1411 is a mixture of dimer (N2) and monomer (N1) quadruplex species. The relative amounts of these must depend on the details of the relative thermodynamic stability and the concentration of oligonucleotide.

Consider the model:
\[
K_2 \quad K_1
\]
\[
N2 \leftrightarrow 2U \leftrightarrow N1
\]  \hspace{1cm} (S5)

We have simulated the distribution of N2, N1 and U as a function of temperature for different total DNA concentrations (from 1 to 100 µM), with the following parameter values.

\[
T_m(N1)+ = 340 \text{ K}, \Delta H = 60 \text{ kcal/mol}. K_2 = 1 \text{ at } 350 \text{ K}, \Delta H = 100 \text{ kcal/mol}.
\]

\[
n_t = n_1 + 2n_2 + u
\]

At low T, \(n_t = n_1 + 2n_2\) and \(n_2 = K_2n_1^2/K_1^2\)

In addition to the expected concentration dependence of the melting profile, the populations of the relative populations of the monomer and dimer species are also temperature dependent blow the global melting transition (Figure S7).

![Figure S7. Unfolding of a monomer-dimer mixture.](image)

U, n1, n2 are the unfolded and folded monomer and dimer forms, respectively. Thermal unfolding was calculated for 1, 10 and 100 µM total strand concentration.