

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

Magdalena M. Dailey<sup>1</sup>, M. Clarke Miller<sup>2</sup>, Paula J. Bates<sup>2</sup>, Andrew N. Lane<sup>2,\*</sup> and John O. Trent<sup>2,\*</sup>

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

Received December 11, 2009; Revised and Accepted March 1, 2010

## ABSTRACT

The remarkable structural polymorphism of quadruplex-forming sequences has been a considerable impediment in the elucidation of quadruplex folds. Sequence modifications have commonly been used to perturb and purportedly select a particular form out of the ensemble of folds for nuclear magnetic resonance (NMR) or X-ray crystallographic analysis. Here we report a simple chromatographic technique that separates the individual folds without need for sequence modification. The sequence d(GGTGGTGGTGGTTGTGGTGGTGGTGG) forms a compact quadruplex according to a variety of common biophysical techniques. However, NMR and chromatography showed that this oligonucleotide produces at least eight monomeric quadruplex species that interconvert very slowly at room temperature. We have used a combination of spectroscopic, hydrodynamic and thermodynamic techniques to evaluate the physicochemical properties of the mixture and the individual species. These species have almost identical thermodynamic, hydrodynamic and electrophoretic properties, but significantly different NMR and circular dichroism (CD) spectra, as well as kinetic stability. These results demonstrate that simple standard low-resolution techniques cannot always be used for quadruplex fold determination or quality control purposes, and that simple thermodynamic analysis does not directly provide interpretable thermodynamic parameters.

## INTRODUCTION

In the presence of appropriate cations, nucleic acid sequences of the general kind (GGGN)<sub>n</sub> spontaneously fold into stable three-dimensional structures. These are characterized by the stacking of two or more successive planes of four guanine residues arranged in a square planar array via Hoogsteen hydrogen bonding (a G-quartet). The structures are stabilized by monovalent ions such as K<sup>+</sup> or Na<sup>+</sup> by coordination to the carbonyl O6 atoms (1). Such sequences are found naturally in the telomeres of chromatin in all species so far examined (2–6), as well as in internal sites of DNA, especially in promoter regions (7–9). The prevalence of these latter sequences is very high (7). In general the stability of quadruplex structures is greater in the presence of potassium rather than in sodium solutions (10,11), and it has been suggested that G-rich sequences may occur *in vivo* under appropriate circumstances, either as regulatory elements (2,12,13), or for as yet unknown reasons.

In addition to the possible biological functional roles of G-quadruplexes (2,12,14), such G-rich sequences have bio/nanotechnological uses, particularly in the area of aptamers (15). There are several G-quadruplex aptamers that have therapeutic uses, and are presently undergoing phase I and phase II clinical trials (16,17). A notable example is AS1411, which is now in phase II clinical trials. Quadruplexes have also been shown to have antiviral activity and have been demonstrated to be effective against HIV-1 *in vitro* (18,19). Because of their ability to recognize both nucleic acids and proteins with a high degree of specificity combined with their stability and nuclease resistance, quadruplexes are rapidly becoming attractive agents for development of novel therapeutics

\*To whom correspondence should be addressed. Tel: +1 502 852 2194; Fax: +1 502 852 7979; Email: john.trent@louisville.edu  
Correspondence may also be addressed to Andrew N. Lane. Tel: +1 502 852 3067; Fax: +1 502 852 7979; Email: andrew.lane@louisville.edu

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

and preclinical studies are underway for several other quadruplex-based therapeutics (20–22).

It has been shown that G-quadruplex-forming sequences are remarkably polymorphic. A single sequence may fold into several structures depending on the counterion residing in its core, the details of the folding conditions, and the environment. For example, the human telomeric sequence 5'dN(GGGTTA)<sub>3</sub>GGGN has been shown to fold into at least four distinguishable structures depending on whether the counterion is Na<sup>+</sup> (23), K<sup>+</sup> (24–26) or in the crystal state (27). Moreover, in K<sup>+</sup>-containing solution, this sequence simultaneously folds into at least two different quadruplex topologies.

Of the 26 possible looping topologies and 8 possible tetrad arrangements available based on the glycosyl bond angles orientations, there are over 200 possible unimolecular conformations for the simplest quadruplex (28). The formation of multistranded structures further increases the possible conformational space. Furthermore, the conformational polymorphism observed under solution conditions does not adequately reflect the conditions *in vivo*, including crowding, specific protein binding and the fact that such sequences are closed at one (telomeres) or both ends (internal DNA sites such as promoter regions) (29,30).

Polymorphism seems to be a common feature of G-rich sequences (1,30). Whether this occurs in biological systems or only *in vitro*, it is important to understand the rules that determine the populations of the various structures under given conditions, especially for biotechnology applications. Many quadruplex systems cannot be described structurally due to high complexity resulting from the simultaneous presence of multiple species. Ideally one would simply separate the various species present in solution in order to analyze them individually. Unfortunately, for G-quadruplexes it is difficult to apply standard separation techniques because, as we show here, different folds of the same sequence often have nearly identical physical properties. As a result, instead of separation, researchers have 'isolated' single structural configurations by modifying the sequence of interest (24–27,31,32), which may artificially force the strand into a distinct conformation not available to the wild-type sequence. This may contribute to the inconsistencies in the reported structural and biophysical data for differently modified systems derived from a single sequence.

The anticancer aptamer AS1411 that has been extensively studied for its biological properties (16,33–35) has the sequence d(GGTGGTGGTGGTTGTGGTGGTGGTGG), which comprises (GGT)<sub>4</sub> and (TGG)<sub>4</sub> repeats separated by a TTGT linker. There are several theoretically possible monomeric, dimeric or tetrameric G-quadruplex topologies for this sequence as shown in Figure 1. This 26-nucleotide DNA sequence forms a mixture of several species when annealed slowly in K<sup>+</sup> buffer. In practice, it folds into at least eight different monomeric quadruplex structures under the same experimental conditions. Furthermore, a different species is produced by rapid cooling, indicating significant differences between kinetic and thermodynamic stability. We will show that many of these species can be resolved by

chromatography. We have characterized individual conformers by several biophysical techniques, and compared them both with one another and with the unresolved mixture. We have demonstrated that each conformation has very similar hydrodynamic and electrophoretic properties, and similar thermodynamic stability, yet can be resolved by chromatography into species that have very distinct nuclear magnetic resonance (NMR). With one exception, the circular dichroism (CD) spectra could not discriminate among the different species. In the absence of sufficient resolution or incomplete theory for most commonly used biophysical techniques (36), and because the G-rich oligonucleotides commonly adopt multiple conformations, analysis of the quadruplex system should begin from rigorous characterization by high information content, high resolution techniques such as NMR, before detailed analyses of other biophysical properties are initiated, as we have previously suggested (30).

