Published online 25 February 2010 *Nucleic Acids Research, 2010, Vol. 38, No. 11* ***3817–3826***
doi:10.1093/nar/gkq102

$[C_7GC_4]_4$ Association into supra molecular i-motif structures

Aude Laisné[1], Denis Pompon[1] and Jean-Louis Leroy[2,*]

[1]Laboratoire d'Ingénerie des Proteines Membranaires, Centre de Génétique Moléculaire, CNRS FRE3144, Unité associé à l'Université Pierre et Marie Curie and [2]Laboratoire de Chimie et Biologie Structurale, Institut de Chimie des Substances Naturelles, Gif-sur-Yvette, 91190, France

Received December 28, 2009; Revised January 29, 2010; Accepted February 3, 2010

ABSTRACT

The self-associative properties of cytidine-rich oligonucleotides into symmetrical i-motif tetramers give to these oligonucleotides the capacity of forming supramolecular structures (sms) that have potential applications in the nanotechnology domain. In order to facilitate sms formation, oligonucleotides containing two cytidine stretches of unequal length (C_nXC_m) separated by a non-cytidine spacer were synthesized. They were designed to associate into a tetramer including an i-motif core built by intercalation of the $C{\cdot}C^+$ pairs of the longer C stretch with the two dangling non-intercalated strands of the shorter C stretch at each end. Gel filtration chromatography shows that the non-intercalated C-rich ends give to this structure the capacity of forming extremely stable sms. Using C_7GC_4 as a model, we find that the sms formation rate varies as the oligonucleotide concentration and increases at high temperature. Competitively with the tetramer involved in sms elongation, C_nXC_m oligonucleotides form i-motif dimers that compete with sms elongation. The dimer stability is strongly reduced when the pH is moved away from the cytidine pK. This results in an equilibrium shift towards the tetramer and in the acceleration of the sms formation rate. The chromatograms of the sms formed by C_7GC_4 indicate a broad distribution. In a 1.5 mM solution incubated at 37°C, the equilibrium distribution is centered on a molecular weight corresponding to the assembly of nine tetramers and the upper limit corresponds to 80 tetramers. The lifetime of this structure is about 4 days at 40°C, pH 4.6.

INTRODUCTION

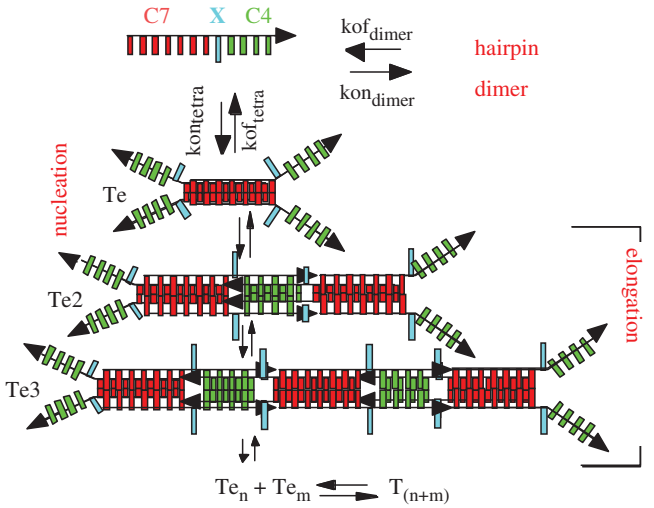
At slightly acid pH, the hemiprotonated $C{\bullet}C^+$ pairs formed by neutral and protonated cytidines induce the association of cytidine-rich oligonucleotides into parallel duplexes. These duplexes are short-lived elusive species, but the tetramers built by head to tail intercalation of two duplexes are extremely stable. This structural arrangement, called i-motif is exceptional in that it involves systematic base-pair intercalation (1). The i-motif structure may include four identical C-rich strands, two hairpins each caring two cytidine stretches or a folded strand caring four cytidine stretches (2,3). The remarkable symmetry of i-motif tetramers is a consequence of the equivalence of the four strands associated in the structure. Face to face intercalation of thymidine in i-motif structures is generally unfavorable (4,5) and purine intercalation seems to be sterically hindered (6,7). A recent kinetics investigation shows that i-motif tetramers with partial or full intercalation topology are formed at comparable rates and that the predominance at equilibrium of the fully intercalated species is due to its longer lifetime (8).

The self-associative properties of C-rich stretches give to i-motif tetramers the potentiality of forming supramolecular structures (sms) by mutual intercalation of the non-paired cytidines of tetramers built by partially intercalated duplexes. This property has been reported in a recent study revealing the association of C7 oligonucleotides into linear repetitive structures that could have biotechnological applications (9).

In order to understand and control the formation process of i-motif sms, we examine in the present paper how temperature, strand concentration, ionic strength and pH influence the sms formation and dissociation rates.

Oligonucleotides of the C_nXC_m series were designed to improve the assembly of i-motif tetramers into sms. They are composed of two cytidine stretches of unequal length separated by a non-cytidine residue i.e. by A, T, G or

*To whom correspondence should be addressed. Tel: +33 1 69 82 3630; Fax: +33 1 69 82 3784; Email: jean-louis.leroy@icsn.cnrs-gif.fr



Scheme 1. Postulated association pathway of C_7GC_4 into i-motif sms. The monomer is in equilibrium with one or several hairpin dimer(s) (cf. Scheme 2) and Te, the tetramer formed by full intercalation of the C_7 stretches. Association and mutual intercalation of the C4 stretches of two Te building blocks result in the formation of the Te_2 species. The i-motif symmetry gives to the assembly of several building blocks the same elongation capacity than that of the building blocks themselves and allows the association of preformed Te_n and Te_m species into structures including (n + m) building blocks.

a diol-spacer. They were synthesized with the longer C stretch either at the 5′- or the 3′-end. The non-cytidine spacer that cannot intercalate into the i-motif was intended to prevent intercalation of both cytidine stretches into a single i-motif core. It was expected that the spacer would enforce the formation of a tetramer (Te in Scheme 1) including the i-motif core built by the $C{\bullet}C^+$ pairs of the longer cytidine stretch and at each end the non-intercalated cytidines of the shorter stretch.

Gel filtration chromatography shows that these oligonucleotides associate into sms. The stability, the formation and dissociation rates of the sms formed by oligonucleotides differing by the C stretch length and the nature of the spacer were qualitatively compared. The relatively fast formation rate and the good stability of the sms formed by C_7GC_4 prompted us to select this sequence as a benchmark model to investigate the effect of pH, strand concentration, temperature and ionic strength on the sms formation and dissociation kinetics. The influence of the C stretch length and of the nature of non-cytidine spacer on the sms stability, formation rate and structure will be examined in separate publications.

