



HAL
open science

Electrospray Mass Spectrometry of Telomeric RNA (TERRA) Reveals the Formation of Stable Multimeric G-Quadruplex Structures

Gavin W. Collie, Gary N. Parkinson, Stephen Neidle, Frédéric Rosu, Edwin de Pauw, Valérie Gabelica

► **To cite this version:**

Gavin W. Collie, Gary N. Parkinson, Stephen Neidle, Frédéric Rosu, Edwin de Pauw, et al.. Electrospray Mass Spectrometry of Telomeric RNA (TERRA) Reveals the Formation of Stable Multimeric G-Quadruplex Structures. *Journal of the American Chemical Society*, 2010, 132 (27), pp.9328-9334. 10.1021/ja100345z . hal-01524165

HAL Id: hal-01524165

<https://hal.science/hal-01524165v1>

Submitted on 17 May 2017

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.

Electrospray mass spectrometry of telomeric RNA (TERRA) reveals the formation of stable multimeric G-quadruplex structures

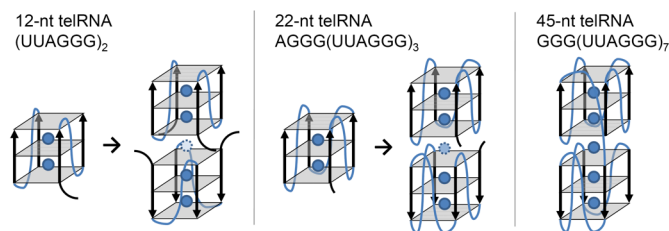
Gavin W. Collie,¹ Gary N. Parkinson,¹ Stephen Neidle,¹ Frédéric Rosu,² Edwin De Pauw,² Valérie Gabelica²

¹ CRUK Biomolecular Structure Group, The School of Pharmacy, University of London, WC1N 1AX, London, UK.

² Physical Chemistry and Mass Spectrometry Laboratory, Department of Chemistry, University of Liège, Belgium, v.gabelica@ulg.ac.be

Abstract

We report on the self-assembled structures formed by 12-mer, 22-mer, and 45-mer telomeric RNA (telRNA/TERRA) sequences compared to their DNA analogues, as studied by electrospray mass spectrometry, circular dichroism, and thermal denaturation. The major difference between telomeric RNA and DNA sequences is the ability of telomeric RNA to form higher-order dimeric assemblies, initiated by cation-mediated stacking of two parallel G-quadruplex subunits. 5'-5' stacking had been observed recently by NMR for the r(GGGUUAGGGU) 10-mer (Martadinata & Phan, *J. Am. Chem. Soc.* **2009**, *131*, 2570); the present work shows that stacking also occurs for the 22-mer containing four G-tracts and for the 45-mer containing eight G-tracts, suggesting a general structural feature of telomeric RNA. The importance of kinetic effects in multimer formation, unfolding, and structural rearrangements is also highlighted.



Introduction

The telomeric ends of mammalian chromosomes comprise tandem repeats of the sequence d(TTAGGG), together with associated proteins. Mammalian telomeric DNA is typically 5-8 kilobases in length, with a 3' single-stranded overhang of 100-200 bases.¹ These G-rich sequences are able to form four-stranded G-quadruplex structures, which have as their core the stacking of at least two guanine quartets, a planar arrangement of four guanine bases held together by eight Hoogsteen hydrogen bonds.^{2,3} Mammalian telomeres are systematically shortened by 100-200 nucleotides during each round of cell division.⁴ This is counteracted in most cancer cell types by the enzyme telomerase, a heterodimeric reverse transcriptase, which catalyses the synthesis of telomeric DNA repeats.^{5,6} One approach to inhibit telomerase activity involves stabilising telomeric DNA as G-quadruplex structures by the use of small-molecule quadruplex-binding ligands.⁷⁻⁹

The assumption that the telomere was transcriptionally silent has been overturned by recent findings that telomeric DNA can be transcribed, the resulting RNA transcripts ranging in size from ~100 nucleotides up to at least 9 kb, composed of tandem repeats of the sequence r(UUAGGG).^{10,11} It has been shown by circular dichroism (CD)¹²⁻¹⁷, NMR^{15,18}, and X-ray crystallography¹⁹ that these G-rich telomeric RNA sequences are able to form parallel G-quadruplex structures in sodium as well as in potassium ionic conditions,²⁰ and are significantly more stable than their DNA counterpart in the presence of potassium,^{13,16,21} but not sodium ions.^{16,21}

The role that telomeric RNA (and by implication, RNA quadruplexes) plays in telomere biology is yet to be fully determined.^{22,23} Long telomeric RNA molecules are compact and there is evidence that they form higher-order assemblies based on G-quadruplex subunits.¹⁷ An NMR study of a 10-mer RNA sequence r(GGGUUAGGGU) revealed that two parallel G-quadruplex units assemble via 5'-5' stacking.¹⁸ The aim of the work presented here is to provide information on the ways in which longer telomeric RNA quadruplexes can interact with each other, using electrospray ionisation mass spectrometry (ESI-MS) and spectrophotometric techniques. ESI-MS enables unambiguous determination of the number of strands in non-covalent assemblies, and can resolve mixtures of stoichiometries.²⁴⁻³¹ Moreover, because G-quadruplexes accommodate cations between consecutive G-quartets, counting the number of strongly bound cations using ESI-MS provides additional insight on the structure and stability of the assemblies.^{31,32} We have therefore studied telomeric RNA sequences with two to eight G-tracts, and compared them with their DNA counterparts.