## MATERIALS AND METHODS

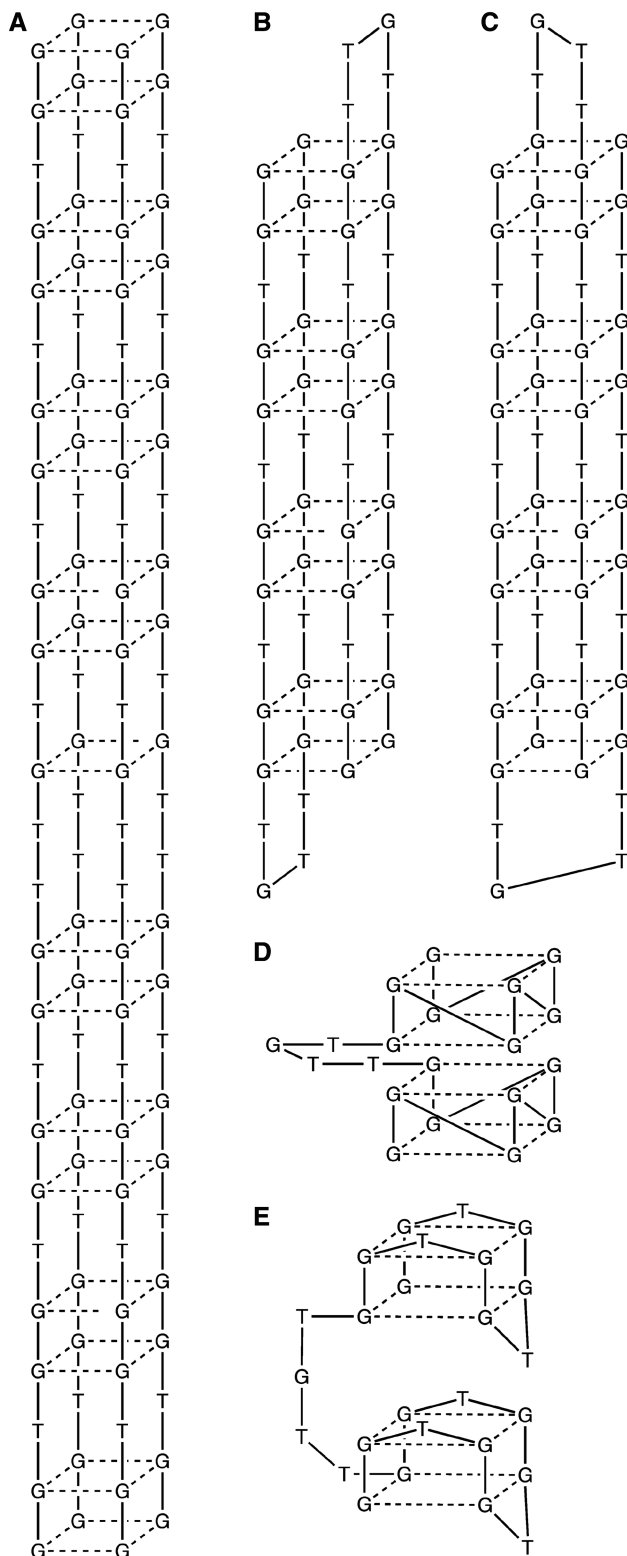
### Materials

The oligodeoxynucleotide AS1411-3'-A, d(GGTGGTGGTGGTTGTGGTGGTGGTGG), was purchased from Integrated DNA Technologies (Coralville, Iowa) and delivered as a desalted dry pellet. The oligomer was shown to be a single chemical species by electrospray mass spectrometry, high performance liquid chromatography (HPLC) and electrophoresis. This oligomer had enhanced ease of purification of fraction 2 that was used for the <sup>1</sup>H{<sup>13</sup>C} HSQC experiments. <sup>13</sup>C/<sup>15</sup>N guanine phosphoramidites were purchased from Cambridge Isotope Laboratories (MA), and incorporated into a series of 16 oligonucleotides by TriLink Biotechnologies (San Diego, CA), each containing a single <sup>13</sup>C/<sup>15</sup>N labeled guanine. The guanine in the linker region (G14) of the sequence was not included in the series. AS1411, d(GGTGGTGGTGGTTGTGGTGGTGGTGG) was also obtained, which had been extensively HPLC purified prior to annealing from Antisoma, UK. The complementary DNA oligomer d(CCACCACCACCACAACCACCACCACC) was purchased from Integrated DNA Technologies.

### Methods

**Preparation of samples.** The purified oligonucleotides were dissolved in a buffer containing 80 mM KCl, 20 mM potassium phosphate buffer (pH = 7.0) at a concentration of 10 mg/ml and allowed to dissolve over several hours. Samples were then equilibrated for 10 min at 100°C in a water bath, followed by spontaneous cooling to room temperature overnight (annealed samples). The DNA was also subjected to rapid cooling by heating a 500 μM sample in the same buffer to 100°C for 10 min, followed by quenching on ice ('quench'). The DNA samples were then refrigerated until use.

**Size exclusion chromatography.** A Superdex 75 10/300 column (GE Healthcare) was installed on an ÄKTA-FPLC with UPC-900 UV absorbance monitor and Frac920 fraction collector (GE Healthcare) and prepared



**Figure 1.** Some possible quadruplex topologies for d(GGTGG TGGTGGTTGTGGTGGTGGTGG). (A) 4-strand parallel quadruplex comprising 104 nucleotides (not observed experimentally). (B and C) Antiparallel dimer basket and chair quadruplexes comprising 52 nucleotides (dimers were observed in the annealed mixture). (D) Monomer comprising two stacks of parallel quartets linked by TTGT (26 nucleotides). The parallel loops contain a single thymine (not shown for clarity). There are four such arrangements denoted by 1234, 1243, 2143 and 3412 where 1, 2, 3, 4 denote the individual quartets reading

with a mobile phase consisting of 100 mM KCl, 25 mM  $K_2HPO_4$  (pH = 8.0). For calibration purposes the column was run at a flow rate of 0.10 ml/min. Elution was monitored at  $A_{280}$ . The elution of proteins was calibrated using standards of known molecular weight from MW-GF-70 and MW-GF-1000 molecular weight calibration kits (Sigma). The void volume was determined by injection of 200  $\mu$ l of 1 mg/ml blue dextran in elution buffer with 5% glycerol. For DNA separation, the flow rate was altered to 0.05 ml/min. DNA was injected at 10 mg/ml in up to 500  $\mu$ l aliquots. 0.2 mL fractions were collected.

*Analytical ultracentrifugation (AUC) and circular dichroism sample preparation.* Samples of purified oligonucleotides were diluted in the FPLC running buffer as described above, to the appropriate absorbance ( $\sim 0.8 A_{260}$ ).

*Electrophoresis.* Samples of AS1411, AS1411-3'-A and different fractions of AS1411 in 80% sucrose and Orange loading dye (Fermentas) were loaded in equal amounts on 5% acrylamide gels in TBA buffer, and run at a constant 100 V/cm for 1 h. The O'RangeRuler 5 bp DNA Ladder (Fermentas) and 20 bp PCR Low Ladder (Sigma) were used as molecular weight markers. Gels were stained with ethidium bromide (Sigma-Aldrich, St. Louis) and visualized with a UV transilluminator (BioRad ChemiDoc XRS).

*NMR sample preparation and data collection.* Initial NMR samples were freshly annealed DNA samples which had been dialyzed into fresh buffer and  $D_2O$  added to 10%. Sample concentrations varied from 100 to 1000  $\mu$ M. Samples of individual fractions were collected from the FPLC, combined, and concentrated using microcon YM-3 (Millipore) centrifugal concentrators. Each fraction was then re-injected into the FPLC, and relevant fractions were concentrated to 150–250  $\mu$ M. Thirty microliters of  $D_2O$  was added, and the samples brought to a total volume of  $\sim 330 \mu$ l with doubly deionized  $H_2O$ . Due to limited supply, individual labeled samples were run through the FPLC only once, concentrated, 30  $\mu$ l  $D_2O$  added, and then brought to a final volume of  $\sim 330 \mu$ l before loading into 5 mm Shigemi NMR tubes.

*CD spectroscopy.* CD spectra of samples having an  $A_{260}$  of 0.8 were recorded on a JASCO J-820 spectropolarimeter at 20°C in the FPLC running buffer from 340 to 220 nm at a scan rate of 200 nm/min using a 1 cm pathlength cuvette. The spectral band-width was set to 2 nm and data pitch to 0.2 nm. Spectra were recorded with a response time of 2 s. Three spectra were averaged, and the buffer baseline was subtracted. The raw CD intensities were normalized to the oligonucleotide concentration determined by absorption spectroscopy using an absorption coefficient of  $250\ 800\ M^{-1}\ cm^{-1}$ .

5' to 3' in the sequence. The rotation between the central quartets can be either positive or negative (at least seven distinguishable monomeric species were observed experimentally). (E) monomer comprising two stacks of antiparallel quartets linked by TTGT. There are four such arrangements denoted by 1234, 1243, 2143 and 3412 where 1, 2, 3, 4 denote the individual quartets reading 5' to 3' in the sequence.



**CD melting.** 3D CD melting curves were recorded in stoppered 1 cm cuvettes over the temperature range 20–90°C using a Peltier thermoelectric temperature controller and the oscillations around every temperature step  $< \pm 1^\circ\text{C}$ . The temperature was ramped at a heating or cooling rate of  $1^\circ\text{C}$  per min, and held for 1 min while a CD spectrum was recorded. The  $T_m$  was estimated from the first derivative of the melting profile at 263 nm. Curves were also fitted to a two-state transition with linearly sloping baselines as previously described (37), to give the apparent enthalpy  $\Delta H$  (assumed independent of temperature) and the melting temperature,  $T_m$  according to:

$$S = \sigma_u \cdot u + \sigma_N \cdot n \quad (1A)$$

$$\frac{\sigma_u \cdot u_l + \sigma_N \cdot u_l \exp(\Delta H/R((1/T) - (1/T_m)))}{1 + \exp(\Delta H/R((1/T) - (1/T_m)))} \quad (1B)$$

The baselines outside of the melting transition were assumed to be linearly dependent on temperature according to:

$$\sigma_u(T) = \sigma_u(293) + b(T - 293); \sigma_N(T) = \sigma_N(293) + a(T - 293) \quad (1C)$$

where  $a$  and  $b$  are constants,  $\sigma_u(293)$  and  $\sigma_N(293)$  are the specific ellipticities at the arbitrarily chosen reference temperature of 293 K.

These equations refer to single melting and unfolding transitions, which are the same if the system approaches equilibrium at each temperature. However, slow annealing from high temperature produced several distinguishable states at thermal equilibrium, which do not readily interconvert at temperatures below 30°C, but as shown from the reversibility of melting and cooling experiments, can do so via the unfolded state (see ‘Results’ section).