MATERIALS AND METHODS

Oligonucleotide synthesis and sample preparation

The oligonucleotides were synthesized on a 2 or 10 μM scale, purified by chromatography on an anionic DEAE column according to procedures already described and extensively dialyzed (5). After dialysis, the oligonucleotide solutions were lyophilized and dissolved in water to make 0.5–1 mM stock solutions. The C_7GC_4 concentration was determined using the A^{260} values, 89600 M^{-1} cm^{-1}, computed according to a nearest neighbor model (10).

Formation and dissociation rate measurements of i-motif sms

Except otherwise stated, all the measurements were performed in 0.4 M NaCl, 10 mM Na acetate and 10 mM Na phosphate, a solution designed as the NAP buffer. During sms formation and dissociation, the samples, typically 50–200 μl, were incubated in a mastercycler Eppendhorf® PCR incubator whose lid temperature was set 5°C above that of the samples in order to avoid condensation on the cap of the Eppendhorf tubes.

To study the sms formation kinetics, the samples were initially melted at 100°C and rapidly cooled at the temperature of the incubator. For sms dissociation time measurements, a solution containing only sms was prepared by pooling the species eluted on a GPC 100 column with a molecular weight larger than that of a tetramer. Afterward, this solution was diluted in the NAP buffer in such way as to allow complete dissociation of the sms at equilibrium. The measurements started right after dilution.

Gel filtration chromatography

During sms formation and dissociation, the sample composition was measured at room temperature by gel filtration chromatography (11) on GPC 100 and GPC 1000 columns (250 × 4.6 mm I.D.) provided by Eprogen®.

Aliquots took off the incubated samples were injected in the column immediately after dilution to 25 μl (the volume of the injection loop) in the NAP buffer. They were eluted with the same buffer at a flow rate of 0.4 ml min^{-1}. Systematic addition of thymidine at a final concentration of about 5 μM in the injected sample provided a reference marker on the chromatograms. The elution times ranged from 4 to 8.5 min. (Supplementary Figure S1). The elution buffer pH was generally 4.6. However, when the evolution of the sample composition during elution on the column at pH 4.6 was not negligible, the elution buffer pH was adjusted to the value of the sample pH. The GPC-100 column whose exclusion limit correspond, according to the manufacturer to an oligonucleotide containing about 300 nt was calibrated using tRNA, thymidine, C-rich non-structured oligonucleotides, C-rich i-motif monomers and tetramers. The GPC-1000 column was calibrated with double-stranded plasmids containing from 800 to 16000 nucleosides, the i-motif tetramer of $T_2C_8T_2$, tRNA and thymidine. The GPC-1000 calibration shows that oligonucleotides containing 50 and 10^4 residues are eluted, respectively, in the permeation and exclusion limits. The calibration curves displayed in Supplementary Figure S1 establishes the reasonable linearity of the measured elution times with the log of the molecular weight of reference oligonucleotides. The dispersion of the data points of the calibration curves suggests that the relative error on the estimated molecular weights is <0.5. The sizes of the sms were characterized by the mass center, M_c, of the sms distribution. M_c was computed according to $M_c = \Sigma A(t)^* N(t)/ \Sigma A(t)$,

where A(t) is the absorbance of the solution and N(t) the number of tetrameric units in the sms eluted on the chromatogram at time t. N(t) was obtained from the calibration curve of the GPC 1000 column (Supplementary Figure S1). The dimer and tetramer stoichiometries were also routinely measured by the slope of the log–log plot of the multimer versus monomer equilibrium concentrations. In pH and temperature conditions favorable to sms formation, an opalescence characteristic of the formation of extremely large structures was observed in concentrated samples (>1 mM) incubated during times much longer than the sms formation time constant. Correlatively with the development of the sample opalescence, the chromatogram of these samples showed a reduction (up to 50%) of the integrated area of the eluted species indicating that a fraction of the sample composed, presumably of the largest sms, was trapped in the column. In that case, the oligonucleotide fraction retained on the column was estimated by comparison with the chromatogram of an identical melted sample and included in the sms fraction. However, in order to have a direct control of all the species formed, we generally restrained the oligonucleotide concentrations and incubation times to a range of values allowing detection of all the species on the chromatograms.

NMR method

The NMR experiments were performed using a 500 MHz Varian Inova spectrometer with the jump and return sequence for water suppression (12). The spectra were accumulated with a repetition time longer than the relaxation times to avoid saturation. The spectral intensity was multiplied by a 1/sin function in order to correct the distortion introduced by the jump and return excitation (13). The spectra collected during the sms formation were normalized using the methyl proton peak of the acetate buffer.

Multimer dissociation constants

The dissociation constant of a multimer may be expressed as a function of α_{eq}, the monomer equilibrium fraction of an oligonucleotide solution at concentration $[M_0]$ by: $K_{dis} = S\,\alpha_{eq}^{s}[M_0]^{s-1}/(1-\alpha_{eq})$ where s is the multimer stoichiometry. We will characterize the multimers stability by Fi, a parameter whose dimension is independent of the stoichiometry which is equal to the free monomer concentration for which $\alpha_{eq} = 0.5$. Fi is related to the multimer dissociation constant by: $Fi = (K_{dis}/s)^{(1/(s-1))}$ and it will be designated as the reduced dissociation constant.

RESULTS

Chromatographic evidences for the formation of i-motif supramolecular structures

Immediately after melting and fast cooling, a 3 μM C_7GC_4 sample at pH 6.2 injected into a GPC 100 column is mainly eluted as a monomer (Figure 1). The chromatograms recorded at different times during incubation at

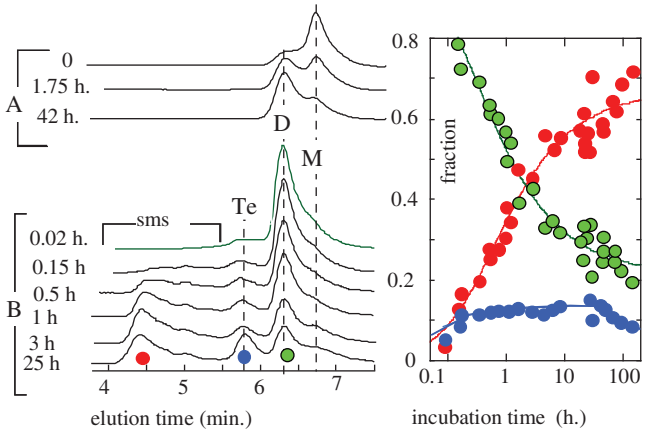


Figure 1. Association of C_7GC_4 into i-motif multimers and supra molecular structures (sms) in the NAP buffer. The samples were initially melted. *Left panel*: GPC-100 chromatograms recorded after incubation during the times indicated. The elution times expected for a monomer (M), a dimer (D) and a tetramer (Te) are indicated. (**A**) the chromatogram of a 3 μM solution, pH 6.2 injected right after melting shows a 20/80 dimer/monomer mixture. The chromatograms recorded after incubation at 20°C show that the dimer is formed with a time constant of about 1.5 h. The monomer and dimer concentrations at equilibrium correspond to the reduced dissociation constant $Fi_{dimer} = 2\,\mu M$. (**B**) The chromatograms of a 0.15 mM C_7GC_4 solution at pH 4.63 collected after melting and incubation at 50°C during the times indicated on the figure show that the dimer is slowly converted into tetramer and supramolecular structures. *Right panel*: evolution of the sms (red), tetramer (blue) and dimer (green) fractions as a function of the incubation time at 50°C.