Materials and Methods

Oligonucleotides and G-quadruplex formation. All sequences were purchased from Eurogentec (Belgium) and used without further purification. The sequences used were: RNA12 (r(UUAGGG)₂, MW = 3892.4 Da), DNA12 (d(TTAGGG)₂, MW = 3756.5 Da), RNA22 (r[AGGG(UUAGGG)₃], MW = 7234.4 Da), DNA22 (d[AGGG(TTAGGG)₃], MW = 6966.6 Da), RNA45 (r[GGG(UUAGGG)₇], MW = 14813.9 Da) and DNA45 (d[GGG(TTAGGG)₇], MW = 14290.3 Da). For the kinetics experiment monitored by mass spectrometry, the internal standard sequence dT₆ (monoisotopic mass = 1762.318 Da) was also used. Ammonium acetate (BioUltra ~5 M, for molecular biology) was provided by Fluka (Sigma-Aldrich NV/SA, Bornem, Belgium). Water was either nuclease-free water from Ambion (Applied Biosystems, Lennik, Belgium), or bi-distilled in house. RNA sequences were not degraded in bi-distilled water for > 6 months when folded into G-quadruplexes (ESI-MS experiments not shown), but nuclease-free water must be used for long-term storage of telRNA in water. Except for the kinetics experiments, G-quadruplexes were formed by annealing (heating the oligonucleotides for 5 minutes at 85°C, followed by cooling at 1°C/min to 75°C, then cooling at 0.2 °C/min to 20°C) in 100 mM ammonium acetate. The faster cooling to 75 °C is to minimize thermal degradation of the samples. Unless stated otherwise, the single-strand concentrations upon annealing were 1 mM (RNA12, DNA12, DNA22 and RNA22), 300 µM (DNA45), or 80 µM (RNA45). The reason for this difference was the paucity of starting material. Annealings at lower strand concentrations of RNA12 and RNA22 were also tested (see results and discussion). For studying the kinetics of G-quadruplex formation at room temperature, the single strands in water were heated at 85 °C for 5 minutes to disrupt any pre-formed structures, and then equilibrated at RT. The starting time of G-quadruplex formation kinetics is the addition of concentrated ammonium acetate to reach a final ammonium acetate concentration of 100 mM and the desired strand concentration.

Mass Spectrometry. All ESI-MS experiments were carried out on a Q-TOF Ultima Global mass spectrometer, fitted with an electrospray source. The instrument was used in the negative ion mode, with the capillary voltage set to 2.2 kV, the cone voltage set to 100 V and the collision voltage set to 2 V. The RF lens 1 voltage varied between 60 V and 210 V. The source temperature was set to 40 °C and the desolvation temperature was set to 60 °C. The source pressure was set to 3.40 mbar. Data processing was performed using MassLynx 4.0. For the analysis of annealed G-quadruplexes, the samples were injected at a final strand concentration of 20 µM for the 12-mers, 10 µM for the 22-mers, and 5 µM for the 45-mers, at a rate of 4.0 µL min⁻¹. In order to improve the electrospray signal, methanol was added to all samples prior to injection, to a final concentration of 20%. The mass spectra were smoothed (mean function, 2 × 20 channels) and background subtracted (order 10, 1% below curve). For monitoring the kinetics of G-quadruplex formation by ESI-MS, unfolded RNA was pre-mixed with the internal standard dT₆ in water/methanol. Concentrated ammonium acetate was added to reach final concentrations of 10 µM RNA, 4 µM dT₆, 100 mM NH₄OAc. The kinetics of self-assembly was tested in 0%, 10%, and 20% volume percentage of methanol. The spectrum recording was started simultaneously

with ammonium addition. The sample was vortexed, loaded into the 250-mL syringe, the spray was initiated as quickly as possible by manually pushing the syringe, and the flow rate was then stabilized at $4.0 \mu\text{L min}^{-1}$. The time lapse between ammonium addition and spray stabilization is typically one minute. The detailed data treatment to reconstruct the evolution of the concentrations from the evolution of the relative intensities and to correct for the different electrospray response factors is described in detail elsewhere.³³ Briefly, for each stoichiometry adopted by the RNA, the ratio between the area of a peak attributable unambiguously to that stoichiometry and the area of a peak of the internal standard (here dT_6^{2-}), is calculated at each time-point. By assuming that the electrospray response of the RNA and dT_6 are proportional throughout the experiment, and by using the mass balance equation applied to the RNA (the total strand concentration being $10 \mu\text{M}$), the relative response factors of each RNA stoichiometry can be determined. These response factors are then used to recalculate the concentration of each RNA stoichiometry at each time point.

Circular dichroism. CD experiments were performed on a Jasco J-810 spectropolarimeter. The CD spectra of the annealed samples were recorded using in 0.4-mm path quartz cuvettes. The annealed samples were diluted to final strand concentration of $20 \mu\text{M}$ for the 12-mers, $10 \mu\text{M}$ for the 22-mers, and $5 \mu\text{M}$ for the 45-mers. The spectra of the annealed structures consisted of an average of 10 scans, recorded from 200 to 400 nm, using a scan rate of 100 nm/min. The spectra recorded as a function of time after ammonium addition to RNA22 consisted of an average of 2 scans, recorded from 200 to 400 nm in 1-cm path quartz cells, using a scan rate of 100 nm/min. The spectra presented in the figures were smoothed (Savitsky-Golay filter, 5% level) using PeakFit 4.11 (Systat Softwares).

UV spectrophotometry. Experiments were carried out on a Uvikon XS spectrophotometer (Secomam). The absorption spectra at 17°C and the thermal difference spectra (difference between absorption spectrum at 90°C and 17°C) of the annealed samples were recorded on the same samples that were subjected to the CD analysis. For thermal denaturation experiments, 1-cm path length black quartz cells were used (Hellma, type No. 115B-QS, France). Annealed oligonucleotides were diluted in $100 \text{ mM NH}_4\text{OAc}$ so as to obtain an absorbance of 1 at 260 nm at 17°C , to ensure a high enough absorbance at 295 nm. For thermal denaturation, absorbance was monitored as a function of the temperature at 240, 260, 275 and 295 nm. The temperature was first increased from 14°C to 92°C , then decreased to 14°C , and increased a second time to 92°C . The gradient was 0.2°C/min in all cases.

Results

ESI-MS shows that telomeric RNA sequences preferably form higher-order assemblies containing two G-quadruplex subunits

We have firstly compared the ESI-MS spectra of telomeric RNA sequences containing two to eight G-tracts with their DNA counterparts. All samples were annealed in ammonium acetate. The smallest G-quadruplex that can be formed with the 12-mers is dimeric, since each strand contains two G-tracts, and four G-tracts are needed to form a telomeric G-quadruplex. The 22-mers, which contain four G-tracts, are expected to form intramolecular (monomeric) G-quadruplexes, and the 45-mers, which contain eight G-tracts, are expected to form two intramolecular G-quadruplex subunits connected by a (T/U)₂A linker.