The apparent free energy of folding was estimated according to the relationship:

$$\Delta G(T) = T\Delta H \left( \frac{1}{T} - \frac{1}{T_m} \right) \quad (2)$$

Non-linear regression analyses were carried out using Kaleidagraph (Synergy Software).

**Analytical ultracentrifugation.** Sedimentation velocity profiles were recorded of unfractionated and fractionated AS1411 using a Beckman Optima XLA ultracentrifuge equipped with absorbance optical system and An-60 Ti rotor. All samples were dissolved in the FPLC buffer to working concentrations resulting in 0.25, 0.5 and 1 A<sub>260</sub> in 1 cm pathlength. Sedimentation velocity profiles were recorded at 20°C and 50 000 rpm. Primary data were processed using DCDT+ (38) and SedFit (39) and fit to appropriate physical models. Sedimentation coefficients showed little dependence on concentration under these conditions. Equilibrium centrifugation provided an independent estimate of  $M$ .

The translational friction coefficient  $f$ , was calculated from  $s$  as:

$$f = \frac{M(1 - v\rho)}{N_A s} \quad (3)$$

where  $\rho$  is the solution density,  $v$  is the partial specific volume and  $N_A$  is Avogadro’s number. A value of 0.54 g/ml was used for  $v$  (40,41) and the molecular weight was set to the chemical value.

**NMR Spectroscopy.** NMR spectra were recorded using a 5 mm inverse triple resonance (HCN) probe on Varian Inova spectrometers at either 18.8 or 14.1 T, the latter using a cold probe.

Spectra were recorded in both 90% H<sub>2</sub>O:10% D<sub>2</sub>O (NOESY) and in 100% D<sub>2</sub>O. Spectra of AS1411-3'-A in H<sub>2</sub>O were recorded over a range of temperature from 10 to 40°C. Spectra of the molecule in D<sub>2</sub>O were recorded at 30°C. <sup>1</sup>H{<sup>13</sup>C} HSQC spectra were recorded of annealed oligonucleotides containing a single uniformly labeled <sup>13</sup>C/<sup>15</sup>N nucleotide.

The kinetic stability of AS1411 was assessed by mixing the slowly annealed G-rich strand with the complementary C-rich strand at 150–200 μM, and monitoring the decay of imino proton resonances at 20°C, by proton NMR at 800 MHz. The imino proton resonances of the quadruplex and of the duplex are completely resolved from one another (see ‘Results’ section). Under these conditions, all of the quadruplex converts to duplex (30), and with a rate constant that largely represents the unfolding rate constant of the quadruplexes under the same conditions (42).

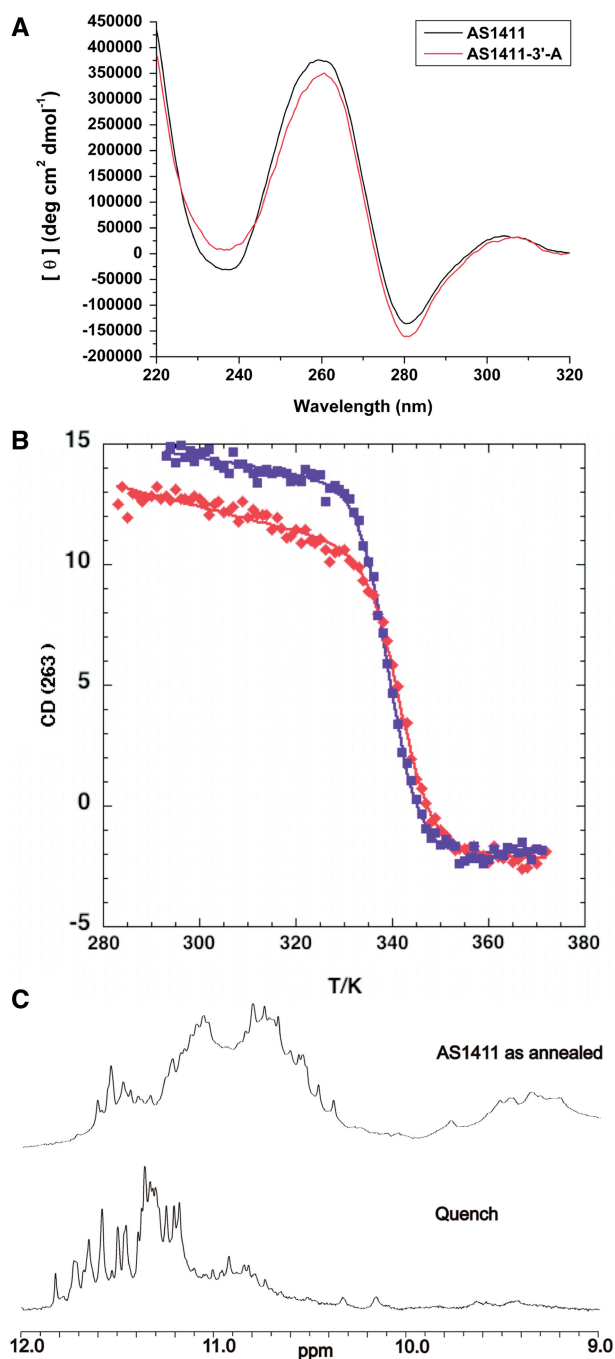
## RESULTS

### The sequence d(GGTGGTGGTGGTTGTGGTGGTGGTGG) forms a stable quadruplex

Figure 2A shows the CD spectrum of AS1411 annealed slowly in K<sup>+</sup> buffer, compared with the nearly identical AS1411-3'-A CD spectrum. Both samples also had the same relative electrophoretic mobility, and appeared as a single band, as shown in Supplementary Figure S1.

The apparent  $T_m$  values from CD melting curves under these conditions are ~338 K, which is in good agreement with expectations (43), indicating the formation of a thermodynamically stable species (cf. Table 1). The melting and annealing curves (Figure 2B) were analyzed as a two-state transition as described in the ‘Materials and Methods’ section, from which the apparent enthalpy change was obtained, and the free energy of folding under these buffer conditions using Equation (2). The averaged  $\Delta G(298)$  for the whole species was 36.8 kJ mol<sup>-1</sup>. The  $\Delta H$  (306 kJ mol<sup>-1</sup>) and  $\Delta G(298)$  values are ~30% higher than quadruplexes containing three G-quartets (30). As expected, slow cooling from the denatured state present at greater than 90°C was independent of the previous history of the samples, as shown by the good agreement among independently determined  $T_m$  and  $\Delta H$  values for the mixture and the fractions (and see below). Furthermore, the melting profile of the annealed oligonucleotide was very similar to that of the cooling profile, suggesting a reversible process at this heating and cooling rate.

The sedimentation velocity profile,  $c(S)$  (Supplementary Figure S2) of the mixture shows two species with



**Figure 2.** Biophysical properties of AS1411 and AS1411-3'-A. (A) CD spectra of AS1411 and AS1411-3'-A. (B) CD melting and annealing curves of AS1411 in  $K^+$  buffer. Annealing was carried out as described in the 'Materials and Methods' section at 263 nm. The CD data were fitted to a two-state transition using Equation (1) as described in the 'Materials and Methods' section. Melting (red)  $T_m = 341.8 \pm 0.24$  K,  $\Delta H = 66.9 \pm 3.2$  kcal/mol, cooling (blue),  $T_m = 339.4 \pm 0.14$  K,  $\Delta H = 72.9 \pm 2.5$  kcal/mol. (C) 1D  $^1H$  NMR spectra of AS1411 and quenched AS1411-3'-A.

sedimentation coefficients of 2.4 and 3.6 S at 293 K. The molecular weights were estimated as  $8900 \pm 500$  Da and  $17600 \pm 1400$  Da (Table 1). The former (major) species has a mass close to the formula weight of a single strand ( $M_r = 8272$ ), 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.

The frictional ratio of the monomers (Equation 3) is close to unity, and given that nucleic acids are typically well hydrated (40,44,45), this suggests that the monomeric species is highly compact, and behaves hydrodynamically as a nearly spherical object. In contrast, the frictional ratio of the dimer is considerably larger than unity, indicating a higher level of hydration (more than one layer of water) and/or a significant deviation from spherical geometry (Table 1).

### The sequence d(GGTGGTGGTGGTTGTGGTGGTGGTGG) forms a mixture of quadruplex species

Although the biophysical data are consistent with the presence of a monomeric quadruplex species and a certain amount of dimer, these low-resolution techniques do not unequivocally discriminate the several possible folding topologies (Figure 1), and thus experimentally generate averages. To reveal the real sample content requires high-resolution structural techniques, such as NMR. The 1D  $^1H$  NMR spectra of the slowly annealed oligonucleotide and the quenched samples are shown in Figure 2C. The frequency range of the imino proton resonances (10.5–12 ppm) is typical of G-quadruplex structures (26,46,47), in agreement with the biophysical data. A large fraction of these resonances persists for extended periods (>24 h) in pure  $D_2O$  solution, implying that the G-quartets are kinetically stable (47). However, there are many more distinguishable features in the imino proton region than the number of GN1H expected for a monomeric quadruplex ( $n = 17$ ), indicating that there are at least two to three quadruplexes present in the slowly annealed sample.