20°C reveal that a dimer is formed with a time constant of about 1.5 h. At equilibrium, the dimer and monomer concentrations correspond to a reduced dissociation constant Fi = 2 μM. It has been shown that the formation rate of i-motif dimers increases with the nucleotide concentration and when the solution pH is shifted near 4.4, the value of the cytidine pK_{N3} (8). This explains that about 1 min after melting, an aliquot of a 0.15 mM C_7GC_4 solution, pH 4.6 is eluted as a dimer (Figure 1). The chromatograms of samples incubated at 50°C show the appearance of two peaks, one corresponding to a tetramer and the other to unresolved heavier species that will be designed as sms.

The evolution versus time of the dimer, tetramer and sms fractions is displayed in Figure 1. After 0.5 h, the tetramer fraction reaches a maximum value corresponding to about 12% of the oligonucleotide concentration. The sms increase with a time constant of 1 h toward an equilibrium level corresponding to 65% of the oligonucleotide concentration. The dimer fraction decreases with a comparable time constant. The sms molecular weights were estimated using a GPC-1000 column. The chromatograms of aliquots taken from a 1.5 mM C_7GC_4 solution incubated at 37°C, pH 4.6 are displayed in Figure 2A. The dimer and the tetramer are eluted together in the permeation volume of the GPC-1000 column. The sms half formation time is 1 h and the oligonucleotide fraction associated in sms at equilibrium is about 75% (Figure 2D). The average size of the sms increases as a function of the incubation time (Figure 2B). After

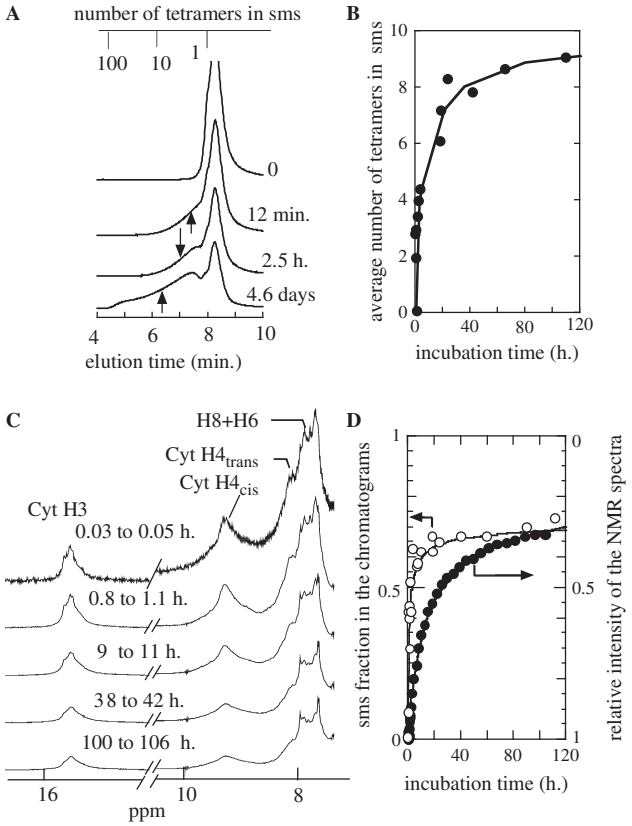
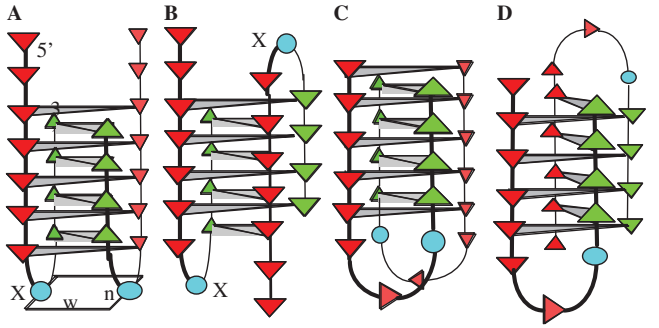


Figure 2. Association of C_7GC_4 into sms as monitored by NMR and gel exclusion chromatography on GPC 1000 column. The oligonucleotide concentration was 1.5 mM. The NMR spectra and the chromatograms were recorded after melting and incubation in the NAP buffer, pH = 4.6 at 37°C during the times indicated. (**A**) The tetramer and the dimer are eluted together in the GPC 1000 chromatograms. The chromatograms collected as a function of the time show the formation of sms. The upper scale shows the elution times expected for sms including the indicated number of tetrameric repeats. Note that one tetramer contains 48 nucleotides. On each chromatogram, an arrow indicates the mass center of the sms distribution. (**B**) the number of tetramer repeats at the center of the sms distributions is plotted as a function of the incubation time. At equilibrium, the mass center of the sms distribution corresponds to the association of nine tetramers (i.e. 432 bases). The top 15% of the sms distribution includes structures that are five times larger. (**C**) exchangeable and aromatic proton regions of NMR spectra collected during sms formation. The line broadening (about 40 Hz) and the reduction of the signal intensity observed as a function of the time reflect the formation of large structures. (**D**) oligonucleotide fractions included in sms as determined from the chromatograms of panel A (open circles) and from the intensity of the NMR peaks of the spectra of panel C (black circles). The time constant derived from the chromatography measurements, 1 hour, concerns the formation of structures heavier than a tetramer, whereas that derived from the NMR experiments, 20 h, is related to structures too large to be detected on the NMR spectra.

incubation over 4 days at 37°C, the sms exhibit an extremely broad distribution centered around 430 residues, a size corresponding to the association of nine tetramers. The elution time of the largest species, 4.5 min, is that expected for a structure including about 4000 residues, i.e. 80 tetrameric units (Figure 2A).

sms formation monitored by NMR

Right after melting and fast cooling at 0°C, the NMR spectrum of C_7GC_4 shows three broad clusters of



Scheme 2. Plausible structures of the C_7XC_4 dimer (X = A, T or G). The number of intercalated C•C$^+$ pairs in each structure is optimal. The narrow (n) and wide (w) grooves are indicated in structure (**A**). One residue is enough to make a loop across the i-motif narrow groove (6) and three residues allow loop formation across the wide groove (5). For each loop topology, the hairpins may be in parallel (**A**, **C**) or in head to tail orientation (**B, D**). Many other topologies can be imagined by changing the intercalation order of the C•C+ pairs and the number of residues in the loops. The poor resolution of the NMR spectrum of C_7GC_4 suggests that several dimers coexist.

exchangeable protons at the positions characteristic of the imino and amino protons of C•C$^+$ pairs (14). The poor spectral resolution suggests the presence of multiple conformations (Supplementary Figure S2). The 10.3 ppm chemical shift and the fast exchange rate, about 40 ms, of the G imino proton establish that the guanine is not H-bonded to the cytidine N3. The number of cytidine imino protons around 15.5 ppm estimated by reference to the guanosine imino proton peak is consistent with the number of C•C$^+$ pairs (seven to nine C•C$^+$ pairs for two guanosine imino protons) expected for the i-motif dimers displayed in Scheme 2.

