DNA i-motif formation at neutral pH is driven by kinetic partitioning

Petra Školáková, Martin Gajarský, Jan Palacký, Denis Šubert, Daniel Renčiuk, Lukáš Trantírek, Jean-Louis Mergny, Michaela Vorlíčková

2950–2962 *Nucleic Acids Research, 2023, Vol. 51, No. 6* *Published online 13 March 2023*
https://doi.org/10.1093/nar/gkad119

DNA i-motif formation at neutral pH is driven by kinetic partitioning

Petra Školáková [1], Martin Gajarský [2], Jan Palacký [1], Denis Šubert [1,3], Daniel Renčiuk [1], Lukáš Trantírek [2], Jean-Louis Mergny [1,4] and Michaela Vorlíčková [1,*]

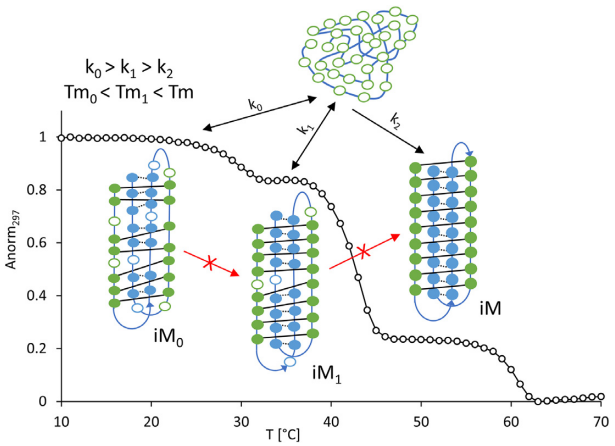
[1]Department of Biophysics of Nucleic Acids, Institute of Biophysics of the Czech Academy of Sciences, Královopolská 135, Brno 612 00, Czech Republic, [2]Central European Institute of Technology, Masaryk University, Kamenice 5, Brno 625 00, Czech Republic, [3]National Centre for Biomolecular Research, Faculty of Science, Masaryk University, Kotlářská 2, Brno 611 37, Czech Republic and [4]Laboratoire d'Optique & Biosciences, Institut Polytechnique de Paris, Inserm, CNRS, Ecole Polytechnique, Palaiseau 91128, France

Received September 16, 2022; Revised January 17, 2023; Editorial Decision February 02, 2023; Accepted February 14, 2023

ABSTRACT

Cytosine-rich DNA regions can form four-stranded structures based on hemi-protonated C.C$^+$ pairs, called i-motifs (iMs). Using CD, UV absorption, NMR spectroscopy, and DSC calorimetry, we show that model $(C_nT_3)_3C_n$ (Cn) sequences adopt iM under neutral or slightly alkaline conditions for $n > 3$. However, the iMs are formed with long-lasting kinetics under these conditions and melt with significant hysteresis. Sequences with $n > 6$ melt in two or more separate steps, indicating the presence of different iM species, the proportion of which is dependent on temperature and incubation time. At ambient temperature, kinetically favored iMs of low stability are formed, most likely consisting of short C.C$^+$ blocks. These species act as kinetic traps and prevent the assembly of thermodynamically favored, fully C.C$^+$ paired iMs. A higher temperature is necessary to unfold the kinetic forms and enable their substitution by a slowly developing thermodynamic structure. This complicated kinetic partitioning process considerably slows down iM folding, making it much slower than the timeframes of biological reactions and, therefore, unlikely to have any biological relevance. Our data suggest kinetically driven iM species as more likely to be biologically relevant than thermodynamically most stable iM forms.

GRAPHICAL ABSTRACT



INTRODUCTION

Non-canonical DNA structures are increasingly being studied by scientists, as they investigate the effects of these structures on functionally important genomic regions as well as their physiological or pathological relevance (1–4). One of the most fascinating structures is the i-motif (iM), arising from cytosine-rich sequences. The iM structure consists of two parallel duplexes of hemiprotonated C.C$^+$ pairs alternately intercalated in a mutually antiparallel fashion. Depending on the intercalation topology, the iMs are of two types denoted as 3′E and 5′E (5–7). Due to the necessity of cytosine protonation, iM formation depends on pH: lowering the pH (up to ∼4) will result in faster iM creation and greater stability (8–10). In contrast to the pK_A of free cytosine, the midpoint of the pH-induced iM formation (pK_T) is shifted toward a less acidic pH (11–15). This shift towards physiological pH makes it possible to speculate on the biological relevance of the iMs. iM-forming sequences are frequently found in gene promoters and may

*To whom correspondence should be addressed. Tel: +420 541 517 188; Email: mifi@ibp.cz

be involved in regulating gene expression (16). One possible way to achieve this regulation would be to stabilize the iM in the promoter regions using specific ligands. Two small molecules (IMC-48 and IMC-76) were found to control the expression of the *BCL2* gene supposedly through the stabilization/destabilization of iM (17). Additionally, *c-myc* and *KRAS* oncogenes may also be regulated by iM modulators (18,19). Several studies provided supporting evidence for the existence of iM in cells. Analysis of iM formation in the nuclei of human cells was performed with in-cell NMR spectroscopy (10,20) or using a specific antibody (21).

iM formation is influenced by many external or internal factors (22), such as the presence of metals (23), crowding agents (24–26), concentration and type of ions (27,28), and ligands (29,30). The structure and stability of iMs strongly depend on primary DNA sequence, such as loop length and composition (10,31,32) and the number of cytosines in blocks (33–35). A higher number of cytosines is associated with higher pK_T and stability of iMs (36). Long cytosine regions or repeated short C blocks can adopt iM even at close-to-neutral pH values (37,38). Some of such long sequences are present in functionally important regions of genomes (34). Four highly stable candidate sequences were identified in the promoter regions of DNA repair genes (39), the most stable having nine consecutive cytosines.

In our previous work (36), we focused on studying iM formation in DNA sequences with long cytosine blocks and tried to clarify the cause of their unusual melting in two steps, as reported by the Waller (34) and Burrows groups (39,40). We phenomenologically described the existence of two iM species differently populated depending on pH, C block length, and temperature. Still, we could not identify their structural variation or the cause of their formation. In this work, we provide a plausible model regarding the nature of these iM states. We show that there are more than two distinctly melting species in long iMs at pH close to neutral values and that this phenomenon is not limited to only long C blocks. The results demonstrate how iM formation is achieved by gradually lengthening $C.C^+$ paired blocks, thus providing insight into whether iM formation could be of biological significance.

MATERIALS AND METHODS

Sample preparation

All oligonucleotides used in this study (Supplementary Table S1) were synthesized and purified by desalting by Sigma-Aldrich. The lyophilized oligonucleotides were dissolved using 1 mM sodium phosphate buffer and 0.3 mM EDTA (pH 8) to give a stock solution concentration of approximately 10 mM in DNA nucleotides. The quality of oligonucleotides was checked by denaturing electrophoresis (20% acrylamide, 7 M urea, 1× TBE; run 1 h at 500 V in 1× TBE).

UV absorption and CD spectroscopy

The precise oligonucleotide concentrations were determined based on their absorbance at 260 nm measured at 90°C in the same solution on a Specord 250 plus spectrophotometer (Germany) using the calculated molar absorption coefficient ε_{260} (Supplementary Table S1) according to Kejnovská *et al.* (41).

Unless stated otherwise, all experiments were performed at 23°C in 0.5× Britton-Robinson buffer (K-RB) initially set to pH 8 (20 mM H_3BO_3, 20 mM H_3PO_4, 20 mM CH_3COOH, 37 mM KOH). For NMR and DSC experiments, in which high DNA concentrations are used, 1× K-RB buffer was chosen. The pH values were then gradually adjusted by adding small amounts of 2 M HCl directly to the cells and measured using a Mettler Toledo pH meter with an InLab micro pH electrode. The total volume of HCl added was ∼11 μl, corresponding to about 1% of the whole volume. Changes in DNA concentration were accounted for. Samples were left to equilibrate, and all measurements were taken the following day without thermal annealing. Annealing, if needed, was performed in a standard manner (90°C for 5′ followed by slow cooling).

CD spectra were measured using a Jasco J815 spectrometer (Japan) in 1 cm (unless stated otherwise) path-length Hellma cells in a Peltier holder. Strand concentrations of oligonucleotides were 1.5–4 μM (∼0.08 mM in nucleoside residues). Circular dichroism was expressed as $\Delta\varepsilon = \varepsilon_L - \varepsilon_R$ in units of $[M^{-1}.cm^{-1}]$ (the molarity being related to DNA nucleotide concentration). The pH titration curves were changes of Δε at 287 nm (CD maximum for iM formation). The experimental data were fitted by a three-parameter sigmoidal function using SigmaPlot 10 software (Systat Software Inc., USA).

UV absorption melting curves

Prior to melting, all samples were left overnight at 23°C to achieve equilibrium. Melting/annealing curves (23–85°C, 85–5°C and again 5–85°C) were measured in 1cm cells using a Varian Cary 4000 spectrophotometer (Australia) or using a Specord 250 plus spectrophotometer (Germany). The temperature was increased/decreased in 1°C steps and, unless stated otherwise, the samples were equilibrated for 2 min before each measurement (together with Peltier holder tempering and the spectra measurement 4 min/1°C) to give an average rate of temperature change of 0.25°C/min. For kinetic experiments, the equilibration time was either shortened or extended. Different average rates of temperature change (time of equilibration at every 1°C) were considered in Figure 2, Supplementary Figures S4, and S9. Melting/annealing curves were monitored by absorbance at 297 nm and expressed as 0–1 normalized data. Melting temperatures (T_ms) were determined using a dual baseline-correction procedure according to (42). The melting/refolding experiments were repeated at least three times, and T_ms (determined for a standardly used rate of temperature change of 0.25°C/min) were calculated as average values. Standard deviations were derived from repeated measurements and were 0.2 – 1.8°C for different sequences.