### Chromatography

Figure 3 shows a chromatogram of the slowly annealed AS1411. There are 10 peaks having different degrees of resolution, covering a range of elution volumes of 18 ml or  $\pm 40\%$  from the mean. The first three fractions account for two-third of the total. For comparison, a 10 kDa protein elutes at the same volume as fraction 2. The large retention volumes of species four and six implying either that their size is equivalent to a mass of <1000 Da, or that these molecules interact with the Superdex matrix. However, the CD and sedimentation velocity profiles show that not only are these species quadruplexes, but also that they are monomers having a Stokes radius and mass the same as that of the other fractions, and thus are not breakdown products (Table 1).

Both AS1411 and the AS1411-3'-A sequence migrates on a polyacrylamide gel as a single band under these conditions, with an apparent mobility of  $\sim 22$  bp duplex DNA ( $M = 13$  kDa) (Supplementary Figure S1). Remarkably, all of the fractions 1–5 show a very similar mobility to that of the parent mixture. As the fractions all have the same frictional ratio (Table 1) this implies that the effective charge distributions are also similar. Fraction 1 is a mixture of monomer and dimer species, and appears as a

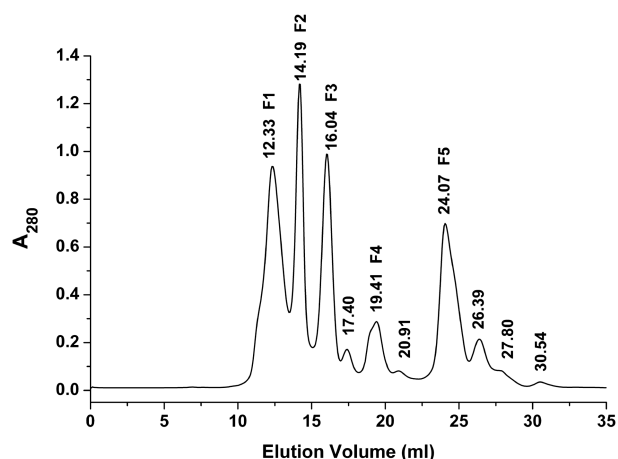
**Table 1.** Biophysical properties of the 26-mer AS1411

Fraction	Fractional area	$T_m$ K	$\Delta H_{\text{vH}}$ kJ mol <sup>-1</sup>	$\Delta G(298)$ kJ mol <sup>-1</sup>	$S_{20,w}$ S	$f/f_0^a$	$R_h$ Å	$Mr^b$
Slow anneal	–	339.7	317 ± 11	38.9	2.49 ± 0.22	1.1	13.1	9100
Slow melt	–	341.8	280 ± 14	35.9	3.56 ± 0.09	1.22	18.3	16940
1	0.29	337.9	291.9	34.5	2.31 ± 0.094	1.02	14.1	8940
2	0.2	338.2	319.3	37.9	3.78 ± 0.32	1.29	17.3	18200
3	0.18	338.5	301.4	36.1	2.35 ± 0.032	1.04	13.9	8740
4	0.05	337.3	310.8	36.2	2.27 ± 0.034	1.00	14.4	8883
5	0.19	338.4	308.9	36.9	2.26 ± 0.006	1.00	14.4	9170
			308.9	36.9	2.26 ± 0.012	1.00	14.4	8693

$T_m$  and  $\Delta H$  were determined from annealing curves.

<sup>a</sup> $f$  is the translational friction coefficient [Equation (3)],  $f_0$  is the expected value for an anhydrous sphere of the same molecular weight.

<sup>b</sup>The chemical formula weight is 8272.

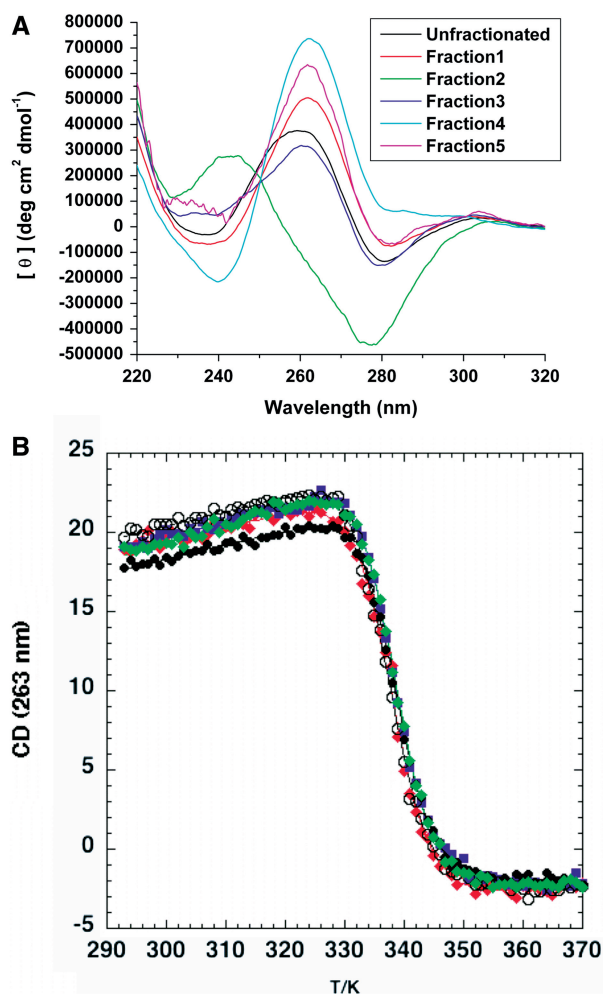


**Figure 3.** Size exclusion chromatography of AS1411. The column was run at 0.05 ml/min as described in the ‘Materials and Methods’ section. Numbers over each peak are the retention volume in milliliter with F1–5 labels representing fractions 1–5.

poorly resolved doublet by gel electrophoresis, with a slower band migrating at ca. 25 bp duplex DNA ( $M = 15$  kDa). Thus, electrophoresis does not necessarily resolve multiple isomeric species; a single band is not conclusive evidence of a pure component.

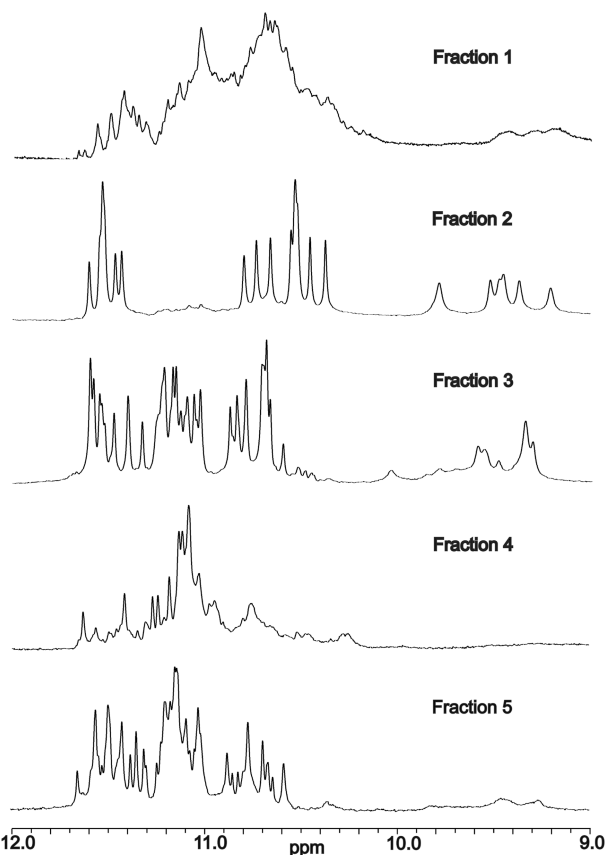
The CD spectra (Figure 4) of fractions 1, 3, 4 and 5 have similar shapes with extrema near 240 (–), 260 (+), 283 (–) and 305 (+) nm (and a third positive band below 220 nm), albeit with a 2-fold range in intensity. These spectra are nearly isodichroic at 250 nm. These bands correspond to Cotton effects in the absorption spectra, which have a maximum at 257 nm, and a shoulder at 293 nm (not shown). These are typical of G-quadruplex CD spectra (36,48). Fraction 2 has a highly anomalous CD spectrum, with extrema at 240 (+), 278 (–) and 308 (+) nm. As we will show, this species has a substantially different NMR spectrum from that of the other fractions. The annealing curves for fractions 1–5 are extremely similar as expected (Figure 4).