The ESI-MS spectra recorded for all six sequences are shown in Figure 1. The two shorter DNA sequences used for the study were either identical (DNA12) or very similar (DNA22) to the sequences studied previously by ESI-MS and ion mobility spectrometry.^{34,35} DNA12 (Figure 1a) was found to form a dimer, but the dimer is much less abundant than the monomer. DNA22 (Figure 1b) was detected mainly as a monomer, but broad peaks corresponding to a dimer can also be seen when zooming on the spectra. The presence of dimer of DNA22 was confirmed by experiments on an electrospray Synapt HDMS instrument, where ion mobility separation allows better differentiation of monomer and dimer signals (data not shown). The DNA45 sequence was also detected exclusively as a monomer (Figure 1c).

The RNA telomeric sequences behave very differently from their DNA counterpart. RNA12 (Figure 1d) forms a dimeric quadruplex much more readily than DNA12, but the most striking difference between them is that the major RNA12 self-assembled structure is a tetramer in ammonium acetate. This sequence has been previously studied by Xu and co-authors using MALDI-TOF mass spectrometry in sodium chloride.¹⁵ The spectra showed mainly monomer and some dimer. Two reasons can explain this difference: either the tetramer can form in ammonium but not in sodium solution, or the tetramer can form in sodium as well, but does not survive the MALDI ionization conditions (50% acetonitrile in the sample preparation,¹⁵ and higher internal energies in MALDI than in electrospray^{36,37}).

In the case of RNA22 (Figure 1e), we found both monomeric (intramolecular) and dimeric G-quadruplex structures, the predominant species being the dimer. A trimer was also detected. The trimer became undetectable when the annealing was performed at 300 μ M strand concentration or lower, but the dimer remained the major species even if the annealing is performed at 10 μ M strand concentration. The dimer formation by the same RNA22 sequence in KCl has previously been reported by Qi and Shafer, from gel electrophoresis and equilibrium ultracentrifugation experiments.¹² We also performed native gel electrophoresis experiments that confirmed the

formation of dimers by RNA22 in sodium, potassium and ammonium ionic conditions (supporting information Figure S1).

The DNA45 (Figure 1c) and RNA45 (Figure 1f) sequences were both detected as monomers. These monomers can contain two G-quadruplex units; the formation of these G-quadruplex subunits can be determined by counting the number of ammonium ions selectively trapped in the structure. Native gel electrophoresis experiments were performed on samples leftover from the ESI-MS analysis of Figure 1. The gels (supporting information Figure S2) reveal the presence of even higher-order assemblies of RNA than could be detected by ESI-MS (either using the Q-TOF Ultima mass spectrometer, or using a Waters Synapt HDMS equipped with an 8-kDa quadrupole, data not shown). Note that the sensitivity of ethidium bromide staining in gel electrophoresis is proportional to the number of bases in the assembly, while the sensitivity in ESI-MS is proportional to the molar concentration of each assembly. Moreover, the electrospray response of non-covalent assemblies usually decreases when the size increases. Another explanation could be that higher-order assemblies survive in the gel, but fragment into dimeric assemblies even in the soft ESI-MS conditions used here.

The number of trapped ammonium ions suggests the cation-mediated interaction of two G-quadruplex subunits

Figure 2 shows the distribution of the number of ammonium ions preserved inside each G-quadruplex form, at selected charge states, as a function of the source RF lens 1 voltage. Low voltage (60 or 100 V) corresponds to soft conditions where inner ammonium ions are more likely to be retained in the interior of the G-quartet core. Increasing the voltage induces more energetic collisions in the source, which cause the loss of these inner ammonium ions. The exact mechanism of ammonium ion loss in the gas phase is still unknown, but it involves proton transfer from the ammonium ions to the oligonucleotide that results in the loss of neutral ammonia (there is no change in total charge of the oligonucleotide ion).

The first row (Figures 2a-2d) shows the strand stoichiometries corresponding to a single quadruplex unit (dimeric for the 12-mers, monomeric for the 22-mers), at charge state 5-. Each quadruplex subunit has three contiguous G-quartets, and is expected to preserve two ammonium ions. For the DNA12 dimeric G-quadruplex (Figure 2a), two ammonium ions are preserved at 100 V, and lost at 140 V, while for the RNA counterpart (Figure 2b), the two ammonium ions are much more stable. The difference is even more striking when comparing DNA22 (Figure 2c), for which it is impossible to preserve the two ammonium ions even at 60 V, and RNA22 for which two ammonium ions are still preserved up to 140 V (Figure 2d) and even 180 V (not shown). A recent comparison of RNA and DNA quadruplexes, using crystallographic and modelling methods, provides a possible explanation for the significant differences in stability of the RNA and DNA quadruplexes seen in the ESI-MS spectra.¹⁹ The study identified the C2' hydroxyl groups of the RNA to play a key role in RNA quadruplex stability, which were seen to

preferentially form intramolecular hydrogen bonds with the RNA itself (as opposed to intermolecular interactions with water molecules) thereby stabilising the quadruplex through an extended intramolecular hydrogen bonding network.

The second row (Figures 2e-2h) shows the strand stoichiometries corresponding to two quadruplex units, at charge state 8-. For these larger species, there is a broad distribution of ammonium ion adducts at low voltage, suggesting either the presence of alternative folds with additional specific coordination sites, or non-specific cation adduction which is typical of ESI-MS. The number of ammonium ion adducts and how they evolve in response to increased voltages does not give information on the ammonium ion binding sites, but it allows the number of weakly bound and tightly bound ammoniums to be determined. At 140 V, only the most tightly bound ammonium ions can be preserved. For the DNA 45-mer (Figure 2e), all ammonium ions are quite labile. The preferred stoichiometry for the ammonium ions in the native state lies between three and five, but cannot be precisely determined. Ammonium ion stability is much higher in the RNA counterparts containing two G-quadruplex subunits. For the unexpected higher-order RNA assemblies formed by RNA12 and RNA22, the most stable adduct peak (i.e. the one which is preserved at high voltage) corresponds to four ammonium adducts. At charge state 8-, the adduct with five ammonium ions seems to be significantly stable (Figure 2g at 100 V), but increasing the collision energy in the source leads to one ammonia loss, leaving the dimer with four ammonium ions (Figure 2g at 140 V). Further voltage increase leads to the complete loss of the ammonium ions, two at a time. The four-ammonium ion stoichiometry suggests two ammonium ions per subunit. To check this structural interpretation, product ion spectra of the RNA22 dimer were recorded for the more highly charged dimer $[M_2(NH_4)_4]^{9-}$. The separation into the monomers is observed instead of ammonia loss owing to greater Coulombic repulsion (supporting information Figure S3). Remarkably, a significant fraction of the monomers keep two ammonium ions. These observations are consistent with RNA22 comprising two rather distinct quadruplex subunits, each containing two ammonium ions, and hence presumably three G-quartets.