Sms formation was followed by NMR in the same experimental conditions ([C_7GC_4] = 1.5 mM, pH 4.6, T = 37°C) than the gel filtration experiments described just above. At 37°C, the guanosine imino proton peak is broadened out by exchange with water (Figure 2C). The NMR spectra recorded as a function of the time after melting show a broadening of about 40 Hz and a reduction of the intensity of all the NMR peaks indicating the formation of large structures with a slow tumbling rate. The initial intensity is fully restored after heating at 100°C. At equilibrium, the reduction of the spectral intensity (~75%) is comparable to the oligonucleotide fraction found associated in sms by chromatography, but the intensity of the NMR peaks decreases with a time constant 20 times longer than that measured by chromatography (Figure 2D). The discrepancy of the time constants measured by chromatography and NMR is merely related to the difference of the species that are taken into account by each method. The time constants derived from the chromatographic investigation are related, according to the definition given to sms, to the formation of structures whose molecular weights are larger than that of a tetramer whereas the time constants derived from the NMR experiments correspond to high molecular weight structures whose proton spectra are broadened out.

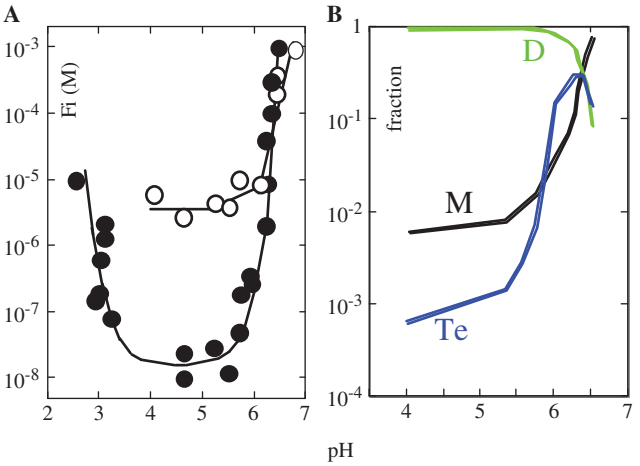


Figure 3. Effect of pH on dimer and tetramer stability at 20°C. (**A**) Reduced dissociation constants versus pH of the dimer(s) of C_7GC_4 (black circles) and of $[C_7T]_4$ (open circles). C_7T that do not associate into dimer or sms was used as a model to evaluate the stability of the tetramer of C_7GC_4 (Te in Scheme 1), the structure assumed to be the building block of the sms elongation pathway. (**B**) Computed monomer (black), dimer (green) and tetramer (blue) fractions in a 0.3 mM C_7GC_4 solution assuming that the tetramer of $[C_7GC_4]_4$ has the same reduced dissociation constant than $[C_7T]_4$. Due to the greater sensitivity of the dimer stability to pH, the tetramer fraction increases by two magnitude orders between pH 5.5 and pH 6.

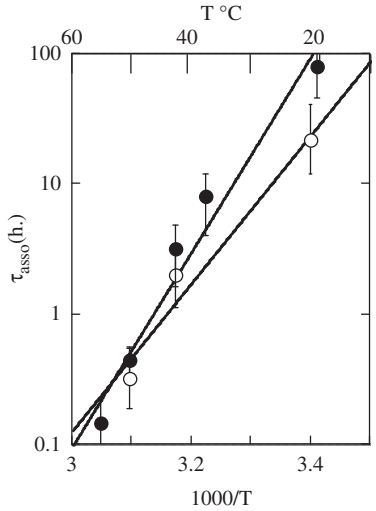


Figure 4. Sms half formation times vs. temperature of C_7GC_4 (black circles) and C_4GC_7 (open circles) 0.3 mM solutions in the NAP buffer at pH 4.6. The activation energies related to sms formation, 143 ± 40 and 110 ± 25 kJ/M are comparable.

Competition between dimer and tetramer formation

The concomitance of the evolution of the dimer and sms fractions (Figure 1) suggests that sms formation is controlled by dimer dissociation. In order to understand how the dimer and the tetramer of C_7GC_4 interfere with sms formation, we measured their reduced dissociation constant at different pH. The chromatograms of C_7GC_4 solutions recorded as a function of the time after melting and incubation at 20°C show the following systematic features: the first multimer formed is always a dimer. When the oligonucleotide concentration is sufficiently high, one observes that a tetramer is formed at a slower rate. Tetramer formation is always followed by the appearance of sms. The dimer reduced dissociation constant, Fi_{dimer}, measured from the dimer and monomer equilibrium concentrations is plotted *vs.* pH in Figure 3. Fi_{dimer} is close to 2 10^{-8} M between pH 3.5 and 5.5 and increases sharply when the solution pH is moved away from this range of values. The C_7GC_4 monomer concentration, which is controlled by the monomer–dimer equilibrium, is extremely small in the range of oligonucleotide concentrations allowing tetramer formation and for this reason the tetramer–reduced dissociation constant is experimentally inaccessible. Assuming that the tetramer of C_7GC_4 is formed by pairing and intercalation of the seven consecutive cytidines, we considered that its reduced dissociation constant should be comparable to that of any i-motif tetramer formed by intercalation of C_7 stretches and we used C_7T as a model to evaluate its reduced dissociation constant. The chromatograms of C_7T show two peaks at the positions expected for a monomer and a tetramer. The reduced dissociation constant, Fi_{tetra}, of $[C_7T]_4$ derived from the tetramer and monomer equilibrium concentrations is displayed versus pH in Figure 3. Between pH 4 and 5.5, the tetramer reduced dissociation constant, Fi_{tetra}, is about 200 times larger than that of the C_7GC_4 dimer. Due to the lower sensitivity of Fi_{tetra} to pH, the dimer and tetramer reduced dissociation constants are comparable around pH 6.5.

Effect of temperature on the sms formation time

We measured by chromatography on GPC 100 column the sms formation times in 0.3 mM C_7GC_4 solutions, pH 4.6 versus temperature. Figure 4 shows that the sms half formation time varies from 0.15 h at 55°C to about 90 h. at 20°C with an activation energy of 143 ± 30 kJ/M.