To obtain interrupted melting or annealing curves, the melting/annealing process ('stopped gradient experiments') was run to the desired temperature (30°C, 38°C, 44°C or 54°C), then the samples were incubated for a defined period of time (usually 24 h) at the respective temperature in a laboratory incubator, and then the melting/annealing process was continued.

The kinetics of iM formation was measured using a Specord 250 plus spectrophotometer. The samples were heated to 90°C and then cooled quickly to room temperature. Transition curves were monitored by following the absorbance at 297 nm, and individual points were recorded every 5 min. The curves were fitted using Sigma Plot software.

Unless stated otherwise, all experiments were undertaken in 0.5× K-RB buffer, pH 6.5.

NMR experiments

NMR experiments were performed on a Bruker Avance III HD 700 MHz NMR spectrometer (USA) using a room temperature inverse (1H-BB) probe. 1D ^{1}H NMR spectra were acquired using the p3919 pulse sequence (standard Bruker library) in K-RB buffer containing 10% D_2O. The spectra were processed with TopSpin v4.0.6 (Bruker) and MNova v14.1.2 (Mestrelab Research, Spain).

Native electrophoresis

Non-denaturing polyacrylamide gel electrophoresis was performed in a thermostatted submersible apparatus (SE-600; Hoefer Scientific, San Francisco, CA) with buffer circulation. Gels (16%, 29:1 acrylamide:bisacrylamide), were run for 19 h at 30 V (2 V/cm) and 23°C in 0.5x K-RB, pH 6.5. 2 μg of the respective oligonucleotides in the same solvent were loaded on the gel. As a marker, we used the GeneRuler Ultra Low Range DNA ladder (bands refer to the number of bases).

The gels were stained with Stains-All (Sigma-Aldrich) and digitalized using Personal Densitometer SI 375-A (Molecular Dynamics, Sunnyvale, CA).

DSC calorimetry

Microcalorimetry experiments were performed using a Nano DSC calorimeter (model 602000) from TA instruments with a sample/reference cell volume of 0.3 ml. The oligonucleotides were prepared in K-RB buffer, pH 6.54 at approximately 60 μM strand concentration. CD spectra of samples were taken before and after DSC measurements to rule out any DNA contamination or degradation. Buffer and sample solutions were degassed for 10 min prior to loading into DSC cells. The experiments were performed at a constant pressure of 6 atm. The excess heat capacity is related to oligonucleotide strand concentration. The samples were first pre-incubated in the DSC cell for defined periods of time at defined temperatures (both data specified in figure legends), then rapidly cooled down to 1°C and finally the thermograms were recorded in the range of 1–75°C for both the heating and cooling direction using a scan rate of 1°C/min. The basic analysis of DSC scans was performed using NanoAnalyze Software (version 3.7.5 from TA instruments). First, heating/cooling buffer scans were subtracted from the corresponding sample scans. These DSC scans (representing heat flux [in μwatts] in the sample cell against temperature) were then normalized concerning scan rate and oligonucleotide concentration to obtain dependences of molar heat capacity [in kJ/(mol.K)] on temperature. Afterward, a manually defined polynomial baseline was subtracted. Finally, thermodynamic parameters were computed for all transitions (DSC bands) by directly integrating molar excess heat capacity against temperature using the conventional relations for enthalpy and entropy without the need to use a particular thermodynamic model. The transition (melting) temperature for each DSC band was then determined as the temperature corresponding to the maximum band signal.

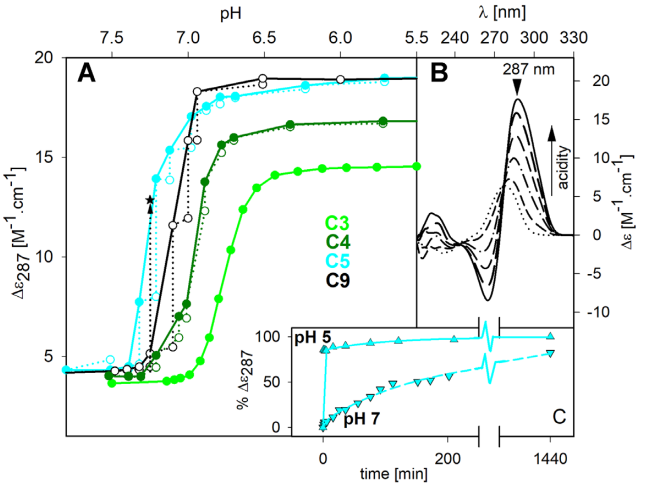


Figure 1. (**A**) $\Delta\varepsilon_{287}$ of **Cn** sequences for $n = 3$, 4, 5 and 9 plotted as a function of pH: vertical dotted lines connecting open and full symbols correspond to values measured immediately and 24 h after pH changes, respectively. (**B**) CD spectra of **C9** measured at pH values decreasing from 8 (dotted) to 5 (solid) immediately after pH change. (**C**) Kinetics of iM formation of **C5** at pH 5 (solid) and pH 7 (dashed) expressed as the % of the complete iM formed at pH 5. All experiments were measured in 0.5× K-RB buffer at room temperature. The pH dependence of **C5** measured 24 h after pH adjustment is the last equilibrium dependence. The **Cn** sequences with $n > 5$ do not reach equilibrium within 24 h (attaining equilibrium at, e.g. pH 7.25 lasts 24 h for **C5**, 4 days for **C6**, ∼10 days for **C7**, ∼20 days for **C8** and >2 months for **C9** (see asterisk at this pH)). This indicates that the iM transitions of **Cn**s with $n > 5$ no longer shift towards more alkaline values.

RESULTS

We studied iM formation by model $(C_nT_3)_3C_n$ DNA sequences (hereafter referred to as **Cn**) composed of four repeated blocks of n ($n = 3$–9) cytosines separated by three T triplets. Three nucleotides between C blocks were found to be optimal: shorter sections lead to cytosine removal into loops (36), while longer loops possibly decrease iM stability (31). iM formation was monitored by CD and UV absorbance changes, as shown in Figure 1 and Supplementary Figure S1.

iM formation of Cn sequences longer than C3 takes place at neutral pH

Figure 1 depicts the pH-induced formation of the iM structures of sequences differing in the length of **Cn** blocks monitored by the Δε values of the positive CD maximum. We followed iM formation by rapid acidification of a basic pH solution (pH 8) containing the unstructured **Cn** sequences: the switch to acidic pH leads to folding all sequences into

iMs. All the iMs formed are intramolecular (Supplementary Figure S2A). The pH-induced transition of **Cn** to iM shifts towards neutral and even slightly alkaline pH values with the increasing length of the C blocks (Figure 1). At room temperature, **C3** transition to iM structure occurs in the pH range 6.5–7.0, **C4** starts converting to iM at pH 7, and longer **Cn**s adopt iM structures even at slightly alkaline pH. All studied sequences adopt iM at pH 6.5 (Figure 1); thus, further experiments were performed (unless stated otherwise) at this pH.

No time-dependent changes were observed with the CD spectra in the course of the pH-induced iM transition of **C3**, which means that all points reach their equilibrium values within less than 4 min at pH $\leq$ 6.5 (4 min is the time of our standardly used repeated CD measurement) (Supplementary Figure S3A). When approaching pH 7 (i.e. under 'proton-deficient' conditions), iM formation becomes slow, especially for sequences with long cytosine blocks (Figure 1 and Supplementary Figure S3). Spectral changes are initially rapid upon acidification, but the final equilibrium structures at room temperature are only adopted with long-lasting kinetics. This can be clearly seen for long sequences ($n > 6$) (Supplementary Figure S3B). iM formation is thus initially kinetically driven, and only later, the thermodynamically stable structures at a given temperature are adopted.

Figure 1 plots $\Delta\varepsilon$ values measured immediately and 24 hours after adjusting each pH. The values differ distinctly due to the long-lasting kinetics. For cytosine blocks longer than 6 (Supplementary Figure S3B), the kinetics are so slow that the points in the course of the pH-induced transitions do not reach equilibrium values within 24 h (Supplementary Figure S3B). This is why the transition of **C9** (Figure 1A) appears at lower pH values than the transition of **C5** (both sequences were kept for 24 h at each pH, **C5** reached equilibrium, while **C9** did not). Attaining equilibrium for **C9** at pH values >7 takes tens of days (Supplementary Figure S3C).