Finally, for the RNA45 sequence (Figure 2h), the adduct with five ammonium ions has particularly high stability, even at 140 V and higher. Importantly, this particular stoichiometry was observed for all charge states, up to 10- (supporting information Figure S4). As for the RNA22 dimer, the ammonium ion loss pattern for RNA45 is $5 \rightarrow 4 \rightarrow 2 \rightarrow 0$, consistent with two G-quadruplex subunits, each having two ammonium ions, plus one ammonium ion tightly stacked between the two subunits. The results therefore suggest a structural model for the higher-order assemblies involving the cation-mediated stacking of G-quadruplex subunits (Figure 3).

Higher-order RNA assemblies formed by annealing in ammonium may not be strictly parallel G-quadruplexes

The remarkable preservation of the ammonium ions at high voltage in each RNA G-quadruplex subunit, and not in each DNA G-quadruplex subunit, is one further example of what appears to be a general trend in the ESI-MS of G-quadruplexes. When sprayed from ammonium acetate solutions, ammonium ions inside mostly parallel G-quadruplexes (e.g. (TGGGGT)₄,^{28,32} Pu22myc and Pu27myc intramolecular quadruplexes³⁴) survive in the gas phase even at relatively harsh conditions, whereas ammonium ions inside antiparallel G-quadruplexes (e.g. (G₃₋₄T₄G₃₋₄)₂,^{28,29} telomeric DNA^{28,34,35}) are less stable and more easily lost upon gentle collisional activation.

This can be explained as follows. In solution, one factor influencing the parallel/antiparallel topology of a telomeric quadruplex structure is the nature of the cation.³⁸ Potassium ions may favor parallel structures because they remain between the G-quartets, whereas sodium ions favor antiparallel structures because they can more readily move through the G-quartet stem. Ammonium ions show intermediate behaviour.³⁰ In ESI-MS, the probability of ammonia loss in the gas phase is dependent on the probability that a proton-accepting group from the oligonucleotide is located at the exit of the G-quadruplex stem. This can explain why ammonia loss is more likely in anti-parallel structures, because the loops, which are stacked onto 3' or 5' ends of the quadruplex stem can serve as proton acceptors.

Because the inner ammonium ions are exceptionally stable in all detected RNA G-quadruplexes, and because the structure of telomeric RNA has been reported to be parallel in both sodium and potassium solutions,¹²⁻¹⁸ we anticipated a CD signature of purely parallel-strand orientation for the RNA sequences with ammonium ions. The CD spectra of the annealed structures are shown in Figure 4, and the thermal difference spectra of the same samples are shown in supporting information Figure S5. The CD confirms that DNA12 is predominantly in an unstructured form, whereas DNA22 and DNA45 are forming mixed or hybrid, predominantly antiparallel G-quadruplexes (highest maximum at 295 nm).

All RNA sequences adopted predominantly parallel structures (highest maximum at 260 nm, minimum at 240 nm), but surprisingly, a significant maximum at 300 nm was also observed for RNA22 and RNA45, suggestive of a non-negligible fraction of base steps in alternating *syn-anti* conformations. This could be due to a partially antiparallel arrangement of the guanine steps, or to particular arrangements and/or stability of the UUA loops. The relative intensity of the CD band at 295 nm is higher than in all previously published CD spectra of annealed RNA22, or of closely related 21-nt or 23-nt sequences in KCl solutions.^{12,13,16,18} This motivated us to further study the influence of sample preparation procedure, namely the influence of annealing and the influence of reaction kinetics, on the structures adopted by telomeric RNA sequences.

Kinetic effects on RNA G-quadruplex folding and dimerisation

To obtain further insight into the mechanism of dimer formation by RNA22, we investigated the kinetics of dimer formation at a 10 μ M total strand concentration. The RNA was heated at 90 °C in water for 5 minutes to unfold any pre-formed structures and thermalized at room temperature (22 °C). G-quadruplex formation was initiated by the addition of ammonium acetate (100 mM final concentration). In an ESI-MS experiment, relative intensities of the different forms were measured as a function of reaction time, and the relative concentrations of monomer and dimer were recalculated at each time point (Figure 5a). The monomer was detected at the earliest time points (0.6 minutes) with two ammonium ions bound, indicating rapid intramolecular folding of RNA22, and suggesting that the dimeric structure then evolves from the pre-formed monomeric G-quadruplex. The dimer formation half-life is approximately 1 hour in pure water. Experiments in 10% and 20% methanol did not reveal dramatic differences in the self-assembly kinetics, in contrast with the effect of methanol on the self-assembly kinetics of the tetramolecular G-quadruplex $[d(TG_5T)]_4$.³⁹ The ammonium ion count in the dimer (Figure 5b) reveals another difference with the annealed structure. In the dimer formed by incubation at room temperature, the adducts with five ammonium ions and more are the only detectable species, while in the annealed dimer (Figure 5c), the ammonium adduct distribution shows also a four-ammonium ion adduct. These results suggest that the dimer that is initially formed at room temperature consists of a cation-mediated stacking of two monomers, and that the dimer formed by annealing has undergone a more complex rearrangement of the strands.