Effect of the oligonucleotide concentration on sms formation

Figure 5 shows that the sms half formation time decreases as the inverse of the oligonucleotide concentration. The oligonucleotide fraction associated at equilibrium in sms increases from 13% in a 28 μM C_7GC_4 solution to about 80% in 0.6 mM solution. The oligonucleotide fraction in tetrameric structures, about 10%, depends weakly on the oligonucleotide concentration. The monomer concentration is always negligible in the concentration range explored.

Effect of pH on the sms formation time

I-motif formation involves association of neutral and protonated cytidines. For this reason, the i-motif half formation time is minimal when the product, $[f_C]\,[f_{C+}]$, of the neutral and protonated cytidine fractions is maximal, i.e when the pH is equal to the cytidine pK_{N3} (8). On the contrary, the sms formation time is maximal when the pH is close to the cytidine pK_{N3} (Figure 6). It is quite puzzling to note that the effect of pH on the sms formation kinetics (shorter sms formation time when

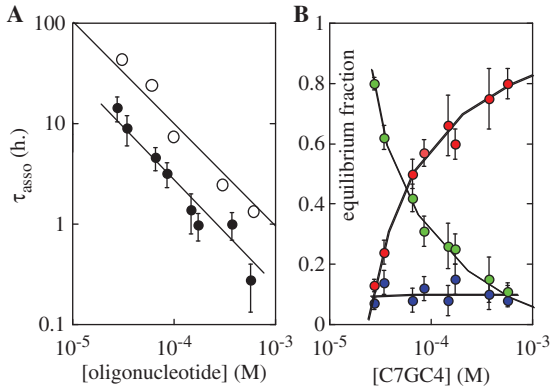


Figure 5. Effect of the oligonucleotide concentration on the sms half formation time at 42°C, pH 4.6, NAP buffer. (**A**) Half formation time of the sms of C_7GC_4 (black circles) and C_4TC_7 (open circles) versus the oligonucleotide concentration. The slope of the lines drawn through the data points is -1. (**B**) Dimer (green), tetramer (blue) and sms (red) equilibrium fractions versus the oligonucleotide concentration. The monomer fraction is negligible in this range of concentrations.

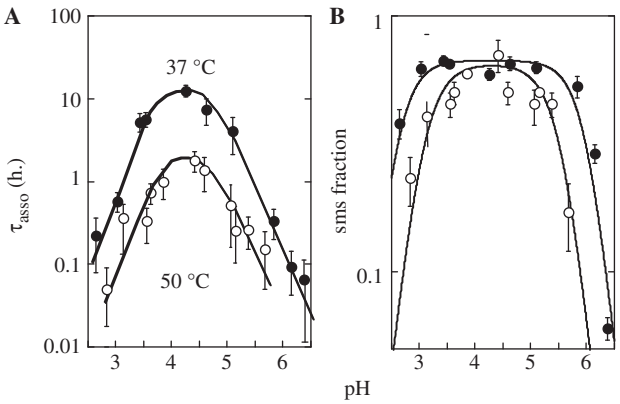


Figure 6. Effect of pH on sms formation in 0.3 mM C_7GC_4 solution at 37 (black circles) and 50°C (open circles) in the NAP buffer. (**A**) sms half formation time versus pH. The sms formation time is maximal when the pH is equal to the cytidine pK_{N3}. (**B**) C_7GC_4 fraction incorporated at equilibrium into sms at 37 and 50°C versus pH.

the pH is shifted away from 4.4) is opposite to that observed for the i-motif formation kinetics (longer i-motif formation time when the pH is shifted away from 4.4). At equilibrium, the oligonucleotide fraction incorporated in sms is maximal around pH = 4.4. The equilibrium sms fraction at pH 4.4 depends weakly on the temperature, as shown by the similarity of the sms proportions measured at 37 and 50°C.

Effect of the NaCl concentration on sms formation

Sms formation was examined in 0.3 mM C7GC4 solutions at 42°C, pH 4.6 as a function of the NaCl concentration in a buffer containing 10 mM sodium phosphate and 10 mM sodium acetate. The sms half formation time, 100 ± 25 h in the absence of added NaCl decreases rapidly to 3 ± 0.5 h in 0.4 M NaCl. It decreases more slowly at higher NaCl concentration and reaches about 1 h in 2 M NaCl. The C7GC4 fraction associated in sms structures

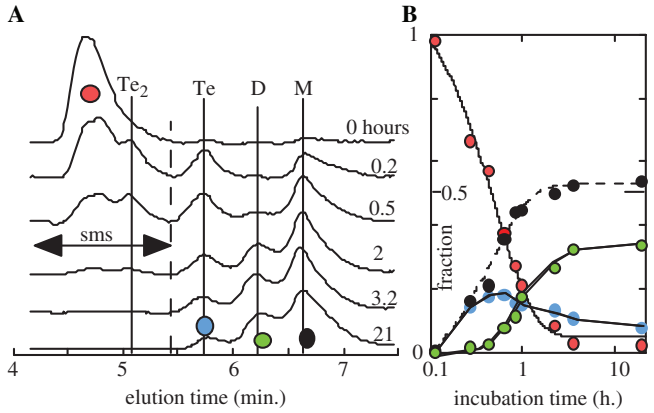


Figure 7. Dissociation of the sms of C_7GC_4 at 40°C, pH 4.63 in the NAP buffer. (**A**) The oligonucleotide concentration is 3 μM. The chromatograms recorded as a function of the time show that sms dissociation first results in the formation of monomer (M), tetramer (Te) and Te_2, a species whose elution time corresponds to that expected for the assembly of two tetramers. At last the chromatograms show the formation of a dimer (D). Note that the dimer fraction increases with an initial zero slope. (**B**) Evolution of the Te (blue), D (green), M (black) and sms (red) fractions as the function of the time.

increases from 10% in the absence of added NaCl to 70% in 0.4 M NaCl solution and decreases slowly to 50% in 2 M NaCl.

sms dissociation kinetics

Figure 7 shows the evolution versus time of the chromatograms recorded during sms dissociation. The sms fraction decreases exponentially with a time constant of 0.46 h at 40°C, pH 4.6. At equilibrium the residual sms fraction is negligible. The chromatograms recorded during the first 30 min of the experiment show the formation of a tetramer (Te), of a species (Te_2) whose molecular weight is about twice that of a tetramer and of monomer. The monomer half formation time, 0.5 h., is close to the sms dissociation time. At last, the chromatograms show the formation of a dimer. The initial zero slope of the plot of the dimer fraction versus time suggests that the dimer is not directly generated by sms dissociation but results of monomer dimerization. The monomer and dimer equilibrium concentrations correspond to a reduced dissociation constant Fi = 2.5 10^{-6} M for the dimer at 40°C, pH 4.6.

Effect of pH and temperature on the sms dissociation time

Figure 8 shows that the sms lifetime is maximal around pH 4.4 and decreases at higher and lower pH. The activation energy related to sms dissociation, E_{ac} = 255 ± 20 kJ/M, seems nearly independent of pH (Figure 8, lower panel). The sms lifetime around pH 4.4 is extraordinary long. The extrapolation of the values measured between 42 and 60°C indicates a lifetime of about 10 years at 20°C.