Length of cytosine blocks affects properties of iM melting and refolding

The thermostability of **Cn** iMs increases with the length of C blocks (Figures 2 and 3). As mentioned above, the **C3** sequence melts fast, and refolding also proceeds rapidly and without any hysteresis between the two processes at the standardly used 0.25°C/min rate of temperature change. **C4** also forms iM structure relatively quickly, and the hysteresis is only modest. When n becomes larger than 4, the hysteresis increases, indicating that the kinetics of both melting and refolding processes become slow (and mutually different). The melting is, however, completely reversible upon returning to starting room temperature. The dependence of hysteresis on the lengths of particular sequences is shown in Figures 2 and 3B–D, and Supplementary Figure S4 for various rates of temperature change. Hysteresis can be expressed as the difference in apparent melting (T_m) and renaturation (T_{ren}) temperatures: $\Delta T = T_m - T_{ren}$. Hysteresis increases with C block length and heating and cooling rates. While choosing slow temperature gradients (0.025°C/min) leads to almost reversible profiles for **C4** (Figure 2) ($\Delta T_m < 1$°C), a large

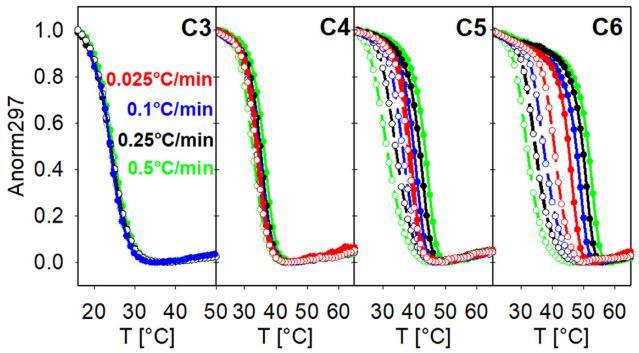


Figure 2. Normalized melting (full symbols) and refolding (open symbols) curves of **Cn** sequences at pH 6.5. Melting/refolding curves were measured at 1°C intervals with different rates of temperature change: 0.5°C/min (green), 0.25°C/min (black), 0.1°C/min (blue) or 0.025°C/min (red).

hysteresis ($\Delta T_m > 10$°C) remains for longer motifs (Supplementary Figure S4). In other words, approaching thermodynamic equilibrium using slower gradients is impractical for long **Cn** oligonucleotides.

Long Cn motifs melt in two distinctly separated steps

Besides playing a role in the magnitude of the hysteresis phenomenon, the length of the C-block significantly impacts the melting profile. Melting of **C3**–**C6** proceeds in a single transition (a single cooperative sigmoidal transition is visible) with T_m values in the range of 27.2–49.6°C that linearly increases with the number n (Figure 3B, Supplementary Table S1). Starting with **C7** (with the 0.25°C/min rate of temperature change), the behavior changes radically, and melting becomes biphasic (34) with two discrete transitions separated by a significant plateau. The first low-temperature melting process (T_{m1}) (to be consistent with tables and dependences throughout the paper, we call the apparent, non-equilibrium melting temperature T_m) and the main, high-temperature melting (T_m) increase in parallel with increasing n with a constant T_m difference ($T_m - T_{m1}$) of ~20°C (Figure 3). At a constant rate of temperature change, the contribution of the first transition (species 1) to the whole melting process increases with increasing n at the expense of the second transition occurring at a higher temperature (Figure 3). Decreasing the rate of temperature change diminishes the extent of the first melting transition for the benefit of the main melting, and the hysteresis of the transition becomes lower (Supplementary Figure S4). The two-step melting remained preserved for analogs of **C9** with mutated loop sequences for AAA, TAA and AAT. Its course was similar, but all three mutations led to a diminished population of the more thermostable iM species (Supplementary Figure S5).

Absorption and CD spectra measured as a function of temperature exhibit isosbestic points (Supplementary Figure S1), as expected for a single two-state transition, although two distinct melting profiles were present. It suggests that the distinctly thermostable iM species only differ quantitatively. The anomalous melting of the long **Cn** motifs is observed at pH approaching neutrality, thus under conditions where protons are in short supply. As the

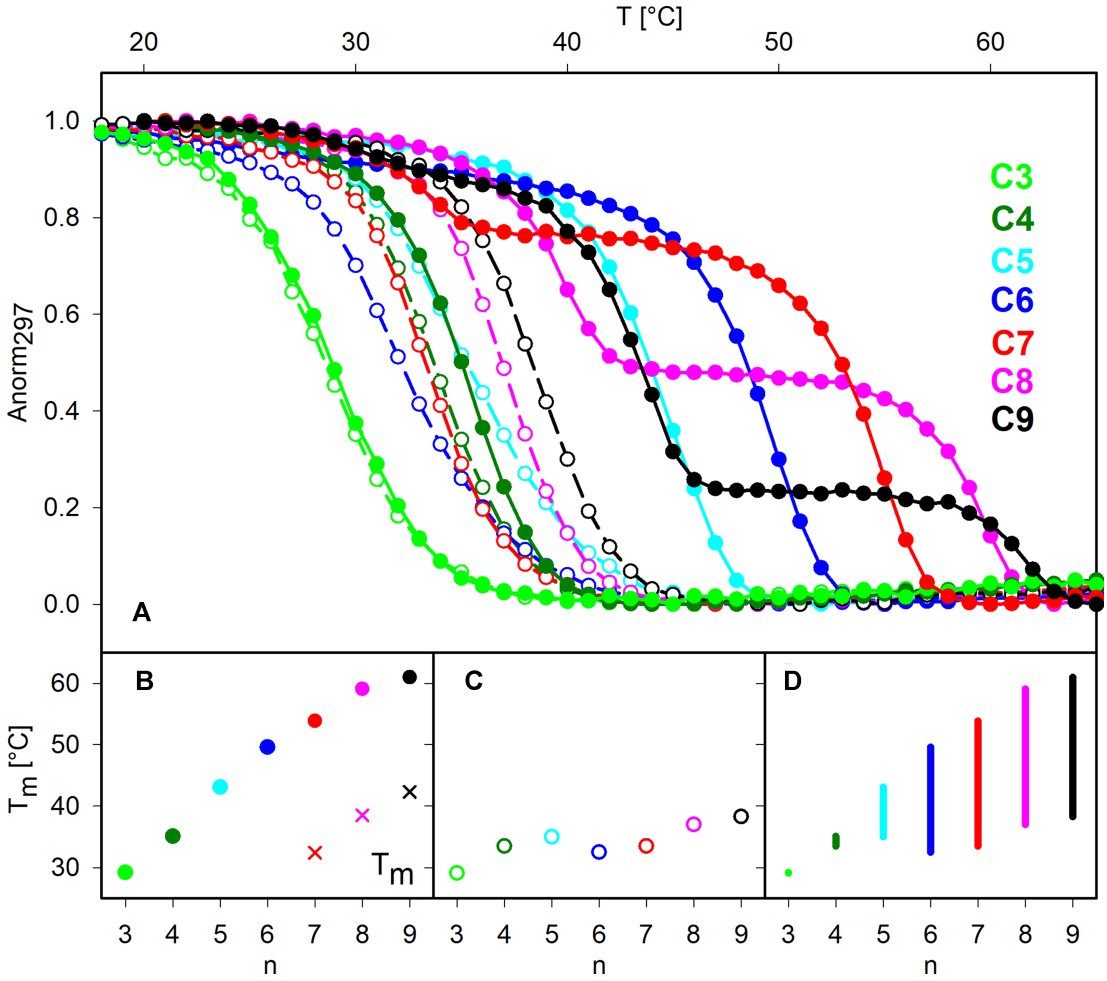


Figure 3. (**A**) Normalized melting (full circles) and refolding (open circles) profiles of studied **Cn** sequences monitored by absorption at 297 nm at pH 6.5, measured with the 0.25°C/min rate of temperature change. (**B**) Apparent melting temperatures of the more stable (circles) and the less stable (crosses) iM species, (**C**) T_{ren}, and (**D**) hysteresis ($\Delta T = T_m - T_{ren}$) plotted against the number (n) of cytosines in blocks. The numerical values (average of three measurements) are in Supplementary Table S1.

pH of the solution becomes acidic, the contribution of the first transition decreases and completely disappears at pH 5 (Supplementary Figure S6). In parallel, the hysteresis of the melting and refolding process gradually reduces.

Interestingly, the first (low temperature) melting transition of **C7** occurs close to the onset of **C4** melting. In the same way, the first step of **C9** melting mirrors that of **C5** (Figure 3), and, similarly, the first transition for **C8** occurs in the same temperature range as **C5**,**4**, a motif with an odd number of C.C$^+$ base pairs (Supplementary Figure S7A). Thus, the first (low temperature) melting process of long **Cn** molecules occurs in the same temperature range as the melting of iM sequences with roughly half the number of cytosines. Substitution of the long **C9** blocks by C_4TC_4 results in the disappearance of the two-step melting, leaving only the low-temperature transition: the sequence $[(C_4TC_4)T_3]_3(C_4TC_4)$ called C_4TC_4 melts like **C4** (Figure 4) and $C_4TC_4T_3C_4TC_4$, which corresponds to half of C_4TC_4 (Supplementary Figure S7B). The same result follows from the experiment when we substitute Ts instead of the middle Cs of **C11**: $[(C_5TC_5)T_3]_3(C_5TC_5)$; thus, C_5TC_5 melts like **C5** (Figure 4).

In contrast to melting, refolding of all studied sequences (cooling curves), exhibits a single transition at all pH and temperature gradients tested (Figure 3 and Supplementary Figure S8). The T_{ren} values do not strictly depend on n (Figure 3C): T_{ren} increases with increasing n up to $n = 5$ but stops increasing for **C6**, and then again increases for n higher than **7**. This T_{ren} decrease of **C6** and **C7** results from the very slow kinetics of their iM formation as it fades away with a decreasing temperature step, and the T_{ren} values become relatively constant for $n > 6$ at prolonged temperature increase (Supplementary Figure S9B). In addition, the slope of the refolding dependencies is less steep for **C5** and **C6** than for the other **Cn**s (Supplementary Figures S8 and S9D), which may indicate a more heterogeneous, multi-step process.