The ESI-MS results show that, by varying the incubation time, it is possible to study the structure of the monomer (< 10 min), or the stacked dimer (longer times at room temperature), and compare them to the annealed structure. Figure 5d shows the evolution of the CD spectra as a function of the incubation time at room temperature. The monomer which is folded after three minutes adopts an almost purely parallel structure. As the incubation time increases, the band indicative of parallel G-quadruplex arrangement (260 nm) further increases, and the smaller band indicative of antiparallel arrangement (300 nm) also grows at the same rate. However, in all cases, the monomers and dimers formed at room temperature are more parallel than those formed by annealing. This shows that contrary to the common belief that they form exclusively parallel structures, telomeric RNA sequences are also significantly polymorphic under these conditions. This also concurs with the NMR study of a 23-nt telomeric RNA in potassium, which showed broad and unresolved spectra in the imino proton region.¹⁸

The substantial effect of the kinetics of multimer formation is also important for studying the thermal stability of telomeric RNA G-quadruplexes by thermal denaturation experiments.^{13,16,21} Previous studies have shown that telomeric RNA G-quadruplexes are significantly more stable than their DNA counterpart in potassium,^{13,16,21} but not in sodium solution.^{16,21} Substantial hysteresis in the denaturation curves of RNA22 in potassium has previously been noted²¹ but the origin of this hysteresis was not elucidated. Here we performed thermal denaturation experiments

on the structures annealed in ammonium acetate solution, and we can use ESI-MS to interpret the results in terms of multimer formation.

Hysteresis was observed in ammonium solution for all three telomeric RNA sequences (Figure 6). No hysteresis is observed for their DNA counterparts (supporting information Figure S6). For RNA22, in the light of the kinetic study at room temperature, the hysteresis can be attributed to slow dimer formation. The bimodal cooling curve can be interpreted by assuming a fast monomer refolding, followed by slower dimerisation and rearrangement. The fact that the heating curves of the first cycle (curve 1) and the second cycle (curve 3) are not superimposed indicates that the structure formed by 0.2 °C/min cooling of the dilute solution (here 10 μ M) can be different from that formed by annealing at 300 μ M strand concentration, but it can also be due to partial strand degradation.

The hysteresis is even more pronounced for RNA12. The starting structure is tetrameric (Figure 1d), and curve 1 leading to the unfolding into single strands appears to be multimodal, indicating a complex dissociation mechanism. There is very large hysteresis between curves 1 and 2, and refolding only occurs at low temperature. Surprisingly, curve 3 superimposes with curve 2, indicating that the structure formed by the cooling cycle is very different from the starting structure. We used ESI-MS to follow the G-quadruplex formation from RNA12 when incubated in ammonium acetate at room temperature (supporting information Figure S7). At 5 μ M strand concentration, the dimer is formed in 20 minutes, but formation of the tetramer has not yet been initiated. We can therefore conclude that curve 1 corresponds to a multistep dissociation of the tetramer, and curves 2 and 3 correspond to the folding/unfolding of the dimeric G-quadruplex.

Interestingly, in the case of RNA45, although the sequence only folds intramolecularly, hysteresis is also observed. Notably, the denaturation is bimodal, while the renaturation is monomodal. Time-resolved ESI-MS analysis of the cation incorporation in RNA45 revealed that all five ammonium ions are incorporated in less than one minute (supporting information Figure S8). It can therefore be concluded that the refolding and cation-mediated stacking is fast, and therefore that the hysteresis arises from denaturation which is slower than the temperature gradient step. This emphasizes the extra stability provided by the cation-mediated stacking of the neighboring G-quadruplex subunits.

Conclusions

Telomeric RNA sequences form dimers of G-quadruplex subunits more readily than the equivalent DNA sequences under the conditions of the experiments reported here. A 5'-5' stacking of two G-quadruplex subunits was previously reported for the telomeric RNA 10-mer GGGUUAGGGU,¹⁸ and our observations suggest that the formation of higher-order assemblies by telomeric RNA is a more general phenomenon, occurring also in longer sequences. ESI-MS also shows that the dimers in the longer sequences are stabilized by cation-mediated stacking of two G-quadruplex subunits. We note that 5'-5' quadruplex-quadruplex stacking has been observed in the crystal structures of both an intramolecular parallel and a dimeric parallel human telomeric DNA quadruplex,⁴⁰ as well as in the crystal structure of a parallel dimer of the human telomeric RNA quadruplex sequence r(U_{Br}AGGGUUAGGGU),¹⁹ suggesting that such stacking is also a consequence of the crowded conditions in the crystalline state, and possibly also in a cellular environment.^{41,42} However, here stacking was found in dilute aqueous solutions, in the absence of crowding agents or of alcohols.

RNA G-quadruplexes are usually considered to form only parallel G-quadruplexes with all-propeller loops, because the C3'-*endo* pucker for the ribose sugars constrains the guanosine nucleotides to be in the *anti* conformation. Our ESI-MS and CD studies of annealed and RT-incubated structures of RNA22 show that the structures adopted are more polymorphic than initially assumed, multimer formation being one of the features of polymorphism. This explains why telomeric RNA analogues with four G-quartets give unresolved NMR spectra.¹⁸ We also suggest that, because of the formation of stable intramolecular dimers, telomeric RNA sequences containing at least eight guanine repeats could be better models for analyzing the structure of long telomeric RNA. The 45-mer should be the shortest sequence to be considered for mimicking the structural arrangement of long telomeric RNA. Future work on longer sequences is needed to determine how these 45-mer subunits interact with each other.

Our results also highlight the influence of the sample preparation procedure on the structures obtained. We therefore suggest that future reports on RNA G-quadruplexes should include detailed information on the concentration-temperature-time profiles during sample preparation. The influence on dimerisation of the bases which are flanking the G-quadruplex subunits also merits further investigation.

Acknowledgements

This work was supported by the European COST action MP0802, by Cancer Research UK (Programme Grant C129/A10902 to S.N.), by the School of Pharmacy (research studentship to G. C.), by the Fonds de la Recherche Scientifique-FNRS (CC.1.5.286.09.F to V.G., research associate position to V.G., postdoctoral fellowship to F.R.), and by the University of Liège (starting grant FRSD-08/10 to V.G.).