Fraction 1 shows a broad overlap of resonances in the NMR spectrum (Figure 5), many in common with those of the mixture. This species is known to be a mixture of a monomeric and dimeric quadruplex forms (Table 1 and



**Figure 4.** Electronic spectroscopy and melting. (A) CD spectra of AS1411 chromatographically resolved fractions. (B) CD annealing curves of chromatographically resolved fractions. The CD data were fitted to a two-state transition using Equation (1) as described in the ‘Materials and Methods’ section, at 263 nm. Fraction 1 (red)  $T_m = 337.9 \pm 0.14$  K,  $\Delta H = 71.2 \pm 2.4$  kcal/mol, fraction 2 (blue),  $T_m = 337.9 \pm 0.1$  K,  $\Delta H = 76.4 \pm 1.6$  kcal/mol, fraction 3 (black),  $T_m = 337.3 \pm 0.1$  K,  $\Delta H = 72.1 \pm 1.6$  kcal/mol, fraction 4 (white)  $T_m = 338.4 \pm 0.14$  K,  $\Delta H = 74.4 \pm 1.5$  kcal/mol and fraction 5 (green)  $T_m = 338.5 \pm 0.1$  K,  $\Delta H = 75.2 \pm 1.5$  kcal/mol.

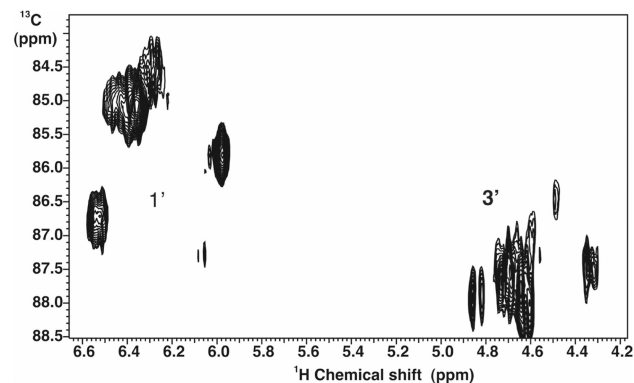




**Figure 5.** 1D  $^1\text{H}$  NMR spectra of AS1411 chromatographically resolved fractions 1–5. Spectra were recorded at 800 MHz and 20°C as described in the ‘Materials and Methods’ section. The free induction decays were apodized using a Gaussian function and a line broadening exponential of 2 Hz. The region shown encompasses the H-bonded GN1H (10.5–12 ppm) and GN2H proton resonances (9–10 ppm).

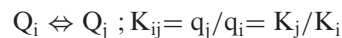
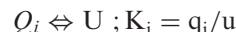
presented AUC data). Fractions 3 to 5 are all monomeric, and the imino proton resonances are sharp and relatively well resolved. Nevertheless, both fractions 3 and 5 have roughly 30 GN1H resonances, indicating that these species too are mixtures. The best resolved spectra are those of fractions 2 and 4, which show a number of GN1H resonances similar to that expected for a monomeric species containing 17 GN1H. Some of the GN1H in Fraction 4, however, are broad and ill-defined, and there are no GN2H resonances (9–10 ppm) visible, suggesting a less stable species. In contrast the NMR spectrum of fraction 2 is sharper, well resolved and contains 14–15 GN1H and 6 GN2H resonances, consistent with a single, monomeric and stable G-quadruplex having four G-quartets.

A further conformational state is accessible to this sequence in the presence of  $\text{Na}^+$  ions, analogous to other quadruplex-forming sequences (27,32,36). In the presence of  $\text{Na}^+$  a different quadruplex forms, as shown by the CD and NMR spectra (Supplementary Figure S3). The sodium form appears to be predominantly a single, monomeric species, which has a considerably lower thermodynamic stability ( $T_m = 314\text{ K}$  at 100 mM  $\text{Na}^+$ ) than the potassium forms. This behavior is analogous to the human telomeric sequences (27,32).



**Figure 6.**  $^1\text{H}$ - $^{13}\text{C}$  HSQC spectrum of  $^{13}\text{C}$ -labeled AS1411-3'-A. The quadruplex sequence containing uniformly  $^{13}\text{C}/^{15}\text{N}$  labeled G at position eight was annealed as described in the ‘Materials and Methods’ section. The 2D  $^1\text{H}\{^{13}\text{C}\}$  HSQC experiment was recorded in  $\text{D}_2\text{O}$  at 30°C at 14.1 T with GARP decoupling with acquisition times of 0.15 s in  $t_2$  and 0.011 s in  $t_1$ . The data tables were zero-filled once in  $t_2$ , linear predicted in  $t_1$ . The regions containing the 1' and 3' C–H interactions are shown, which comprise at least seven resonances (See Supplementary Table S1).

The chromatographic fractions clearly correspond to different G-quadruplex species all derived from a single oligonucleotide chain. Under slow annealing conditions the final populations should approach the equilibrium distribution, as indicated in Table 1. The free energy differences between these species are very small, i.e. they have very similar thermodynamic stability. Moreover, they are kinetically very stable, as they do not interconvert on the timescale of the chromatography experiment or for extended periods in the NMR (days). Independent analysis shows that only fractions 4 and 5 interconvert over a time scale of 1 week or more at 37°C (see Supplementary Data). At 90°C, the CD spectrum is very weak, indicating that at this temperature the strand is essentially a ‘random coil’ with little base-base stacking. This suggests that the folding process from an unfolded strand can be represented by the following model.



where  $Q_i$ ,  $Q_j$  are different folded quadruplex species and  $U$  is the unfolded ensemble and  $q$ ,  $u$  are the concentrations of these species.

The fractional population of any folded species is then

$$f_i = \frac{K_i}{(1 + \sum K_i)} = \frac{K_i}{\sum K_i} \text{ at low } T \quad (4)$$

This was further demonstrated by NMR analysis of AS1411-3'-A in which G8 was uniformly labeled with  $^{13}\text{C}$  and  $^{15}\text{N}$  (>98%). The  $^{15}\text{N}$ -edited 1D spectrum of the unfractionated samples shows multiple GN1H resonances, that represent different conformations in slow exchange on the chemical shift time scale ( $\Delta\delta = 0.05\text{ ppm}$  hence  $k \ll 30\text{ s}^{-1}$ ). In contrast, fraction 2 shows a single GN1H resonance (not shown). The  $^1\text{H}\{^{13}\text{C}\}$ -HSQC 2D spectrum (Figure 6) of the sample shows several cross peaks for C8-H, and for the sugar

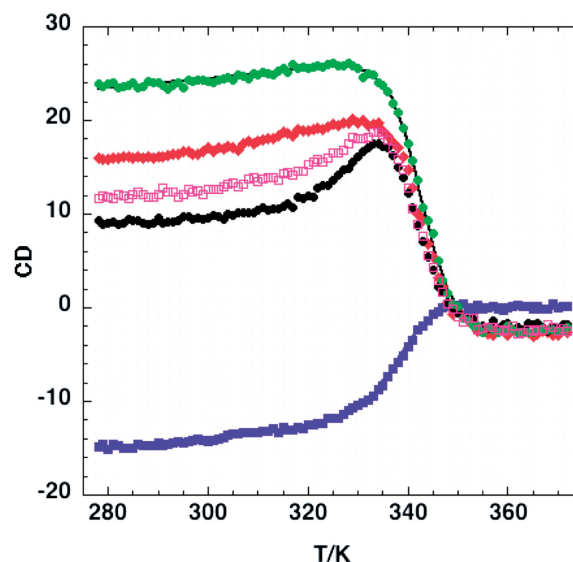
protons, again due to the presence of multiple conformations that are in slow exchange on the chemical shift time scale ( $k \ll 10 \text{ s}^{-1}$ ). As the relaxation properties and magnetization transfer function to be essentially the same for each conformation, we have volume integrated the peaks in each cluster, which were then used to estimate the relative populations of the species present.

Figure 6 shows at least eight distinguishable states for G8, which is comparable to the number resolved by chromatography. We have obtained similar results for other G resonances, implying that the conformational heterogeneity extends throughout the structure (e.g. different topologies). The eight most abundant species have mole fractions ranging from 0.03 to 0.36, equivalent to  $\Delta\Delta G \sim 2.5 \text{ RT}$  (Supplementary Table S1). The 3' region showed 12 peaks ranging in fractional intensity from 0.06 to 0.18 ( $\Delta\Delta G \sim 1.1 \text{ RT}$ ). Thus the free energy of folding of these states spans a range of  $<6 \text{ kJ/mol}$ . For comparison, the overall free energy of folding under these conditions is of the order  $36 \text{ kJ mol}^{-1}$  (cf. Table 1).