Differential scanning calorimetry (DSC) of selected sequences was performed to determine whether the spectral changes correspond to thermodynamically different structures, which is vital to know (43). According to spectro-

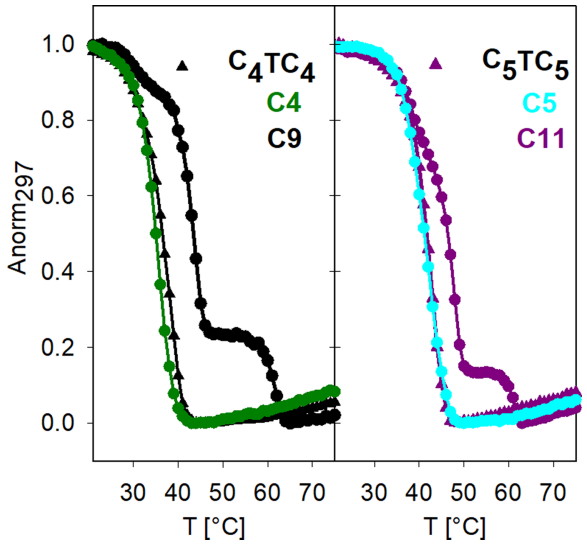


Figure 4. Normalized melting curves of **Cn** sequences (indicated) monitored by absorption changes at pH 6.5. With **C9**, a slight amount of unfolding takes place at **C4** melting, then at **C5** (as shown in Supplementary Figure S7A), and the main melting of the whole molecule takes place separately at high temperatures. A similar course of melting occurs with **C11**, and its analog with middle Cs substituted by Ts melts identically as **C5**.

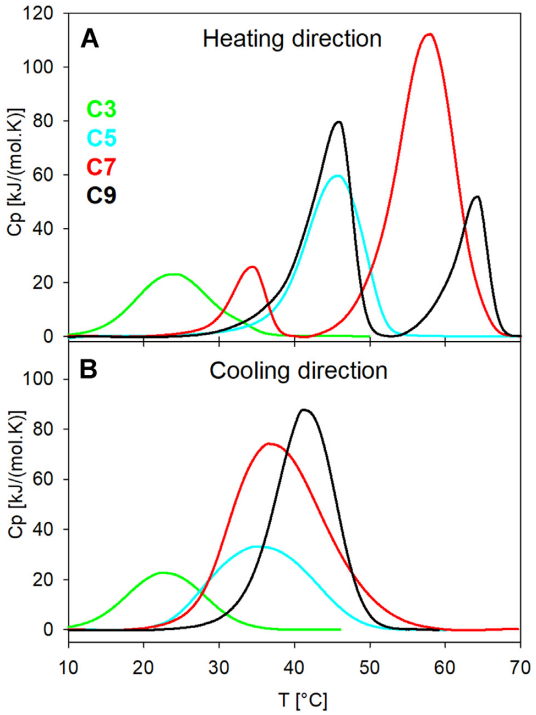


Figure 5. Heating (**A**) and cooling (**B**) scans measured in K-RB, pH 6.5 correspond to keeping the samples of **C3**, **C5**, **C7**, **C9** for 4 h at 20°C. The cooling scans show only one transition whose maximum is shifted to lower temperatures relative to the transitions of the more stable forms. No hysteresis between the heating and cooling scans is observed for **C3**. The maxima of cooling scans are generally smaller at the expense of their greater width, indicating less cooperativity during the renaturation process.

scopic measurements, thermal denaturation of **C3** and **C5** gives a single transition when heating the samples at pH 6.5. Contrastingly, **C7** and **C9** sequences show two distinct peaks upon heating (Figure 5), as in UV absorption-melting experiments. All the sequences show only one peak upon cooling (Figure 5B), in line with absorbance measurements. The first peak of **C7** appears somewhere in the middle between the melting profile of **C3** and **C5**, while the second, more populated one culminates at around 55°C. The first calorimetric peak of **C9** appears at the same temperature as the peak of **C5**, and it is more populated than the high-temperature peak. The calorimetric results are, therefore, in excellent agreement with the spectroscopic ones (Figure 5). Noteworthy, only the DSC peak of **C3** is symmetrical; all calorimetry peaks of the other sequences are asymmetrical (Figure 5). In contrast, no noticeable asymmetry is observed upon cooling the samples.

Checking the reproducibility of the two-step iM melting process reveals a new transition

The following experiments were performed with the **C9** sequence, for which the studied phenomena are most noticeable. The two-step melting at pH 6.5 remains unchanged when the process is repeated (rounds of heating and cooling). Similar profiles are obtained when using various sample preparation protocols, e.g. when measured immediately after preparation, one day or one week before the measurement, with fast or slow return from denaturation to room temperature (23°C), or after prior acidification to pH 5 and return to pH 6.5, as shown in Supplementary Figure S10A, B. In addition, fast cooling in an ice bath after denaturation, followed by 24 h incubation at 23°C does not change the profile of the two-step melting process (Supplementary Figure S10C).

In contrast, fast cooling to 0°C from denaturation immediately followed by melting experiment results in a new temperature-induced transition around 30°C (Figure 6 and Supplementary Figure S10C). We call this transition a 'zero transition' and the corresponding melting species a 'zero iM'.

The reversibility of temperature-induced changes differs for particular melting steps

Using the **C9** sample prepared by flash cooling from high temperature to 0°C (see above), we examined the reversibility of these three separated denaturation steps monitored by absorption changes. We call them step 0, and the least stable iM species 0, then step 1, and iM species 1, and the transition occurring at the highest temperature, which takes place with all **Cn** sequences, and after which all strands are completely melted, is called the main step (Figure 6).

Keeping the sample 24 h at 30°C, i.e. in the middle of the zero transition, led to an increase in absorption up to the original value, and then, omitting the zero transition, the two-step melting took place. Waiting one day at 38°C, i.e. at the end of the zero transition, resulted in an increase in absorption again, and then only a tiny fraction of the iM structures melted in the melting step 1, while the rest became more thermostable. We checked that all iM structures were intramolecular (Supplementary Figure S2B, C). A similar observation was made when interrupting the melting

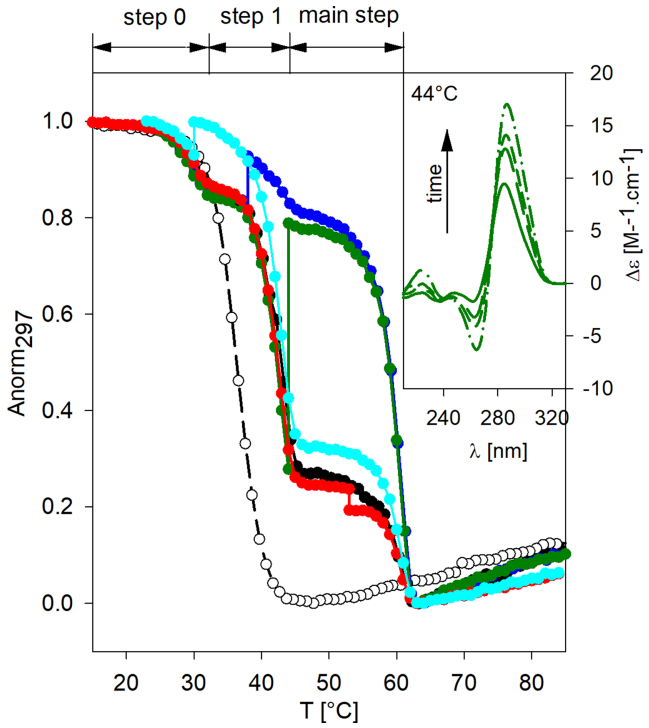


Figure 6. Normalized melting curves (black, filled circles – a control without interruptions) and refolding (black; open circles) of the **C9** sequence at pH 6.5 monitored by absorption at 297 nm. Temperature ramp was interrupted at 30°C (cyan), 38°C (blue), 44°C (green), or 54°C (red), keeping the respective temperature constant for 24 h, and then proceeding with thermal denaturation. (insert) CD spectra of **C9** as a function of time after interrupting the melting process at 44°C for 24 h.

process for one day at 44°C, i.e. once the first melting transition was nearly completed. We observed a time-dependent isothermal increase in absorption and CD, which indicated that iM was partially refolded (Figure 6 and insert), leading to the formation of iM species that melted at a temperature corresponding to the main melting transition. In other words, this prolonged incubation at an intermediate temperature led to the substitution of the kinetically preferred iMs for the most thermally stable iM structure.

Interrupting the melting process by waiting for 24 h at 54°C, i.e. in the middle of the plateau between the last two melting steps, does not lead to an increase in absorbance: no structure is reformed under these conditions. The kinetically favored structures have already melted, and only the remainder of the most stable part of iM species were present in the sample, which melts close to 60°C. The particular iM species may thus be formed depending on the incubation temperature. The formation of these structures takes some time, as they are constructed more slowly at elevated temperatures, and perhaps also because shorter sections need to be melted first so that longer iM blocks can be formed.