Supporting information available:

Figures S1 to S8. This information is available free of charge via the Internet at <http://pubs.acs.org/>

References

1. Wright, W. E.; Tesmer, V. M.; Huffman, K. E.; Levene, S. D.; Shay, J. W. *Genes Dev.* **1997**, *11*, 2801-2809.
2. Parkinson, G. N. In *Quadruplex Nucleic Acids*, Neidle, S., Balasubramanian, S., Eds.; The Royal Society of Chemistry: Cambridge, 2006; pp 1-30.
3. Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. *Nucleic Acids Res.* **2006**, *34*, 5402-5415.
4. Blackburn, E. H. *Cell* **2001**, *106*, 661-673.
5. Collins, K.; Mitchell, J. R. *Oncogene* **2002**, *21*, 564-579.
6. Kim, N. W.; Piatyszek, M. A.; Prowse, K. R.; Harley, C. B.; West, M. D.; Ho, P. L.; Coviello, G. M.; Wright, W. E.; Weinrich, S. L.; Shay, J. W. *Science* **1994**, *266*, 2011-2015.
7. Mergny, J.-L.; Riou, J.-F.; Mailliet, P.; Teulade-Fichou, M.-P.; Gilson, E. *Nucleic Acids Res.* **2002**, *30*, 839-865.
8. Hurley, L. H. *Nature Rev. Cancer* **2002**, *2*, 188-200.
9. Neidle, S.; Parkinson, G. *Nature Rev. Drug Discov.* **2002**, *1*, 383-393.
10. Azzalin, C. M.; Reichenbach, P.; Khoraiuli, L.; Giulotto, E.; Lingner, J. *Science* **2007**, *318*, 798-801.
11. Schoeftner, S.; Blasco, M. A. *Nat. Cell Biol.* **2008**, *10*, 228-236.
12. Qi, J.; Shafer, R. H. *Biochemistry* **2007**, *46*, 7599-7606.
13. Arora, A.; Maiti, S. *J. Phys. Chem. B* **2009**, *113*, 10515-10520.
14. Xu, Y.; Kaminaga, K.; Komiyama, M. *Nucleic Acids Symp. Ser. (Oxf)* **2008**, 175-176.
15. Xu, Y.; Kaminaga, K.; Komiyama, M. *J. Am. Chem. Soc.* **2008**, *130*, 11179-11184.
16. Joachimi, A.; Benz, A.; Hartig, J. S. *Bioorg. Med. Chem.* **2009**, *17*, 6811-6815.
17. Randall, A.; Griffith, J. D. *J. Biol. Chem.* **2009**, *284*, 13980-13986.
18. Martadinata, H.; Phan, A. T. *J. Am. Chem. Soc.* **2009**, *131*, 2570-2578.
19. Collie, G. W.; Haider, S. M.; Neidle, S.; Parkinson, G. N. *Nucleic Acids Res.* **2010**, doi: 10.1093/nar/gkq259.
20. Phan, A. T. *FEBS J.* **2010**, *277*, 1107-1117.
21. Sacca, B.; Lacroix, L.; Mergny, J.-L. *Nucleic Acids Res.* **2005**, *33*, 1182-1192.
22. Luke, B.; Lingner, J. *EMBO J.* **2009**, *28*, 2503-2510.
23. Schoeftner, S.; Blasco, M. A. *EMBO J.* **2009**, *28*, 2323-2336.
24. Goodlett, D. R.; Camp, D. G., II; Hardin, C. C.; Corregan, M.; Smith, R. D. *Biol. Mass Spectrom.* **1993**, *22*, 181-183.
25. Hofstadler, S. A.; Griffey, R. H. *Chem. Rev.* **2001**, *101*, 377-390.

26. Beck, J.; Colgrave, M. L.; Ralph, S. F.; Sheil, M. M. *Mass Spectrom. Rev.* **2001**, *20*, 61-87.
27. David, W. M.; Brodbelt, J.; Kerwin, S. M.; Thomas, P. W. *Anal. Chem.* **2002**, *74*, 2029-2033.
28. Rosu, F.; Gabelica, V.; Houssier, C.; Colson, P.; De Pauw, E. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1729-1736.
29. Guo, X.; Liu, S.; Yu, Z. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 1467-1476.
30. Smargiasso, N.; Rosu, F.; Hsia, W.; Colson, P.; Baker, E. S.; Bowers, M. T.; De Pauw, E.; Gabelica, V. *J. Am. Chem. Soc.* **2008**, *130*, 10208-10216.
31. Rosu, F.; De Pauw, E.; Gabelica, V. *Biochimie* **2008**, *90*, 1074-1087.
32. Gros, J.; Rosu, F.; Amrane, S.; De Cian, A.; Gabelica, V.; Lacroix, L.; Mergny, J. L. *Nucleic Acids Res.* **2007**, *35*, 3064-3075.
33. Gabelica, V.; Rosu, F.; De Pauw, E. *Anal. Chem.* **2009**, *81*, 6708-6715.
34. Gabelica, V.; Baker, E. S.; Teulade-Fichou, M.-P.; De Pauw, E.; Bowers, M. T. *J. Am. Chem. Soc.* **2007**, *129*, 895-904.
35. Baker, E. S.; Bernstein, S. L.; Gabelica, V.; De Pauw, E.; Bowers, M. T. *Int. J. Mass Spectrom.* **2006**, *253*, 225-237.
36. Gabelica, V.; De Pauw, E. *Mass Spectrom. Rev.* **2005**, *24*, 566-587.
37. Gabelica, V.; Schulz, E.; Karas, M. *J. Mass Spectrom.* **2004**, *39*, 579-593.
38. Lane, A. N.; Chaires, J. B.; Gray, R. D.; Trent, J. O. *Nucleic Acids Res.* **2008**, *36*, 5482-5515.
39. Rosu, F.; Gabelica, V.; Poncelet, H.; De, P. E. *Nucleic Acids Res.* **2010**, doi:10.1093/nar/gkq208.
40. Parkinson, G. N.; Lee, M. P. H.; Neidle, S. *Nature* **2002**, *417*, 876-880.
41. Renciuik, D.; Kejnovska, I.; Skolakova, P.; Bednarova, K.; Motlova, J.; Vorlickova, M. *Nucleic Acids Res.* **2009**, *37*, 6625-6634.
42. Xue, Y.; Kan, Z. Y.; Wang, Q.; Yao, Y.; Liu, J.; Hao, Y. H.; Tan, Z. *J. Am. Chem. Soc.* **2007**, *129*, 11185-11191.