Other oligonucleotides

All the oligonucleotides of the C_7XC_4 family, where X is either A, T, G or an ethane-diol spacer, as well as C_4GC_7, C_4TC_7, C_6TC_3, C_7TC_3 and C_5TC_5 associate into sms. The

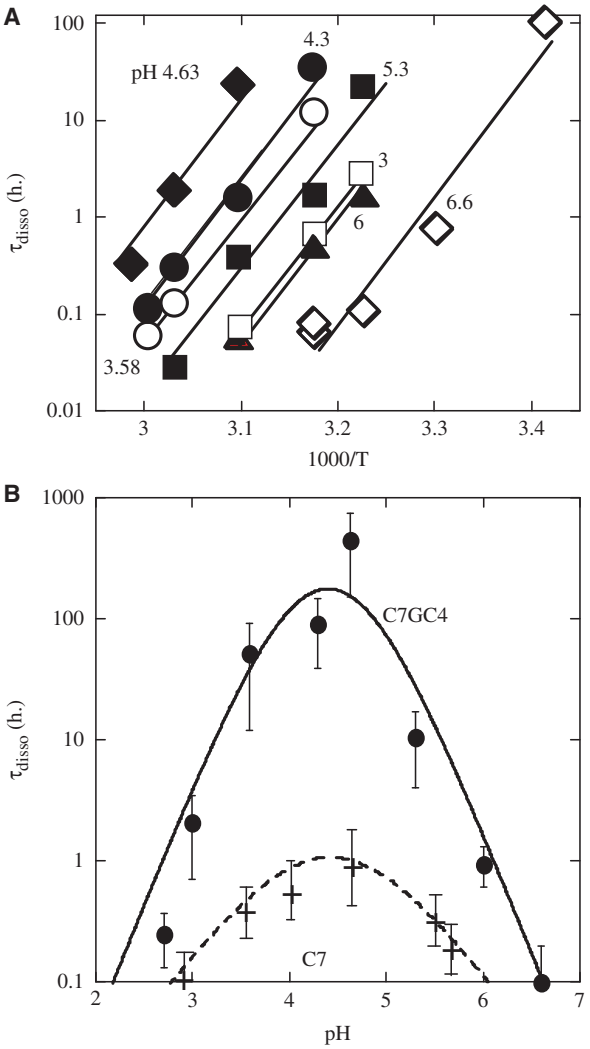


Figure 8. Effect of pH and temperature on the sms lifetimes in the NAP buffer. (**A**) Half dissociation time of the sms of C_7GC_4 versus temperature at pH 4.6 (black diamonds), pH 4.3 (black circles), pH 3.58 (open circles), pH 5.3 (black squares), pH 3 (open squares), pH 6 (black triangles) and pH 6.6 (open diamonds). (**B**) Lifetimes at 40°C versus. pH of the sms of C_7GC_4 (open circles) and C_7 (crosses). The sms lifetimes are maximal when the pH is equal to the cytidine pK_{N3} and decrease symmetrically at lower or higher pH.

sms molecular weights, formation and dissociation rates and the oligonucleotide fractions associated into sms at equilibrium depend on the length of the C stretches and on the nature of the X spacer. All the C_nXC_m oligonucleotides examined form preferentially a dimer rather than a tetramer. The dimer may be the thermodynamically stable multimer, as this seems to be the case for C_7GC_4 (Figure 3) or it may be kinetically trapped during a time exceeding experimental investigations. We also observed the assembly of C_7 into sms as originally reported by Yamuna Krishnan and collaborators (9).

C_4GC_7. The reduced dissociation constants of the dimer of C_4GC_7 is comparable to that measured for the reference C_7GC_4 oligonucleotide. The formation and dissociation rates of the sms of C_4GC_7 (Figure 4 and Supplementary Figure S4) are very similar to the rates measured for C_7GC_4.

C_7TC_4, C_4TC_7 and C_5TC_5. At pH 4.6, the reduced dissociation constants of the dimers formed by these oligonucleotides are at least 10 times smaller than that of the reference C_7GC_4 oligonucleotide**.** This shows that the nature of the non-cytidine spacer contributes to the dimer stability and therefore interferes with sms formation. The half association times of the sms of C_7TC_4 and C_4TC_7 are 20 and 5 times, respectively longer than that of C_7GC_4. At 42°C, pH 4.6 the formation rate of the sms of C_5TC_5 is about 100 times slower than that measured for the reference C_7GC_4 oligonucleotide. After incubation during 5 days, in condition close to equilibrium, the chromatogram of a 5 mM C_5TC_5 solution shows that the dimer, tetramer and sms fractions are 60, 20 and 20%, respectively.

C7. By comparison with C_7GC_4, the association rate of C_7 into sms is faster but the lifetimes of the sms of C_7 measured at different pH is much shorter (Figure 8).

DISCUSSION

A good knowledge of the factors that influence the sms growth is essential to control the assembly of C-rich oligonucleotides into structures with specific applications (16). The detail of the sms formation pathway is beyond the scope of this article; nevertheless some indications on the sms growing process may be derived from the experiments presented above.

In summary, we have shown that all the C_nXC_m oligonucleotides investigated associate into sms. The sms formation rates and equilibrium proportions depend on the C stretch length and on the nature of the spacer. It is interesting to note that the influence of temperature and pH on the sms formation kinetics is quite different from the effect of these factors on the i-motif formation kinetics (5). In contrast with the formation time of i-motif tetramers, which increases as the power of –2 of the monomer concentration (5), the sms formation time increases as the power of –1 (Figure 5). Moreover, we have observed that sms disruption releases short sms fragments, among which only Te_2 may be resolved on the GPC-100 column, and tetrameric structures (Te) that in turn dissociate into monomer. The similarity of the time constant for sms dissociation and monomer formation together with the exponential character of the evolution of these species indicate that tetramer dissociation is not limiting.

Competition between i-motif dimers and tetramers

A constant property of C_nXC_m oligonucleotides is their capacity to form stable i-motif dimers. It must be noticed that these dimers are certainly not parallel hemiprotonated duplexes. There is no experimental indication in the literature in favor of the existence of stable C-rich hemiprotonated duplexes in oligo-C solutions and

recent attempts to enforce their formation led to the conclusion that these species must be considered as short-lived precursors of i-motif tetramers in fast exchange with the monomer (7).

Two families of i-motif dimer have been described (2,6,17). One is formed by the parallel arrangement of two hairpins whose loops are on the same side of the i-motif core, the other by two hairpins in head to tail orientations (Scheme 2). Many conformers may coexist in each family, depending on the loop position with respect to the narrow and wide grooves, on the intercalation topology of the $C{\bullet}C^+$ pairs and on the loop composition. If one considers the structures displayed in Scheme 2, it is hard to imagine that they could induce or contribute to sms formation. It is therefore tempting to regard the dimer as a dead-end way in the sms formation pathway.