We have analyzed the unfolding of the different fractions using CD, as described in the 'Materials and Methods' section. Fits to a two-state transition for the melting for each fraction are shown in Figure 7, with the derived  $T_m$  and van't Hoff enthalpies given in Table 1. This corresponds to an equilibrium folding experiment, and the CD spectra maintain isodichroic points throughout the folding profile, despite the end state being a mixture of different species (albeit most having similarly shaped spectra). As expected, all five fractions showed the same thermodynamic folding parameters within experimental error ( $T_m = 338 \pm 0.5 \text{ K}$ ,  $\Delta H_{\text{vH}} = 306.5 \pm 10.3 \text{ kJ mol}^{-1}$ ). The apparent free energy of folding was  $36.3 \pm 1.2 \text{ kJ/mol}$  at 298 K, and is in the range of estimates for other quadruplexes of comparable size (30,49). We note that this free energy of folding is similar to that observed for globular proteins of the same molecular weight (50).

The shape of the refolding curve indicates that the  $T_m$  and  $\Delta H$  values for each species cannot differ greatly from one another, as this would give rise to quite different shapes (cf. Supplementary Figure S4). This is in agreement with the observation that the major species reach similar populations under slow annealing conditions.

The melting and annealing profiles of the unfractionated mixture are essentially superimposable (cf. Figure 2B and Table 1). In contrast, the individual fractions showed quite different annealing and melting profiles, as expected because the isolated fractions are not at equilibrium. Furthermore, during melting the isolated fractions do not maintain an isodichroic point (Supplementary Figure S5). Thus the starting point for melting is different for each fraction, though the end point is the same. 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. For example, starting with pure fraction 2, as the temperature is increased so that a significant amount unfolds, some will re-equilibrate into all of the other fractions, which then also eventually melt; the concentration of these states is then expected first to increase and then decrease with



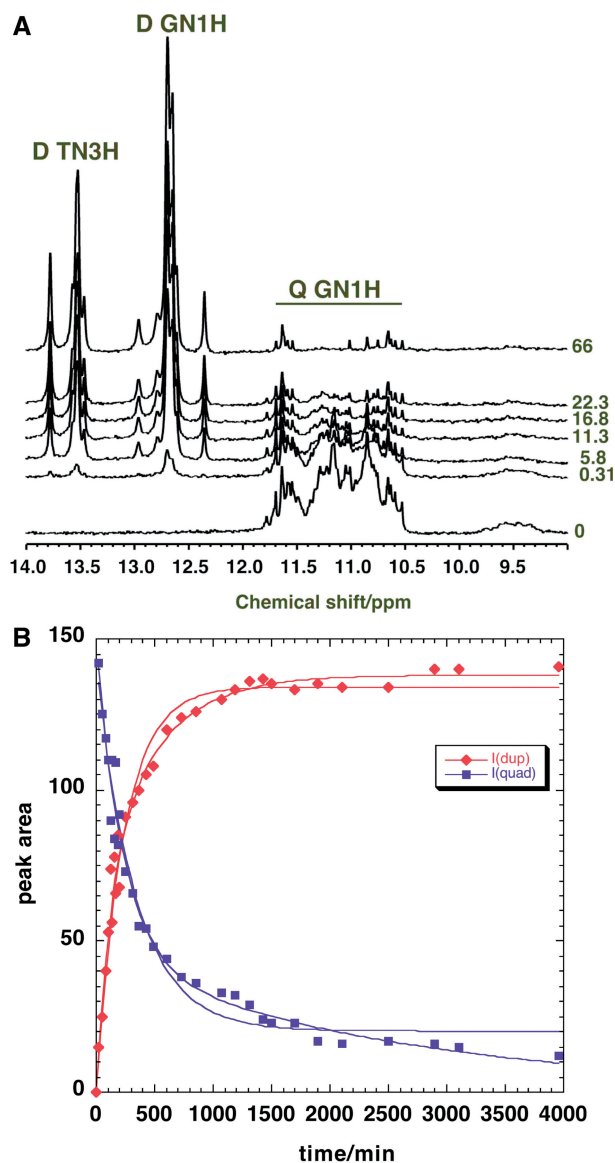
**Figure 7.** Thermodynamics of folding and unfolding of purified fractions of AS1411. Individual fractions were monitored by CD as a function of temperature as described in the 'Materials and Methods' section. The monitoring wavelength was 263 nm for all fractions except number 2, which was monitored at 277 nm (cf. Figure 2). Fraction 1 (red), fraction 2 (blue) fit to Equation (1) with  $T_m = 339 \pm 0.2 \text{ K}$  and  $\Delta H = 60 \pm 2 \text{ kcal/mol}$ , fraction 3 (black), fraction 4 (green) fit to Equation (1) with  $T_m = 343 \pm 0.1 \text{ K}$  and  $\Delta H = 70 \pm 1.4 \text{ kcal/mol}$  and fraction 5 (magenta).

increasing temperature. The effect on the CD intensity will also be determined by the specific CD intensity of each contributing species.

#### Unfolding kinetics of unfractionated AS1411

The folding rate constant for intramolecular G-quadruplexes in  $\text{K}^+$  buffers is typically fast (48), whereas the unfolding rate constant is slow (42,51,52). Figure 8A shows the imino proton spectrum of AS1411 at 30°C before and after adding excess complementary C strand. Over time, the quadruplex imino proton resonances decrease in intensity, with a concomitant increase in the intensity of the duplex resonances. The complementary C-rich strand under these conditions showed no imino proton resonances, and the non-exchangeable protons were sharp, consistent with a flexible unstructured oligonucleotide (data not shown). Substantial intensity of the quadruplex forms was retained up to 20 h, and by 66 h there was a single especially resistant species remaining that accounts for 10–15% of the initial intensity. The proton spectrum of the duplex form has GC:AT imino proton intensities in the ratio 2:1, as expected from the sequence. Figure 8B shows the time dependence of the area of the duplex resonances (12–14 ppm) and of the quadruplex GN1H resonances (10.2–12 ppm). The data cannot be fitted to a single exponential, and there are at least two classes of processes. The biexponential fits returned rate constants of  $0.0073 \pm 0.002$  and  $0.0016 \pm 0.0006 \text{ min}^{-1}$  (duplex) and  $0.0043 \pm 0.0006$  and  $0.00038 \pm 0.0001 \text{ min}^{-1}$  for the quadruplex resonances. An additional experiment carried out using a lower concentration of the complementary strand gave similar





**Figure 8.** Dissociation kinetics of AS1411. (A) Imino proton spectrum of AS1411 with excess complementary strand. NMR spectra were recorded at 30°C at 18.8T with a 3 s recycle time. The first spectrum began recording 4 min after mixing 152  $\mu$ M quadruplex DNA with 381  $\mu$ M complementary strand in 70 mM KCl, 14 mM K-phosphate buffer, 7% D<sub>2</sub>O, pH = 7.2. Spectra were recorded every 30 min for the first 24 h, and then every 2 h for the subsequent 18 h. A final spectrum was recorded at 66 h. (B) Time dependence of the area of the duplex and quadruplex GN1H resonances. The duplex and quadruplex imino proton regions were integrated at each time point. The data were fitted to different models using Kaleidagraph. Blue squares (quadruplex GN1H) were fitted to the sum of two exponentials or an exponential with a linear slope. The two exponential fit was superior. The duplex data (red triangles) were also fit to the sum of two exponentials or to a form of the kind  $a(1 - e^{-kt}) + bt$ .

kinetics, suggesting that these rate constants do reflect the unfolding of the quadruplexes. The most recalcitrant species still present at 66 h appears to decay considerably more slowly than this. However, warming to 50°C for 24 h caused the complete disappearance of the quadruplex peaks, indicating that the persistence at 30°C was kinetic rather than thermodynamic. Furthermore, the spectrum

of this long-lived species was essentially identical to that of isolated fraction 2, which accounted of ~20% of the peak area in Figure 3. Hence, the half-life of this quadruplex must be of the order 66 h under these conditions. This implies that the kinetic stability of the different species varies over at least a 10-fold range. As for other such intramolecular quadruplexes (42,51,53), the AS1411 quadruplexes are extremely kinetically stable, but not especially thermodynamically stable.

## DISCUSSION

The NMR spectroscopy and chromatography conclusively demonstrate the co-existence of at least eight distinguishable quadruplex species under slow annealing conditions in K<sup>+</sup> buffer whereas in Na<sup>+</sup> buffer, mainly one species is formed. As the individual fractions have essentially the same electrophoretic mobility and sedimentation coefficient, it is clear that the chromatography is not separating them based on size/shape alone, but perhaps subtle differences in interaction with the matrix.