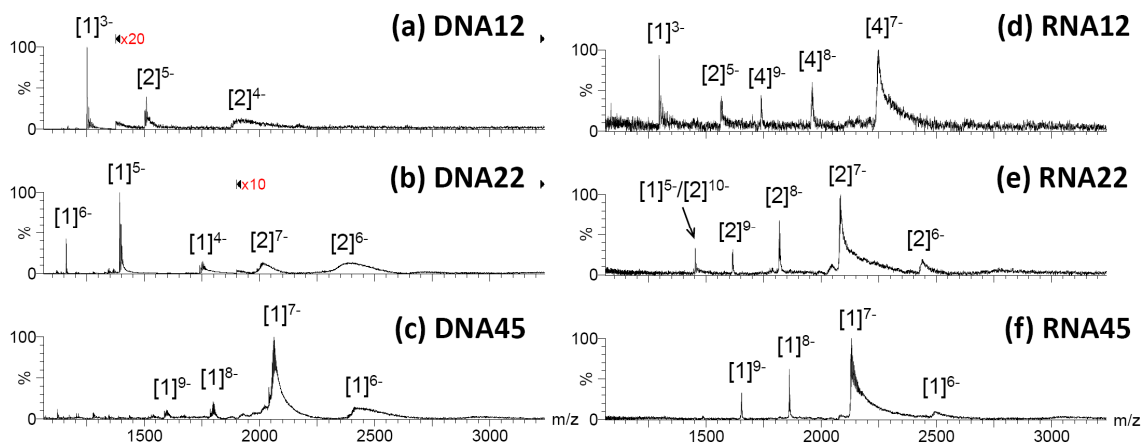


Figure 1: Electrospray mass spectra of annealed (a) DNA12, (b) DNA22, (c) DNA45, (d) RNA12, (e) RNA22, and (f) RNA45. Final strand concentrations are 20 μM for the 12-mers, 10 μM for the 22-mers, and 5 μM for the 45-mers, in 100 mM NH_4OAc . 20 % MeOH was added just before injection. The RF Lens 1 voltage is 100 V for the DNAs (a-c), 120 V for RNA12 and RNA22 (d-e), and 140 V for RNA45 (f). The peak annotations $[n]^z$ indicate the number of strands (n) and the total charge (z). Note the 20-fold magnification of the DNA12 dimer in panel (a) and the 10-fold magnification of the DNA22 dimer in panel (b).

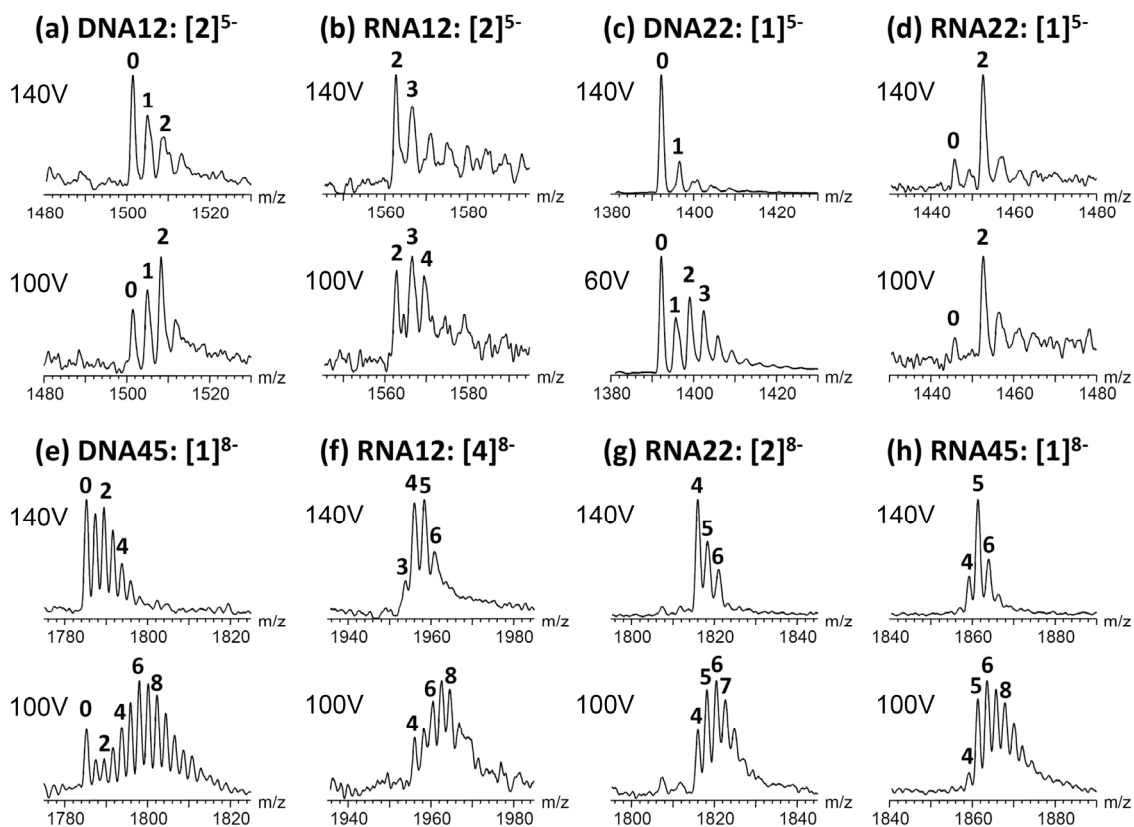


Figure 2: Distribution of the number of ammonium ions preserved in the structure of (a) dimeric DNA12, (b) dimeric RNA12, (c) intramolecular DNA22, (d) intramolecular RNA22, (e) intramolecular DNA45, (f) tetrameric RNA12, and (g) dimeric RNA22, and (h) intramolecular RNA45. The distribution is shown at two RF lens voltages for each species, and the number of ammonium ions are indicated on each peak.