Several examples have been reported showing that formation of i-motif dimers is kinetically favored by comparison with tetramer formation. The thermodynamically stable multimer of $T_2C_8T_2$ (14), 5mCCTCTCC (5) and 5mCCTCTCTCC (7) are tetrameric. Nevertheless, after melting and fast cooling these oligonucleotides are kinetically trapped into hairpin dimers and it is observed that the dimer/tetramer conversion is extremely slow.

The formation time of i-motif multimers should increase in proportion with the number of the steps leading to the stable final structure. It is therefore reasonable to assume that this is the difference of the formation order of dimer and tetramer, 2 and 3, respectively (6,8), that accounts for the faster dimer formation rate.

Dissociation of $[C_nXC_m]_2$ dimers is the limiting step of sms elongation

The study of C_7TC_4 and C_4TC_7 gives indirect arguments suggesting that the dimer stability and the sms formation rate are correlated. The slow formation rate of the sms of these sequences may be a consequence of the stability of the dimers formed by these oligonucleotides. It may be also noticed about the stability of the dimers with a T spacer that the loop topology of the structure of Scheme 2C allows the formation of a cross-loop $T{\bullet}T$ pair (5) that could account for the enhanced stability of the duplexes with a thymidine spacer.

If one supposes that the competition between dimer and tetramer hinders sms growing, it is predictable that the pH and temperature conditions favorable to tetramer formation should accelerate the sms elongation. The simulations displayed in Figure 3b indicate that in response to the effect of pH on the stability of the dimer of C_7GC_4, the tetramer fraction increases between pH 5.3 and 6.3 by more than two magnitude orders and suggests that the acceleration of the sms formation rate observed in this pH range (Figure 6) results of dimer dissociation.

Figure 4 shows that the sms formation time decreases when the temperature is raised (E_{act} = 143 kJ/M for the sms of C_7GC_4, Figure 4). By contrast, the i-motif tetramer formation times increase with temperature (E_{act} = –197 and –306 kJ/M for $[TC_3]_4$ and $[TC_5]_4$, respectively (8). The lengthening of the formation time of i-motif tetramers at high temperature has been ascribed to the reduction of the proportion of an intermediate dimer or trimer species (8). By analogy with the interpretation given above to the effect of pH on the sms formation rate, the acceleration of the sms formation rate at high temperature may be a consequence of the preferential dimer dissociation at high temperature.

By contrast with the formation time of i-motif tetramers that varies as the power of –2 of the monomer concentration, the sms formation time varies as the power of –1 of the concentration (Figure 5). This indicates that the sms formation kinetics is not directly controlled by the formation rate of the i-motif building blocks (Te in Scheme 1). This argues also against an sms growing process involving successive addition of monomer units to a tetrameric nucleus whose rate should be strongly dependent on the oligonucleotide concentration.

A model for sms formation

The sms formation pathway displayed in Scheme 1 is in part inspired by that proposed to account for C_7 association into i-wire structures (9). All the arguments developed above together with the similarity of the time constants for sms formation and dimer dissociation suggest that dimer dissociation is the limiting step of the sms formation pathway. According to Scheme 1, it is noteworthy that since dimer formation is faster than tetramers formation (i.e. $kon_{dimer} > kon_{tetra}$), the dimer concentration should decrease at a rate $k = kof_{dimer} \times kon_{tetra}/(kon_{dimer} + kon_{tetra})$ which is slower than the dimer dissociation rate, kof_{dimer}.

As shown in a recent study of the formation pathway of i-motif tetramers, structures with partial or full intercalation topologies are formed at comparable rates and the species with optimal intercalation topology prevails at equilibrium due to its longer lifetime. The X spacer was inserted in the C_nXC_m sequence to prevent intercalation of the m + n cytidines into a single i-motif core.

The tetramer formed by intercalation of the longer C stretch (Te in Scheme 1) is expected to be the thermodynamically stable species and we postulate that it is the building block of the sms growing process. We assume that after the nucleation step, association by pairing and intercalation of the terminal C_4 stretches of two Te building blocks initiates sms formation. The i-motif symmetry gives to the assembly of several building blocks the same overhanging C_4 terminations than the building block itself. This allows sms elongation by association of preformed Te_n and Te_m assemblies into structures including (n + m) building blocks (Scheme 1) or by successive association of Te blocks.

The lifetime of i-motif tetramers increases strongly with the number of intercalated $C{\bullet}C^+$ pairs (8). It is therefore predictable that the connecting i-motif blocks formed by the C_4 stretches should be the weak sms links. However, it is interesting to note that the lifetime of the sms of C_7GC_4, about 100 h at 40°C, pH 4.5 (Figure 8), is about 600 times longer than that of the i-motif tetramer of TC_4 (8). This suggests that the i-motif block formed by intercalation of the C_4 stretch is stabilized in the sms by staking

interactions with the outer $C{\bullet}C^+$ pairs of the i-motif sections built by the C_7 moiety.

The structure of supramolecular i-motif assemblies

$C{\bullet}C^+$ intercalation into i-motif structures requires an extension of the helical rise at the C-C steps up to 6.3 Å, a value close to the maximum accessible without base-pair disruption (18). Due to this extreme helical stretch and compactness, the i-motif should be an unbendable structure and it is predictable that the i-motif sms should look like extended stiff structures.

We have little information at this time about the structure of supramolecular i-motif assemblies. The formation pathway proposed just above, supposes implicitly that the sms formed by C_nXC_m oligonucleotides are linear structures. Preliminary AFM investigations indicate that the structures seem markedly influenced by the nature of the X spacer. AFM images of C_7TC_4 samples deposited on mica show bent structures as long as micrometer with unexpected 120° bifurcations, those of C4AC7 and C4GC7 form rigid rods whereas the structures formed by of C_7GC_4 look like aggregated shapeless pellets. It is clear that further investigations are required to give a reliable description of the supra molecular formed by C_nXC_m oligonucleotides.

NMR is obviously not appropriated to structural studies of large sms. Nevertheless, the long sms lifetimes should allow purification of tetramer assemblies short enough to be accessible to structural investigations by NMR methods. It should be specially interesting to determine the arrangement of the X spacer in the structure (Te_2, Scheme 1) formed by association of two building blocks.

It is noteworthy that regardless of their structural differences, the G-quartet (19–21) and i-motif tetramers have similar formation pathways (8,15) and associate into supramolecular structures. While it is assumed that i-motif sms are generated by the association of incompletely intercalated preformed tetrameric building blocks (9), G-wire elongation is described as a process involving association of parallel duplexes to an out-of–register G-quadruplex matrix (22).