NMR spectroscopy also confirmed the reversibility of iM melting steps (Figure 7). Series of 1D ^{1}H NMR spectra were recorded on a gradually heated **C9** sample. iM unfolding was monitored via temperature or time-induced changes in both, the integral volumes of imino signals at ∼15.5 ppm specific to the imino protons involved in C.C^{+} base pairs (Figure 7A) and the intensities of aromatic signals corresponding to the unfolded specie(s) (Figure 7B). No intermolecular species were formed at the high DNA concentration used for NMR measurement and increased temperature (Supplementary Figure S2D). Analogously to the absorption-based experiment (Figure 6), the melting was interrupted at 38°C for 17 h. During this period, the intensity of the imino signals gradually increased. The increase paralleled the gradual decrease in the volume of aromatic signals corresponding to the unfolded specie(s) (Figure 7A, B). These observations jointly confirm the refolding of iM at 38°C (Figure 7C, D). Although the entire melting curve could not be obtained due to the technical limitations of the NMR spectrometer, the NMR data fully agree with the conclusions reached by CD and absorption experiments.

The population of particular iM species can be temperature controlled

The population of the particular iM species can be thus modified by maintaining the sample for some time at a given temperature; When this incubation is performed at a mid-high temperature (≈ 40°C) and for a long time, a subsequent melting experiment indicates that a higher proportion of iM with high thermal stability is formed (Figure 8).

Here we start the experiment from a denatured state resulting from incubation at 90°C. iM is formed quickly after rapid cooling to 10°C (before the subsequent CD measurement can be taken) (Figure 8, insert). The fast-forming structure (zero species) melts in three separated steps (∼26°C, 42°C and 61°C; purple curve), and a similar situation takes place upon jumping from the denaturation temperature to 20°C (i.e. slightly below room temperature; red curve): iM again forms quickly, and the T_ms of particular species remain the same as with the previous sample. A return from the denaturation state to 30°C leads to the slightly slower formation of iM (insert in Figure 8), which no longer contains the zero iM form since the temperature is above its T_m and it has melted. During the 24 h, the previous form is substituted with a more stable one. A change in temperature from 90°C to 40°C, which corresponds to the midpoint of the first melting transition, leads to a much slower formation of an iM structure, which is, however, well ordered in longer C.C^{+} blocks according to its high thermostability. The same is true for the iM structure, which forms very slowly (Figure 8, insert) on returning the denatured sample to a temperature of 44°C. Its refolding is not completed within 24 h. The resulting iM structure melts in a single step with the highest T_m value corresponding to the melting of the whole molecule. Upon a temperature jump to 50°C from the denatured state, the iM is no longer reformed.

The populations of the particular iM species can also be modulated by changing the temperature gradient: much slower rates of temperature change than standardly used ones (0.1 or 0.025 instead of 0.25°C/min) lead to a relative increase in more stable iMs. This means that the less stable iM species are being slowly exchanged for more stable ones. iM folding is thus driven by kinetic partitioning. An even slower increase in the rate of temperature change

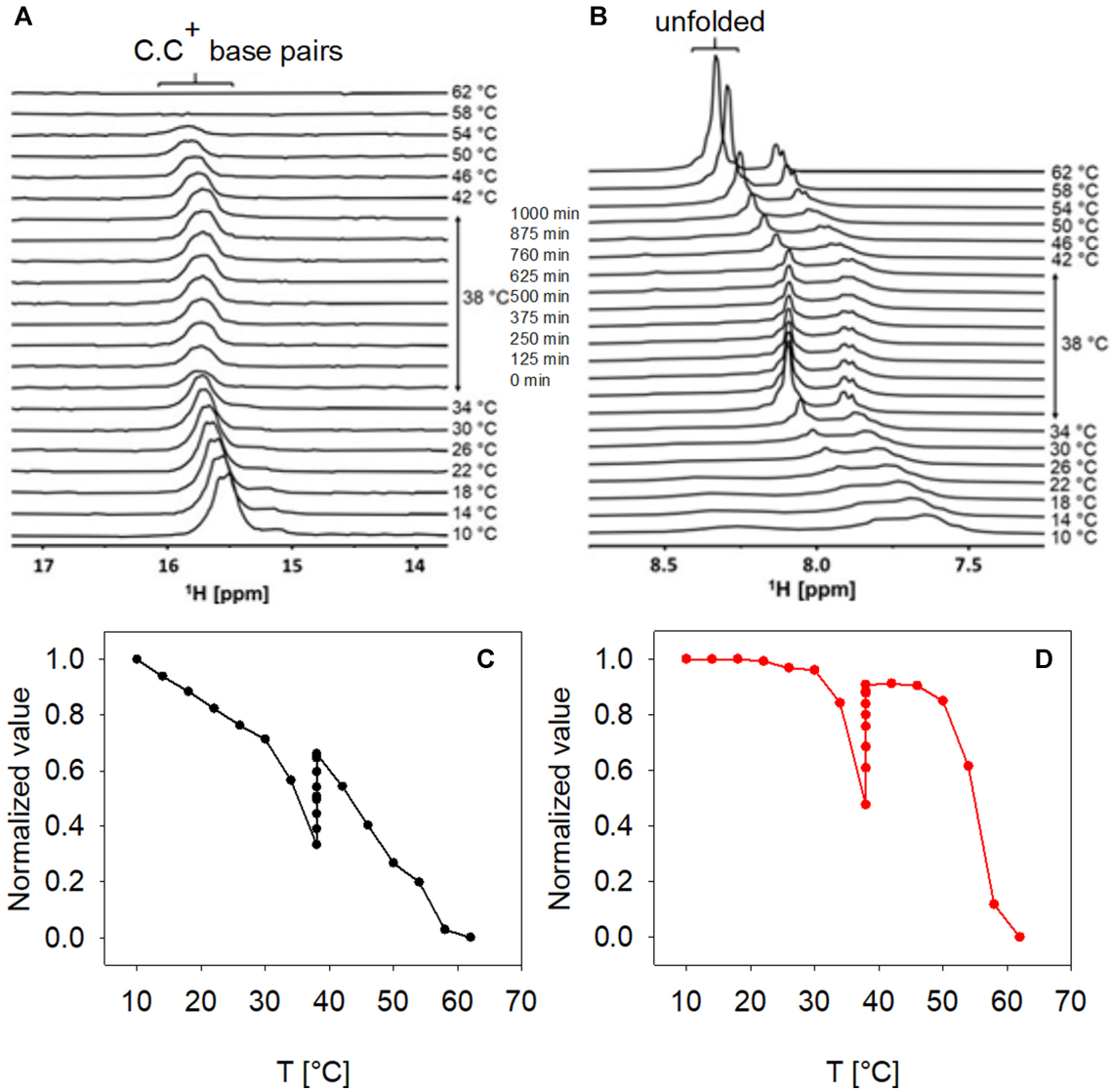


Figure 7. Imino (**A**) and aromatic (**B**) regions of 1D ^{1}H NMR spectra of **C9** acquired as a function of a temperature (indicated) in K-RB, pH 6.5. Temperature ramp was interrupted at 38°C keeping the respective temperature constant for 1000 min (~17 h). (**C**) The plot of normalized volumes of the iM-specific imino signals derived from 1D ^{1}H NMR spectrum of **C9** at pH 6.5 as a function of the temperature. (**D**) The plot of inversed normalized intensities of the degenerated aromatic signals from the unfolded form of **C9** at pH 6.5. Note: Temperature ramp was interrupted at 38°C keeping the respective temperature constant for 17 h.

(0.016$\bar{6}$°C/min, 60 min waiting at every 1°C) can even filter out the melting of the less stable species, leading to a continual formation of the most stable form, thus melting in a single step (Supplementary Figure S4, shown with **C9**). Interestingly, this transition is still not at thermodynamic equilibrium, as the denaturation and renaturation profiles taken at this slowest rate of temperature change are not superimposable (14).

The gradual formation of **C9** iMs was also studied using DSC, which provides model-independent thermodynamic values of the thermal transitions. The DSC profile of the **C9** sequence kept at 43°C is shown in Figure 9, while in particular time intervals, a heating/cooling cycle was run. Due to technical reasons (explained in the M + M section), only the temperatures above room temperature could be recorded. Thus, in line with previous results, we can see only two DSC peaks for heating, corresponding to the two melting transitions. The lower-temperature DSC transition peak considerably dominates over the higher-temperature one for shorter incubation times at 43°C (Figure 9A). At the same time, again in line with spectroscopic results, its relative weight gradually decreases as a function of the incubation time at the expense of the proportionately increasing weight of the higher-temperature transition. The sum of enthalpies of the two transitions remains constant irrespective of the ratio of the low/high stability peaks and is equivalent to the enthalpy of the single cooling transition (Figure 9B, insert). Interestingly, the low stability peak of **C9** practically disappears after maintaining the sequence for 8 hours at 43°C. Both DSC peaks are asymmetrical, and this asymmetry can be captured by a linear combination of at least 2–3 Gaussian components (left insert of Figure 9A). By con-