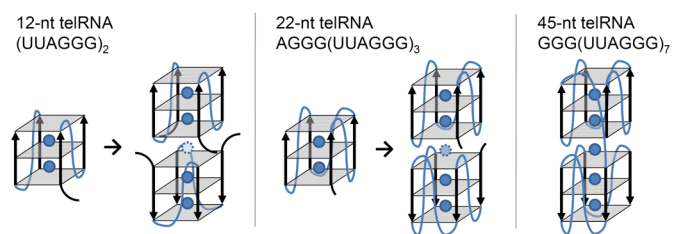


Figure 3: Self-assembly of telomeric RNA as deduced from the number of strands and of bound cations in each assembly, as detected by electrospray mass spectrometry. Darker circles indicate cations remaining tightly bound in the gas phase, and lighter circles indicate more labile cations.

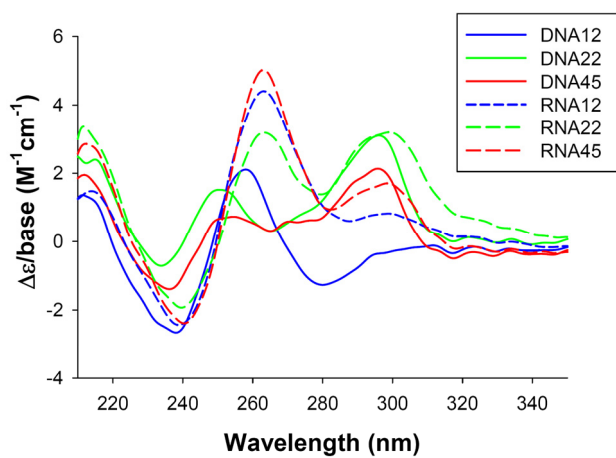


Figure 4: Circular dichroism spectra of all annealed sequences in 100 mM ammonium acetate. The spectra were recorded on a Jasco J810 spectropolarimeter, with 2 mm pathlength cuvettes filled with 20 μM monomer concentrations of DNA12 or RNA12, 10 μM monomer concentrations of DNA22 or RNA22, and 5 μM monomer concentration of DNA45 or RNA45. DNA12 is not predominantly in a G-quadruplex form, DNA22 and DNA45 are assigned as predominantly antiparallel quadruplexes, and all RNA sequences are assigned as predominantly parallel quadruplexes.

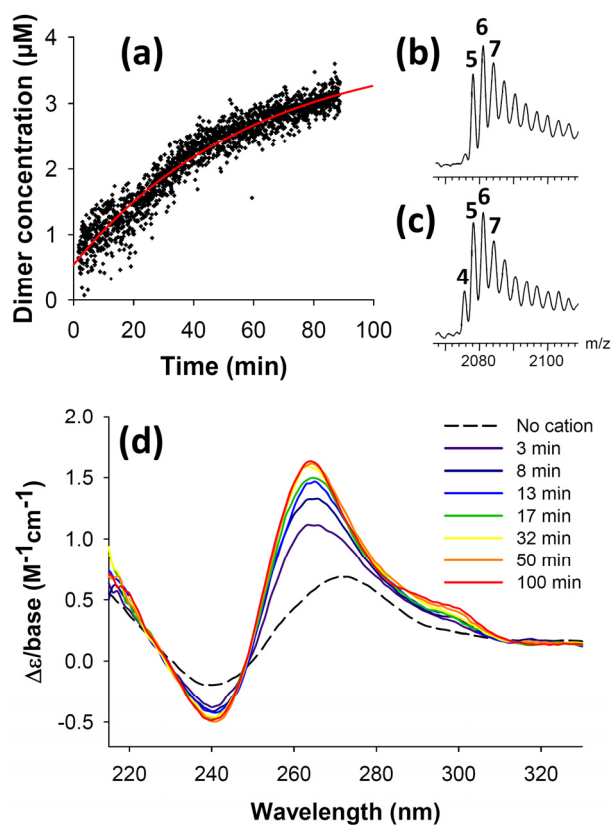


Figure 5: Kinetics of RNA22 folding and dimerisation. (a) Time evolution of the dimer concentration followed by ESI-MS (140 V on the RF lens 1, total strand concentration = 10 μM , no methanol). Time $t = 0$ corresponds to ammonium acetate addition to the RNA solution. The red line is a fit with first-order kinetics (dimer formation rate of 0.017 min^{-1}). (b) Ammonium ion distribution on the M_2^{7-} ion of the dimer, obtained after 50 min incubation at room temperature. (c) Ammonium ion distribution on the M_2^{7-} ion of the annealed dimer, with the identical source conditions as in panel b. (d) Time evolution of the CD spectrum of RNA22 after cation addition to 10 μM unfolded monomer.

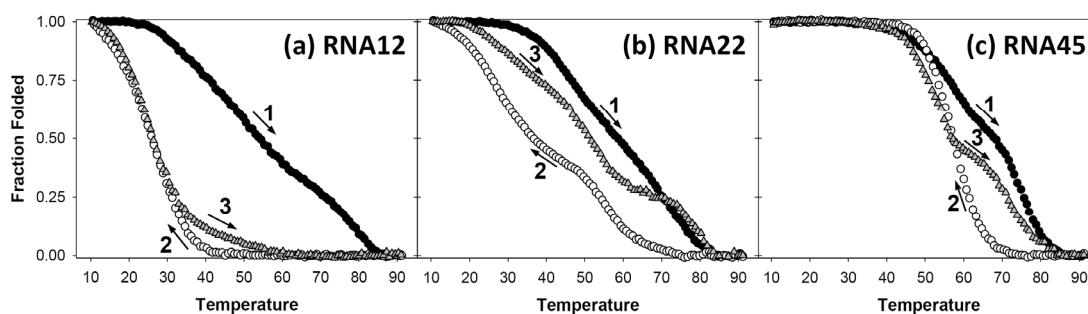


Figure 6: Thermal denaturation experiments on the annealed RNA G-quadruplexes, monitored by UV absorbance at 295 nm. The absorbance data are transformed into fraction folded. The final monomer concentrations were 20 μM for RNA12 (a), 10 μM for RNA22 (b), and 5 μM for RNA45 (c). The sample were heated at 0.2 $^{\circ}\text{C}/\text{min}$ from 17 $^{\circ}\text{C}$ to 92 $^{\circ}\text{C}$ (curve 1, black circles), then cooled down to 10 $^{\circ}\text{C}$ at the same rate (curve 2, white circles), then heated again to 92 $^{\circ}\text{C}$ (curve 3, grey triangles).