The annealing and melting profiles are deceptively simple, in part because at the lower temperatures, the folding and unfolding is not always close to equilibrium. When cooling slowly from the strand state, the exchange among folded forms via the unfolded state remains rapid compared with the cooling rate at least above 40°C, which is below the apparent melting temperature (67°C, Table 1). Under these conditions, the distribution of folded forms approaches the expected thermodynamic distribution. As the temperature is lowered further these become locked-in kinetically. In contrast, individual conformers isolated chromatographically are no longer at equilibrium, but exchange among conformers is very slow. However, the equilibrium distribution should be a mixture of several states, so that as a purified conformer unfolds (well below the global  $T_m$  where the concentration of the unfolded state is infinitesimal), the unfolded state refolds into all of the other conformations, with a relatively high rate constant. This renders the unfolding of the purified conformer essentially irreversible, i.e. the re-equilibration is determined by the unidirectional unfolding rate constant, which is very small under these conditions (see above). We have simulated the unfolding of a purified species under certain assumptions, and compared the kinetic unfolding curves with those expected for a purely thermodynamic unfolding of the mixture (Supplementary Figure S6). Such simulations can account for the observed unfolding profiles (cf. Figure 7), but complete quantitative analysis is not possible with this current experimental design.

Furthermore, one of the fractions contains a contribution from a dimeric species. We have also simulated the expected melting of a dimer in equilibrium with a monomer via the unfolded state, as a function of concentration (Supplementary Figure S7). This indicates that as expected, the amount of the dimer present depends on many factors including total concentration, and the relative enthalpy differences and melting temperatures of the dimer and monomer states.

Although it is well known that many G-rich oligonucleotide species exist as mixtures in solution (1,26,30), analyses of biophysical properties such as thermodynamic stability, kinetic pathway or spectroscopic features typically ignore this complication. When multiple, non-independent species form, the apparent two-state melting data do not yield thermodynamic parameters that are characteristic of an individual species, but a complex average over the entire system. This may in part explain the substantial discrepancies in the literature regarding even folding enthalpies (30,49).

This study makes clear the need for separation and high-resolution techniques such as NMR when studying quadruplex formation, structure, and biophysical properties. Structural and biophysical studies of quadruplex DNA have been hindered by our lack of ability to isolate, study, and understand the individual species that may occur in solution or *in vivo*. Low resolution techniques such as gel electrophoresis, AUC, CD, UV or even calorimetry are not always able to distinguish between separate (isomeric) species in mixtures of quadruplex configurations from the same sequence.

While many NMR and crystallization studies have yielded quadruplex structures, the relevance of these results is uncertain, especially with respect to the biologically relevant species. Every sample that has yielded a 'pure' conformation for structural studies has been the result of sequence modification. In fact it has become a commonly accepted practice to ignore the majority of species in solution and to focus on the result of sequence modifications that single out a single species (24–26,46). Furthermore crystal structures may be the result of selective crystallization of one species over another (21). A well known example is the crystal structure for the human telomere resulting from a sequence that has been shown by NMR to generate multiple conformations in solution (24,27). Thus, the reported crystal and NMR structures should be treated for what they are, isolated structures from an ensemble. This has obvious implications for designing quadruplex binding ligands and how the targets relate to the form(s) assayed either *in vitro* or *in vivo*. In addition to simple factors such as ion concentration and temperature, there are additional complexities *in vivo* such as supercoiling, the G-rich sequences are tethered at one or both ends, and the influence of non-specific high protein concentrations, all of which are difficult to replicate *in vitro* (30). Under these circumstances, it is uncertain which conformation of a specific sequence is the correct or biologically relevant one.

Despite our lack of understanding of these complex systems quadruplex DNA is rapidly becoming an extremely attractive target for clinical and nanotechnology applications and significant strides have been made towards therapies ranging from novel anticancer drugs to new antiviral treatments. Thus the implications of this research for drug discovery, nanotechnology, clinical and other applications are clear.

## CONCLUSIONS

Given the remarkable structural polymorphism that a single quadruplex-forming sequence has displayed, the work presented here clearly shows some potential pitfalls in using low-resolution techniques to characterize a single quadruplex from any given sequence. Techniques such as circular dichroism, electrophoresis, UV or CD melting, and analytical ultracentrifugation are useful and valuable in the analysis of quadruplex structures. However, these individual techniques alone cannot characterize a particular quadruplex fold, or more importantly, clearly identify the presence of more than one species. A higher resolution technique, such as NMR, is essential for this purpose.

We have reported a simple separation method that gives potential access to the isolation of the individual quadruplexes from any given sequence, notably without any sequence modification. The sequence reported here is not unique in its structural polymorphism and may be more the rule than the exception.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

The authors would like to thank Dr Nichola C. Garbett of the University of Louisville James Graham Brown Cancer Center Biophysics Core Facility for assistance in obtaining the biophysical data.

## FUNDING

National Institutes of Health (CA113735-01 to J.O.T.), National Institutes of Health Grant Number P20RR018733 from the National Center for Research Resources, and the Kentucky Challenge for Excellence. Funding for open access charge: National Institutes of Health (CA113735-01), NCI grant.

*Conflict of interest statement.* P.J.B. and J.O.T. have a financial interest in Antisoma, PLC (London, England) as a shareholder. P.J.B. is also a consultant and recipient of research support from Antisoma, PLC.

## REFERENCES

1. Patel, D.J., Phan, A.T. and Kuryavyy, V. (2007) Human telomere, oncogenic promoter and 5'-UTR G-quadruplexes: diverse higher order DNA and RNA targets for cancer therapeutics. *Nucleic Acids Res.*, **35**, 7429–7455.
2. Kipling, D. (1995) *The Telomere*. Oxford University Press, Oxford.
3. Okazaki, S., Tsuchida, K., Maekawa, H., Ishikawa, H. and Fujiwara, H. (1993) Identification of a pentanucleotide telomeric sequence, (TTAGG)<sub>n</sub>, in the silkworm *Bombyx mori* and in other insects. *Mol. Cell. Biol.*, **13**, 1424–1432.
4. Rawal, P., Kummarasetti, V.B., Ravindran, J., Kumar, N., Halder, K., Sharma, R., Mukerji, M., Das, S.K. and Chowdhury, S. (2006) Genome-wide prediction of G4 DNA as regulatory