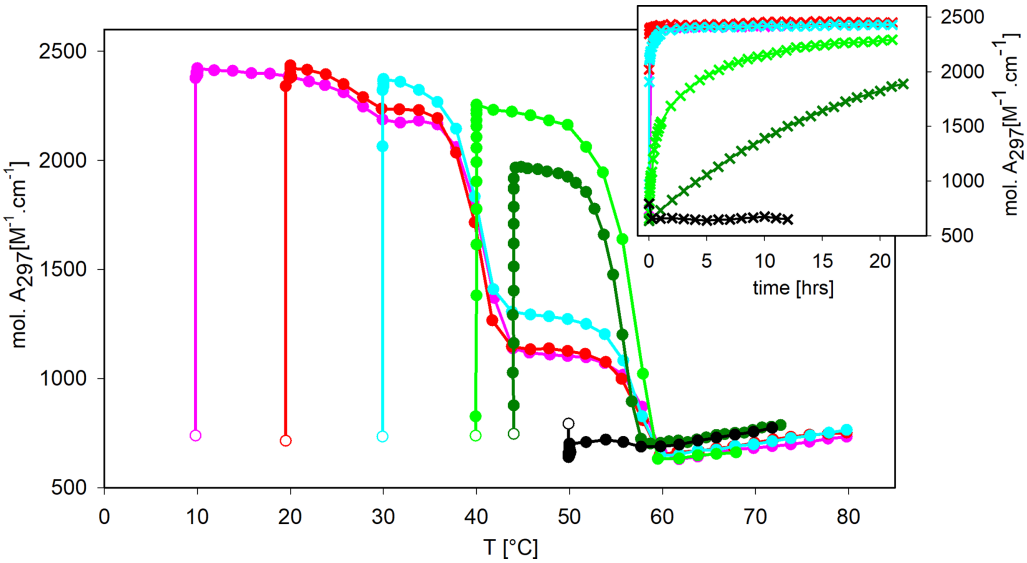


Figure 8. Melting curves of the **C9** sequence were monitored by molar absorption at 297 nm at pH 6.5. The experiment was started after heating the sample at 90°C, fast cooling it to particular temperatures (empty circles), and keeping it there for 24 h: 10°C (purple), 20°C (red), 30°C (cyan), 40°C (green), 44°C (dark green) and 50°C (black). (insert) Kinetics of iM formation of **C9** sequence (from the denatured state) at particular temperatures starting from fully denatured species at 90°C followed by flash cooling to the indicated temperature, where it is kept constant for 24 h.

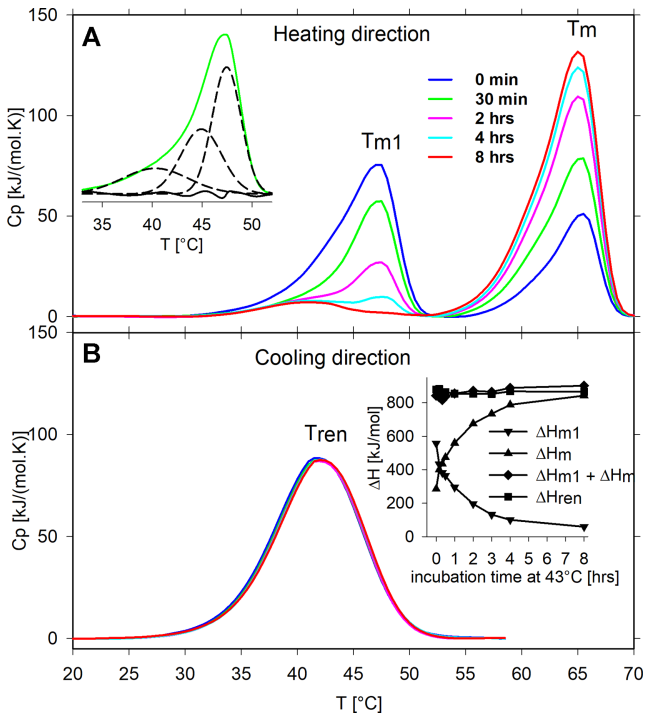


Figure 9. DSC of **C9** measured in K-RB pH 6.5. Heating (**A**) and cooling (**B**) scans correspond to particular time intervals of keeping the sample of **C9** at 43°C (according to the color legend in A). Both DSC peaks obtained upon heating exhibit marked asymmetry, further exemplified by the decomposition of the thermogram (corresponding to 30 min of incubation at 43°C) into three Gaussian peaks (inset in A). Enthalpy changes ΔH_1 and ΔH of both peaks obtained during heating (down and up triangles, respectively) and the corresponding enthalpy change ΔH_{ren} (squares) of the (cooling/annealing) peak against incubation time at 43°C are shown together with the sum of ΔH_1 and ΔH enthalpies in the inset of (B).

trast, no noticeable asymmetry is observed upon cooling at the same rate of temperature change (Figure 9B).

Distinct iM species differ in their interaction with proteins

To ensure pH stability, all the experiments in the paper are pursued in the non-physiological K-RB buffer. Principally the same temperature dependences as in the K-RB buffer are also observed (Supplementary Figure S11) in the physiological buffer mimicking ionic composition in the intracellular space (IC buffer: 25 mM sodium phosphate, 110 mM KCl, 10 mM NaCl, 1 mM $MgCl_2$, 130 nM $CaCl_2$). At the same pH, the iM is formed much faster in the IC than in the K-RB buffer (Supplementary Figure S11B), but the resulting iM structures are less stable (Table in Supplementary Figure S11). The shift of the IC buffer toward neutral pH leads to further significant iM destabilization (Supplementary Figure S11C). However, the distinct melting steps remain preserved (Supplementary Figure S11B, C).

We describe iM formation on model sequences with long C stretches, namely on **C9**, because it is stable at neutral pH and enables clear differentiation of particular iM forms during temperature changes due to their slow formation. The same temperature and time-driven iM formation mechanism also occurs for shorter **Cn** sequences (Supplementary Figure S4C). The exchange between distinctly melting iM species is much faster, so we can distinguish them only at high-speed temperature changes (1°C/min or 2°C/min) (Supplementary Figure S4C). The mechanism of iM formation via kinetic traps is thus likely generally valid for all iMs stable at particular pH and temperature values.

iMs incubated at various temperatures preceding the main melting (Figure 8) or for various time intervals at the increased temperature preserve the respective, once stabilized iM structure upon returning to low temperature (Supplementary Figure S11B). It allows for isolating different

iM states and assessing their biological properties under the same experimental conditions. Supplementary Figure S12 shows that the differently induced iMs are distinctly resistant to P1 nuclease digestion, and, namely, as demonstrated in Supplementary Figure S13, the particular iMs differ in their binding capacity with cellular proteins: The protein binding ability decreases with increasing iM thermodynamic stability.

DISCUSSION

Model $(C_nT_3)_3C_n$ (**Cn**) sequences with $n = 3 – 9$ form intramolecular iMs at room temperature: **C3** at slightly acidic pH but longer sequences at neutral or even slightly alkaline pH values (Figure 1). The iM of **C3** is formed quickly but, when approaching pH 7 with longer **Cn** sequences, the kinetics slows down. The time needed for reaching equilibrium states increases with increasing *n*. The stability of iM sequences depends nearly linearly on the length of their C blocks (Figure 3B) (36). At pH 6.5, **Cn** sequences longer than 6 melt in two (34) or even three separated steps: part of their iM molecules starts melting at about 20°C before the main melting of the long iM molecules, and simultaneously with the melting of shorter **Cn** sequences (Figure 3). This suggests that the long **Cn** motifs allow the formation of less ordered iM species formed by correspondingly shorter C blocks. The absorption and CD spectra during both melting steps intersect at joint isosbestic points (Supplementary Figure S1), which are characteristic features of a single two-state transition between ordered and unordered forms. It is in line with an SVD analysis of these spectra (36); alternative structures are thus not likely. It means that the structures within both transitions only differ in a quantitative parameter—in the population of the ordered and unordered forms. Proton NMR also fails to evidence any conformational difference between the two iM species (Figure 7). Formation of the low-melting fraction as well as the presence of hysteresis are visible with long **Cn** sequences and in a scarcity of protons; both phenomena disappear or are less distinct at lower pH values and shorter **Cn** blocks (Supplementary Figure S6). For this reason, we chose to concentrate our efforts on the longest sequence considered, **C9**.

The course of the iM melting process at a given pH, thus the type of the respective iM species, depends on the temperature of its induction and the time the sample is kept at the given temperature. At low temperatures, the iMs form rapidly, the process is kinetically driven, and the species created probably involve short, $C.C^+$ paired blocks. $C.C^+$ formation likely takes place in a concerted and cooperative fashion, as for duplex or triplex folding (44). The smallest unit capable of forming an intramolecular iM is four $C.C^+$ base pairs (33), i.e. two parallel and antiparallel $C.C^+$ pairs. The low melting iM species likely correspond to the structures consisting of short iM blocks, however, their CD spectrum observed at low temperatures is relatively high. Thus, the structure cannot contain longer unfolded regions. Our previous results indicate (36,45) that the thermodynamic stability of iM is mainly determined by the number of consecutive regularly alternating intercalated parallel and antiparallel $C.C^+$ pairs rather than by the total number of $C.C^+$ pairs in the structure. It is why C_4TC_4 (a variant of **C9** with the middle C substituted by T) melts identically to **C4** (Supplementary Figure S7B). In the same way, C_5TC_5 (the same variant of **C11**) melts the same as **C5**.