The lifetimes of partially matched G-quadruplexes and i-motif multimers may be extremely long (6,8,23). For this reason, the formation of both tetramer families is conditioned by the evolution of species that are kinetically trapped (8,15). It is therefore predictable that as observed for the supramolecular assemblies of i-motif, the kinetic trapping of mismatched species is also a crucial limitation to the elongation of G-wire structures.

The excellent yield of automated chemical DNA synthesis together with the availability of a large variety of artificial residues with functional specificities make DNA oligonucleotides an attractive building material for sms (24–28). The DNA supramolecular assemblies based on Watson-Crick pairs or on G-quartets must be heated at melting temperature to be dissociated. In contrast, the extreme i-motif sensitivity to pH allows dissociation of i-motif sms by a mild pH change from pH 6 to 7. This property may be extremely interesting in assemblies including standard DNA and i-motif sections by allowing a conformational change triggered by pH.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

Yamuna Krishna is gratefully acknowledged for discussion and information on i-wire formation. Eric Lesniewska is acknowledged for communication of AFM images.

FUNDING

Funding for open access charge: Centre National de la Recherche Scientifique.

Conflict of interest statement. None declared.

REFERENCES

1. Gehring,K., Leroy,J.-L. and Guéron,M. (1993) A tetrameric DNA structure with protonated cytidine-cytidine base pairs. *Nature*, **363**, 561–565.
2. Nonin,S., Phan,A.T. and Leroy,J.-L. (1997) Solution structure and base-pair opening kinetics of the i-motoif dimer of d(5mCCT TTACC): a non canonical structure with possible role in chromosome stability. *Structure*, **5**, 1231–1246.
3. Phan,A.T., Guéron,M. and Leroy,J.-L. (2000) The solution structure and internal motions of a fragment of the cytidine-rich strand of the human telomere. *J. Mol. Biol.*, **299**, 123–144.
4. Nonin,S. and Leroy,J.-L. (1996) Structure and conversion kinetics of a bi-stable DNA i-motif: broken symmetry in the [d(5mCCTCC)]4 tetramer. *J. Mol. Biol*, **261**, 399–414.
5. Leroy,J.-L. (2004) T.T pair intercalation and duplex interconversion within i-motif tetramers. *J. Mol. Biol.*, **333**, 25–139.
6. Canalia,M. and Leroy,J.-L. (2005) Structure, internal motions and association-dissociation kinetics of the i-motif dimer of d(5mCCT CACTCC). *Nucleic Acids Res.*, **33**, 5471–548.
7. Canalia,M. and Leroy,J.-L. (2009) $[5mCCTCTCTCC]_4$: an i-motif tetramer with intercalated T•T pairs. *J. Am. Chem. Soc.*, **131**, 12870–12871.
8. Leroy,J.-L. (2009) The formation pathway of i-motif tetramers. *Nucleic Acids Res.*, **37**, 4127–4134.
9. Ghodke,H.B., Krishnan,R., Vignesh,K., Kumar,P., Narayana,C. and Krishnan,Y. (2007) The I-Tetraplex building block: Rational design and controlled fabrication of robust 1D DNA scaffolds through non-Watson-Crick interactions. *Angew. Chem. Int. Ed.*, **46**, 2646–2649.
10. Cantor,C.R. and Warshaw,M.M. (1970) Oligonucleotide interactions. III. circular dichroism studies of the conformation of deoxyoligonucleotides. *Biopolymers*, **9**, 1059–1077.
11. Bhatia,D., Mehtab,S., Krishnan,R., Indi,S., Basu,A. and Krishnan,Y. (2009) Icosahedral DNA Nanocapsules by Modular Assembly. *Angew Chem.*, **48**, 4134–4137.
12. Plateau,P. and Guéron,M. (1982) Exchangeable proton NMR without base-line distortion, using strong pulse sequences. *J. Amer. Chem. Soc.*, **104**, 7310–7311.
13. Guéron,M., Plateau,P. and Decorps,M. (1991) Solvent signal suppression in NMR. *Prog. NMR spectr.*, **23**, 135–209.
14. Leroy,J.-L., Gehring,K., Kettani,A. and Guéron,M. (1993) Acid multimers of oligo-cytidine strands: stoichiometry, base-pair characterization and proton exchange properties. *Biochemistry*, **32**, 6019–6031.
15. Bardin,C. and Leroy,J.-L. (2008) The formation pathway of tetramolecular G-quadruplexes. *Nucleic Acids Res.*, **36**, 477–488.

16. Davis,A.V., Yeh,R.M. and Raymond,K.N. (2002) Supramolecular assembly dynamics. *Proc. Natl Acad. Sci. USA*, **99**, 4793–4796.
17. Gallego,J., Chou,S.-H. and Reid,B. (1997) Centromeric pyrimidine strands fold into an intercalated motif by forming a double hairpin with a novel T:G:G:T tetrad: solution structure of the d(TCCCGTTTCCA) dimer. *J. Mol. Biol.*, **273**, 840–856.
18. Lebrun,A. and Lavery,R. (1996) Modelling extreme stretching of DNA. *Nucleic Acids Res.*, **24**, 2260–2267.
19. 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.
20. Burge,S., Parkinson,G.N., Hazel,P., Todd,A.K. and Neidle,S. (2006) Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.*, **34**, 5402–5415.
21. Davis,T. and Spada,G.P. (2007) Supramolecular architectures generated by self-assembly of guanosine derivatives. *Chem. Soc. Rev.*, **36**, 296–313.
22. Marsh,T.C., Vesenka,J. and Henderson,E. (1995) A new DNA nanostructure, the G-wire, imaged by scanning probe microscopy. *Nucleic Acids Res.*, **23**, 696–700.
23. Mergny,J.L., De Cian,A., Ghelab,A., Saccà,B. and Lacroix,L. (2005) Kinetics of tetramolecular quadruplexes. *Nucleic Acids Res.*, **33**, 81–94.
24. Krishnan-Ghosh,Y., Dongsheng,L.D. and Balasubramanian,S. (2004) Formation of an interlocked quadruplex dimer by d(GGGT). *J. Am. Chem. Soc.*, **126**, 11009–11016.
25. Miyoshi,D., Karimata,H., Wang,Z.M., Koumoto,K. and Sugimoto,N. (2007) Artificial G-wire switch with 2,2′-bipyridine units responsive to divalent metal ions. *J. Am. Chem. Soc.*, **129**, 5919–5925.
26. Niemeyer,C.M. (2000) Self-assembled nanostructures based on DNA: towards the development of nanobiotechnology. *Curr. Opin. Chem. Biol.*, **4**, 609–618.
27. Shen,Z., Yan,H., Wang,T. and Seeman,N.C. (2004) Paranemic crossover DNA: a generalized Holliday structure with applications in nanotechnology. *J. Am. Chem. Soc.*, **126**, 1666–1674.
28. Pitchiaya,S. and Krishnan,Y. (2006) First blueprint, now bricks: DNA as construction material on the nanoscale. *Chem. Soc. Rev.*, **35**, 1111–1121.