- motifs: role in Escherichia coli global regulation. *Genome Res.*, **16**, 644–655.
5. Robertson, H.M. and Gordon, K.H. (2006) Canonical TTAGG-repeat telomeres and telomerase in the honey bee, *Apis mellifera*. *Genome Res.*, **16**, 1345–1351.
  6. Sahara, K., Marec, F. and Traut, W. (1999) TTAGG telomeric repeats in chromosomes of some insects and other arthropods. *Chromosome Res.*, **7**, 449–460.
  7. Huppert, J.L. and Balasubramanian, S. (2005) Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.*, **33**, 2908–2916.
  8. Siddiqui-Jain, A., Grand, C.L., Bearss, D.J. and Hurley, L.H. (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-MYC transcription. *Proc. Natl Acad. Sci. USA*, **99**, 11593–11598.
  9. Simonsson, T., Pecinka, P. and Kubista, M. (1998) DNA tetraplex formation in the control region of c-myc. *Nucleic Acids Res.*, **26**, 1167–1172.
  10. Hud, N.V., Smith, F.W., Anet, F.A.L. and Feigon, J. (1996) The selectivity for K<sup>+</sup> versus Na<sup>+</sup> in DNA quadruplexes is dominated by relative free energies of hydration: A thermodynamic analysis by H-1 NMR. *Biochemistry*, **35**, 15383–15390.
  11. Engelhart, A.E., Plavec, J., Persil, O. and Hud, N.V. (2008) In Hud, N.V. (ed.), *Nucleic Acid-Metal Ion Interactions*. Royal Society of Chemistry, London, pp. 118–153.
  12. Sun, D., Liu, W.J., Guo, K.X., Rusche, J.J., Ebbinghaus, S., Gokhale, V. and Hurley, L.H. (2008) The proximal promoter region of the human vascular endothelial growth factor gene has a G-quadruplex structure that can be targeted by G-quadruplex-interactive agents. *Mol. Cancer Ther.*, **7**, 880–889.
  13. Hurley, L.H. and Siddiqui-Jain, A. (2005) Developing therapeutics to target oncogenes. *Genet. Eng. News*, **25**, 26.
  14. Huppert, J.L. (2007) Four-stranded DNA: cancer, gene regulation and drug development. *Philos. Trans. Roy. Soc. A – Math. Phys. Eng. Sci.*, **365**, 2969–2984.
  15. Herbert, B.-S., Huppert, J.L., Johnson, F.B., Lane, A.N. and Phan, A.T. (2009) Meeting Report: Second International Meeting on Quadruplex DNA. *Biochimie*, **91**, 1059–1065.
  16. Miller, D.M., Laber, D.A., Bates, P.J., Trent, J.O., Taft, B.S. and Kloecker, G.H. (2006) Extended phase I study of AS1411 in renal and non-small cell lung cancers. *Ann. Oncol.*, **17**, 147–148.
  17. Bates, P.J., Kahlon, J.B., Thomas, S.D., Trent, J.O. and Miller, D.M. (1999) Antiproliferative activity of G-rich oligonucleotides correlates with protein binding. *J. Biol. Chem.*, **274**, 26369–26377.
  18. Bishop, J.S., Guy-Caffey, J.K., Ojwang, J.O., Smith, S.R., Hogan, M.E., Cossam, P.A., Rando, R.F. and Chaudhary, N. (1996) Intramolecular G-quartet motifs confer nuclease resistance to a potent anti-HIV oligonucleotide. *J. Biol. Chem.*, **271**, 5698–5703.
  19. Suzuki, J., Miyano-Kurosaki, N., Kuwasaki, T., Takeuchi, H., Kawai, G. and Takaku, H. (2002) Inhibition of human immunodeficiency virus type 1 activity in vitro by a new self-stabilized oligonucleotide with guanosine-thymidine quadruplex motifs. *J. Virol.*, **76**, 3015–3022.
  20. Cogoi, S., Quadrioglio, F. and Xodo, L.E. (2004) G-rich oligonucleotide inhibits the binding of a nuclear protein to the Ki-ras promoter and strongly reduces cell growth in human carcinoma pancreatic cells. *Biochemistry*, **43**, 2512–2523.
  21. Jing, N., Sha, W., Li, Y., Xiong, W. and Twardy, D.J. (2005) Rational drug design of G-quartet DNA as anti-cancer agents. *Curr. Pharm. Des.*, **11**, 2841–2854.
  22. Qi, H., Lin, C.P., Fu, X., Wood, L.M., Liu, A.A., Tsai, Y.C., Chen, Y., Barbieri, C.M., Pilch, D.S. and Liu, L.F. (2006) G-quadruplexes induce apoptosis in tumor cells. *Cancer Res.*, **66**, 11808–11816.
  23. Yang, Y. and Patel, D.J. (1993) Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex. *Structure*, **1**, 263–282.
  24. Dai, J., PUNCHIHEWA, C., Ambrus, A., Chen, D., Jones, R.A. and Yang, D. (2007) Structure of the intramolecular human telomeric G-quadruplex in potassium solution: a novel adenine triple formation. *Nucleic Acids Res.*, **35**, 2440–2450.
  25. Dai, J., Carver, M., PUNCHIHEWA, C., Jones, R.A. and Yang, D. (2007) Structure of the Hybrid-2 type intramolecular human telomeric G-quadruplex in K<sup>+</sup> solution: insights into structure polymorphism of the human telomeric sequence. *Nucleic Acids Res.*, **35**, 4927–4940.
  26. Phan, A.T., Kuryavyi, V., Luu, K.N. and Patel, D.J. (2007) Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K<sup>+</sup> solution. *Nucleic Acids Res.*, **35**, 6517–6525.
  27. Parkinson, G.N., Lee, M.P. and Neidle, S. (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature*, **417**, 876–880.
  28. Webba da Silva, M. (2007) Geometric formalism for DNA quadruplex folding. *Chemistry*, **13**, 9738–9745.
  29. Li, J., Correia, J.J., Wang, L., Trent, J.O. and Chaires, J.B. (2005) Not so crystal clear: the structure of the human telomere G-quadruplex in solution differs from that present in a crystal. *Nucleic Acids Res.*, **33**, 4649–4659.
  30. Lane, A.N., Chaires, J.B., Gray, R.D. and Trent, J.O. (2008) Stability and kinetics of G-quadruplex structures. *Nucleic Acids Res.*, **36**, 5482–5515.
  31. Luu, K.N., Phan, A.T., Kuryavyi, V., Lacroix, L. and Patel, D.J. (2006) Structure of the human telomere in K<sup>+</sup> solution: an intramolecular (3 + 1) G-quadruplex scaffold. *J. Am. Chem. Soc.*, **128**, 9963–9970.
  32. Wang, Y. and Patel, D.J. (1993) Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex. *Structure*, **1**, 263–282.
  33. Dapic, V., Bates, P.J., Trent, J.O., Rodger, A., Thomas, S.D. and Miller, D.M. (2002) Antiproliferative activity of G-quartet-forming oligonucleotides with backbone and sugar modifications. *Biochemistry*, **41**, 3676–3685.
  34. Girvan, A.C., Teng, Y., Casson, L.K., Thomas, S.D., Juliger, S., Ball, M.W., Klein, J.B., Pierce, W.M., Barve, S.S. and Bates, P.J. (2006) AGRO100 inhibits activation of nuclear factor-kappa B (NF-kappa B) by forming a complex with NF-kappa B essential modulator (NEMO) and nucleolin. *Mol. Cancer Ther.*, **5**, 1790–1799.
  35. Teng, Y., Girvan, A.C., Casson, L.K., Pierce, W.M., Qian, N., Thomas, S.D. and Bates, P.J. (2007) AS1411 alters the localization of a complex containing protein arginine methyltransferase 5 and nucleolin. *Cancer Res.*, **67**, 10491–10500.
  36. Dapic, V., Abdomerovic, V., Marrington, R., Peberdy, J., Rodger, A., Trent, J.O. and Bates, P.J. (2003) Biophysical and biological properties of quadruplex oligodeoxyribonucleotides. *Nucleic Acids Res.*, **31**, 2097–2107.
  37. Asensio, J.L., Lane, A.N., Dhesi, J., Bergqvist, S. and Brown, T. (1998) The contribution of cytosine protonation to the stability of parallel DNA triple helices. *J. Mol. Biol.*, **275**, 811–822.
  38. Philo, J.S. (2000) A method for directly fitting the time derivative of sedimentation velocity data and an alternative algorithm for calculating sedimentation coefficient distribution functions. *Anal. Biochem.*, **279**, 151–163.
  39. Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.*, **78**, 1606–1619.
  40. Bonifacio, G.F., Brown, T., Conn, G.L. and Lane, A.N. (1997) Comparison of the electrophoretic and hydrodynamic properties of DNA and RNA oligonucleotide duplexes. *Biophys. J.*, **73**, 1532–1538.
  41. Hellman, L., Rodgers, D. and Fried, M. (2009) Phenomenological partial-specific volumes for G-quadruplex DNAs. *Eur. J. Biophys.*, **39**, 389–396.
  42. Zhao, Y., Kan, Z.Y., Zeng, Z.X., Hao, Y.H., Chen, H. and Tan, Z. (2004) Determining the folding and unfolding rate constants of nucleic acids by biosensor. Application to telomere G-quadruplex. *J. Am. Chem. Soc.*, **126**, 13255–13264.
  43. Stegle, O., Payet, L., Mergny, J.L., MacKay, D.J. and Leon, J.H. (2009) Predicting and understanding the stability of G-quadruplexes. *Bioinformatics*, **25**, i374–382.
  44. Olsen, C.M., Gmeiner, W.H. and Marky, L.A. (2006) Unfolding of G-quadruplexes: Energetic, and ion and water contributions of G-quartet stacking. *J. Phys. Chem. B*, **110**, 6962–6969.
  45. Fernandes, M.X., Ortega, A., Martinez, M.C.L. and de la Torre, J.G. (2002) Calculation of hydrodynamic properties of small nucleic acids from their atomic structure. *Nucleic Acids Res.*, **30**, 1782–1788.



46. Dai, J.X., Chen, D., Jones, R.A., Hurley, L.H. and Yang, D.Z. (2006) NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region. *Nucleic Acids Res.*, **34**, 5133–5144.
47. Feigon, J., Koshlap, K.M. and Smith, F.W. (1995) <sup>1</sup>H NMR spectroscopy of DNA triplexes and quadruplexes. *Methods Enzymol.*, **261**, 225–255.
48. Gray, R.D. and Chaires, J.B. (2008) Kinetics and mechanism of K<sup>+</sup>- and Na<sup>+</sup>-induced folding of models of human telomeric DNA into G-quadruplex structures. *Nucleic Acids Res.*, **36**, 4191–4203.
49. Chaires, J.B. (2010) Human telomeric G-quadruplex: Thermodynamic and Kinetic Studies. *FEBS J*, **277**, 1098–1106.
50. Robertson, A.D. and Murphy, K.P. (1997) Protein Structure and the Energetics of Protein Stability. *Chem. Rev.*, **97**, 1251–1268.
51. Green, J.J., Ying, L.M., Klenerman, D. and Balasubramanian, S. (2003) Kinetics of unfolding the human telomeric DNA quadruplex using a PNA trap. *J. Am. Chem. Soc.*, **125**, 3763–3767.
52. Brown, N.M., Rachwal, P.A., Brown, T. and Fox, K.R. (2005) Exceptionally slow kinetics of the intramolecular quadruplex formed by the *Oxytricha* telomeric repeat. *Org. Biomol. Chem.*, **3**, 4153–4157.
53. Green, J.J., Ladame, S., Ying, L.M., Klenerman, D. and Balasubramanian, S. (2006) Investigating a quadruplex-ligand interaction by unfolding kinetics. *J. Am. Chem. Soc.*, **128**, 9809–9812.