Kinetics of iM formation has been extensively studied in a number of papers, *e.g.* (6,7,40,46). For the telomeric DNA sequence (6,7) 3′E and 5′E iM topologies are formed (upon transfer to acidic environments) with fast and slow kinetics, respectively, while the 5′E gradually replaces the kinetically driven 3′E structure. Furthermore, the thermodynamically more stable 5′E structure is supported by the formation of T-T wobble pairs arising in this arrangement. This explanation for the slow formation of iM, involving the presence of two competing structures, has also been accepted for long cytosine sequences (40): Detailed investigations of hysteresis at a constant temperature as a function of pH revealed that in addition to the different course of iM folding and unfolding, pH-induced unwinding can be, in some cases, multiphasic depending on the rate of iM induction by transferring the sample from pH 8 to 4.5. The observed changes were reported to be also influenced by the presence of a small population of dimers formed mainly at low temperatures and during rapid conversion to acidic pH (47). The observed differences in the pH-induced unwinding of rapidly folded species compared to the slowly annealed ones is similar to our findings and may even be of the same nature. We, however, provide here an alternative explanation for these phenomena.

At ambient temperatures, iMs with short $C.C^+$ blocks are formed with fast kinetics. The core motif is complemented by further $C.C^+$ stacks conforming to a condition of alternating parallel and antiparallel pairs (Figure 10A). The joining $C.C^+$ pairs need not immediately include neighboring cytosines (Supplementary Figure S14). This partially paired structure containing mismatches is then kinetically trapped. It does not change with time, not even after a month unless the temperature is increased. At higher temperatures, the lifetime of this imperfect iM is shorter, allowing for its slow dissociation and the creation of longer iM blocks leading to more thermostable structures (Figure 8). Formation of longer iM blocks may be kinetically disfavored as it first requires a reorganization, including the melting of the previously formed shorter iM blocks. It slows down the iM formation. iM creation follows the kinetic partitioning folding model. In contrast, the renaturation curve is continuous. The course of renaturation with the standard rate of temperature changes (0.25°C/min) and renaturation with waiting at particular temperatures is shown and explained in Supplementary Figure S15.

The length of C blocks (number of consecutive $C.C^+$ pairs) forming iMs at particular temperatures may be estimated from Supplementary Figure S16. The necessity of melting shorter blocks to enable the formation of longer ones can explain the separated melting steps of long iM molecules. Examples of structures that may correspond to forms iM_0, iM_1, and the wholly formed iM in our experiment with **C9** are schematically drawn in Figure 10B.

The presence of distinctly stable iM species in long **Cn** molecules and the dependence of their occurrence at a given pH on temperature and time of incubation at the fixed temperature were also deduced from DSC experiments. The asymmetry of the DSC peaks and their extension toward

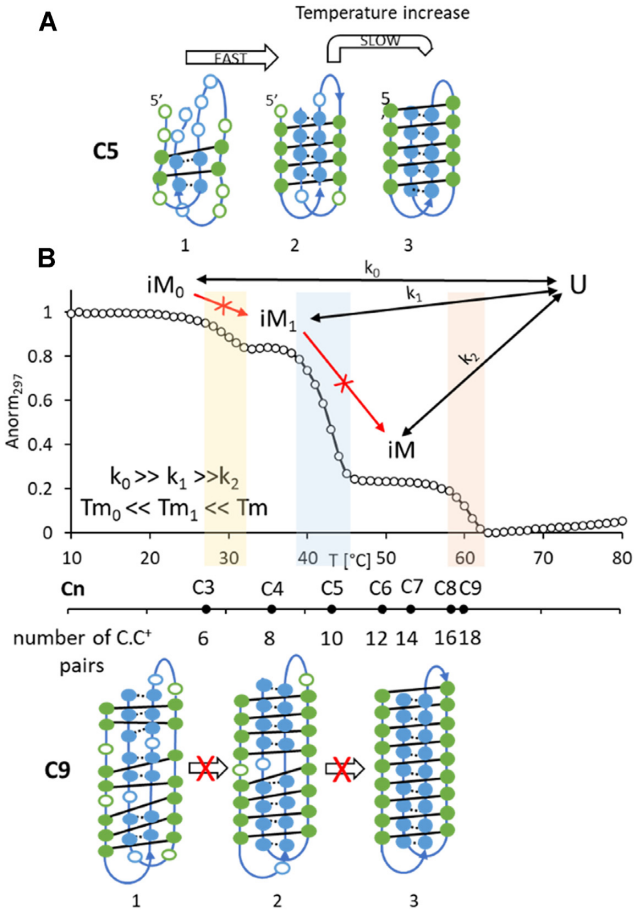


Figure 10. (**A**) The formation of a minimal iM nucleus (**1**) (paired and free cytosines are shown by filled and open circles, respectively) is followed by a fast zipping of other C.C$^+$ pairs conforming to a regular alternation of intercalating parallel and antiparallel strands (**2**, see also Supplementary Figure S11). The structure (**2**) has to melt to enable the formation of the perfectly paired iM structure (**3**). (**B**) The three-step melting profile of **C9** and corresponding possible iM species: The structure (**1**) may be an example of the imperfectly paired iM_0. Here it consists of three blocks of ∼4–6 consecutive C.C$^+$ pairs. The structure is then kinetically trapped and cannot accept further C.C$^+$ pairs as the unpaired cytosines are misaligned. The structure has to melt (U – unstructured form) to enable a proper rearrangement of the strands resulting in a structure (**2**) with longer iM blocks. Another melting is needed (here connected with the shift of the 5′E form for the 3′E one in both halves of the structure) until all cytosines are paired (**3**). The black points under the temperature dependence correspond to the T_m values of particular **Cn** iMs. The region of iM_0 melting (yellow column) may correspond to the melting of four to six consecutive C.C$^+$ pairs, and the melting of iM_1 (light blue column) to 8–10 consecutive pairs. These values match closely with the apparent melting temperatures of particular **Cn** iMs (Supplementary Table S1).

lower temperatures reveal the melting iM species' heterogeneity and the presence of shorter iM blocks in particular melting processes. The partitioning model of iM formation also applies to shorter **Cn** sequences with $n > 3$, as it follows from Supplementary Figure S4C and their hysteresis and the asymmetry of their DSC peaks.

Long C-rich regions able to form iM at physiological conditions have been identified in genomes (34,39). We performed our own analysis (Supplementary Figure S17), searching for sequences $(C_{3-9}N_{1-7})_3C_{3-9}$ differing in C block and loop lengths. The longest C-run capable of iM formation found in the human genome is $(C_7N_{1-7})_3C_7$ if one excludes purely polyC runs as the one found in chromosome 2. We have used the model **C9** sequence to demonstrate the partitioning mechanism of iM formation. The mechanism remains valid also for shorter sequences that are frequent in the human genome.

i-motif-prone sequences have been suggested to contribute to the regulation of transcription of genes (17,48) and DNA biogenesis (49). These complex biological processes are stringently orchestrated in space and time and occur within narrow limits of cellular homeostasis. Important homeostatic parameters concerning the iM formation include an intracellular pH, which is reported to vary between 7.0 and 7.4 (50), and physiological temperature (∼37°C). A time factor is particularly important: During transcription/replication, the proteinaceous factors bind their specific DNA sites on a time scale of seconds, and their residence time usually varies from seconds to minutes (51–53). Our data thus suggest that the biological function of i-motif forming regions in genomes could be connected with kinetically favored folding intermediate states rather than with the thermodynamically most stable iM conformations. Further comparative studies *in cellulo—in vitro* will be necessary to understand iM formation under biologically relevant conditions.

CONCLUSION

The main objective of our work was to investigate the properties of iM formation under physiological conditions to contribute to the question of its possible biological relevance. We studied at a constant pH, close to neutrality, temperature-induced folding and refolding of a set of oligonucleotides differing in the length of the C blocks. We provide a quantitative assessment of what 'slow' means for iM formation at near-neutral pH. Depending on the rate of temperature changes, partially ordered iM species with shorter C.C$^+$ blocks may melt separately at lower temperatures than their fully hemi-protonated iM forms, explaining the presence of multiphasic melting observed with longer **Cn** sequences. Depending on the temperature and incubation time, iM species with different length of their continuous C.C$^+$ blocks are formed, which determines their thermostability. Short C.C$^+$ blocks are formed rapidly at low temperatures, as they are kinetically driven and act as kinetic traps unless the temperature is raised. The respective iM species must first melt for the more stable iM with longer C.C$^+$ blocks at higher temperatures to form. Different iM species are therefore associated with particular temperatures. Overall, as iM formation can take *months* at pH 7.4 (Figure 1 and Supplementary Figure S3), this timeframe is incompatible with the time scale of biological processes: equilibrium structures cannot be reached in the cell. Nevertheless, kinetically driven species form quite efficiently and within the time scales compatible with iM-postulated biological roles. This suggests that kinetic species are more likely biologically relevant players than previously considered thermodynamically favored iM forms.

DATA AVAILABILITY

All data are available in the article or Supplementary information. Raw data will be shared on request to the corresponding/first author.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We acknowledge core-facility [Josef Dadok NMR Centre] of CIISB, Instruct-CZ Centre, supported by MEYS CR (LM2018127) and European Regional Development Fund-Project 'UP CIISB' (No. CZ.02.1.01/0.0/0.0/18_046/0015974) for access to research infrastructure. J.L.M. would like to dedicate this MS to the memory of Prof. Michel Rougée, a pioneer in the analysis of hysteresis of nucleic acids.

FUNDING

Czech Science Foundation [19-17063S to M.V., 19-26041X to L.T. and M.G.]; SYMBIT project [CZ.02.1.01/0.0/0.0/15_003/0000477] financed by the ERDF and ANR grant [ANR-21-CE44-0005-01] 'ICARE'. Funding for open access charge: SYMBIT project [CZ.02.1.01/0.0/0.0/15_003/0000477] financed by the ERDF.
Conflict of interest statement. None declared.

REFERENCES

1. Sen,D. and Gilbert,W. (1988) Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature*, **334**, 364–366.
2. Wells,R.D. (2007) Non-B DNA conformations, mutagenesis and disease. *Trends Biochem. Sci*, **32**, 271–278.
3. Balasubramanian,S. and Neidle,S. (2009) G-quadruplex nucleic acids as therapeutic targets. *Curr. Opin. Chem. Biol.*, **13**, 345–353.
4. Brooks,T.A., Kendrick,S. and Hurley,L. (2010) Making sense of G-quadruplex and i-motif functions in oncogene promoters. *FEBS J.*, **277**, 3459–3469.
5. Gehring,K., Leroy,J.L. and Gueron,M. (1993) A tetrameric DNA structure with protonated cytosine.cytosine base pairs. *Nature*, **363**, 561–565.
6. Lieblein,A.L., Buck,J., Schlepckow,K., Fürtig,B. and Schwalbe,H. (2012) Time-resolved NMR spectroscopic studies of DNA i-motif folding reveal kinetic partitioning. *Angew. Chem. Int. Ed. Engl.*, **51**, 250–253.
7. Lieblein,A.L., Furtig,B. and Schwalbe,H. (2013) Optimizing the kinetics and thermodynamics of DNA i-motif folding. *Chem. Biol. Chem.*, **14**, 1226–1230.
8. Fernandez,S., Eritja,R., Avino,A., Jaumot,J. and Gargallo,R. (2011) Influence of pH, temperature and the cationic porphyrin TMPyP4 on the stability of the i-motif formed by the 5'-$(C_3TA_2)_4$-3' sequence of the human telomere. *Int. J. Biol. Macromol.*, **49**, 729–736.
9. Cheng,M., Chen,J., Ju,H., Zhou,J. and Mergny,J.L. (2021) Drivers of i-DNA Formation in a Variety of Environments Revealed by Four-Dimensional UV Melting and Annealing. *J. Am. Chem. Soc.*, **143**, 7792–7807.
10. Cheng,M., Qiu,D., Tamon,L., Ištvánková,E., Víšková,P., Amrane,S., Guédin,A., Chen,J., Lacroix,L., Ju,H. *et al.* (2021) Thermal and pH Stabilities of i-DNA: confronting in vitro Experiments with Models and In-Cell NMR Data. *Angew. Chem. Int. Ed Engl.*, **60**, 10286–10294.
11. Manzini,G., Yathindra,N. and Xodo,L.E. (1994) Evidence for intramolecularly folded i-DNA structures in biologically relevant CCC-repeat sequences. *Nucleic Acids Res.*, **22**, 4634–4640.
12. Simonsson,T., Pribylova,M. and Vorlickova,M. (2000) A nuclease hypersensitive element in the human c-myc promoter adopts several distinct i-tetraplex structures. *Biochem. Biophys. Res. Commun.*, **278**, 158–166.
13. Dai,J., Hatzakis,E., Hurley,L.H. and Yang,D. (2010) I-motif structures formed in the human c-MYC promoter are highly dynamic–insights into sequence redundancy and I-motif stability. *PLoS One*, **5**, e11647.
14. Iaccarino,N., Di Porzio,A., Amato,J., Pagano,B., Brancaccio,D., Novellino,E., Leardi,R. and Randazzo,A. (2019) Assessing the influence of pH and cationic strength on i-motif DNA structure. *Anal Bioanal Chem*, **411**, 7473–7479.
15. Abdelhamid,M.A.S. and Waller,Z.A.E. (2020) Tricky Topology: persistence of Folded Human Telomeric i-Motif DNA at Ambient Temperature and Neutral pH. *Front. Chem.*, **8**, 40.
16. Hurley,L.H. (2001) Secondary DNA structures as molecular targets for cancer therapeutics. *Biochem. Soc. Trans.*, **29**, 692–696.
17. Kang,H.J., Kendrick,S., Hecht,S.M. and Hurley,L.H. (2014) The transcriptional complex between the BCL2 i-motif and hnRNP LL is a molecular switch for control of gene expression that can be modulated by small molecules. *J. Am. Chem. Soc.*, **136**, 4172–4185.
18. Shu,B., Cao,J., Kuang,G., Qiu,J., Zhang,M., Zhang,Y., Wang,M., Li,X., Kang,S., Ou,T.M. *et al.* (2018) Syntheses and evaluation of new acridone derivatives for selective binding of oncogene c-myc promoter i-motifs in gene transcriptional regulation. *Chem. Commun. (Camb.)*, **54**, 2036–2039.
19. Kaiser,C.E., Van Ert,N.A., Agrawal,P., Chawla,R., Yang,D. and Hurley,L.H. (2017) Insight into the complexity of the i-motif and G-quadruplex DNA structures formed in the KRAS promoter and subsequent drug-induced gene repression. *J. Am. Chem. Soc.*, **139**, 8522–8536.
20. Dzatko,S., Krafcikova,M., Hänsel-Hertsch,R., Fessl,T., Fiala,R., Loja,T., Krafcik,D., Mergny,J.L., Foldynova-Trantirkova,S. and Trantirek,L. (2018) Evaluation of the stability of DNA i-motifs in the nuclei of living mammalian cells. *Angew. Chem. Int. Ed Engl.*, **57**, 2165–2169.
21. Zeraati,M., Langley,D.B., Schofield,P., Moye,A.L., Rouet,R., Hughes,W.E., Bryan,T.M., Dinger,M.E. and Christ,D. (2018) I-motif DNA structures are formed in the nuclei of human cells. *Nat. Chem.*, **10**, 631–637.
22. Abou Assi,H., Garavís,M., González,C. and Damha,M.J. (2018) i-Motif DNA: structural features and significance to cell biology. *Nucleic Acids Res.*, **46**, 8038–8056.
23. Day,H.A., Huguin,C. and Waller,Z.A. (2013) Silver cations fold i-motif at neutral pH. *Chem. Commun. (Camb.)*, **49**, 7696–7688.
24. Rajendran,A., Nakano,S. and Sugimoto,N. (2010) Molecular crowding of the cosolutes induces an intramolecular i-motif structure of triplet repeat DNA oligomers at neutral pH. *Chem. Commun.*, **46**, 1299–1301.
25. Bhavsar-Jog,Y.P., Van Dornshuld,E., Brooks,T.A., Tschumper,G.S. and Wadkins,R.M. (2014) Epigenetic modification, dehydration, and molecular crowding effects on the thermodynamics of i-motif structure formation from C-rich DNA. *Biochemistry*, **53**, 1586–1594.
26. Pramanik,S., Nagatoishi,S. and Sugimoto,N. (2012) DNA tetraplex structure formation from human telomeric repeat motif (TTAGGG):(CCCTAA) in nanocavity water pools of reverse micelles. *Chem. Commun.*, **48**, 4815–4817.
27. Nguyen,T., Fraire,C. and Sheardy,R.D. (2017) Linking pH, temperature, and K(+) concentration for DNA i-motif formation. *J. Phys. Chem. B*, **121**, 7872–7877.
28. Saxena,S., Joshi,S., Shankaraswamy,J., Tyagi,S. and Kukreti,S. (2017) Magnesium and molecular crowding of the cosolutes stabilize the i-motif structure at physiological pH. *Biopolymers*, **107**, e23018.
29. Chen,Y., Qu,K., Zhao,C., Wu,L., Ren,J., Wang,J. and Qu,X. (2012) Insights into the biomedical effects of carboxylated single-wall carbon nanotubes on telomerase and telomeres. *Nat. Commun.*, **3**, 1074.
30. Spence,P., Fielden,J. and Waller,Z.A.E. (2020) Beyond solvent exclusion: i-motif detecting capability and an alternative DNA light-switching mechanism in a ruthenium(II) polypyridyl complex. *J. Am. Chem. Soc.*, **142**, 13856–13866.

31. Gurung,S.P., Schwarz,C., Hall,J.P., Cardin,C.J. and Brazier,J.A. (2015) The importance of loop length on the stability of i-motif structures. *Chem. Commun. (Camb.)*, **51**, 5630–5632.
32. McKim,M., Buxton,A., Johnson,C., Metz,A. and Sheardy,R.D. (2016) Loop sequence context influences the formation and stability of the i-motif for DNA oligomers of Sequence $(CCCXXX)_4$, where X = A and/or T, under slightly acidic conditions. *J. Phys. Chem. B*, **120**, 7652–7661.
33. Mergny,J.-L., Lacroix,L., Han,X., Leroy,J.-L. and Helene,C. (1995) Intramolecular folding of pyrimidine oligodeoxynucleotides into an i-DNA motif. *J. Am. Chem. Soc.*, **117**, 8887–8898.
34. Wright,E.P., Huppert,J.L. and Waller,Z.A.E. (2017) Identification of multiple genomic DNA sequences which form i-motif structures at neutral pH. *Nucleic Acids Res.*, **45**, 2951–2959.
35. Iaccarino,N., Cheng,M., Qiu,D., Pagano,B., Amato,J., Di Porzio,A., Zhou,J., Randazzo,A. and Mergny,J.L. (2021) Effects of Sequence and Base Composition on the CD and TDS Profiles of i-DNA. *Angew. Chem. Int. Ed Engl.*, **60**, 10295–10303.
36. Skolakova,P., Renciuk,D., Palacky,J., Krafcik,D., Dvorakova,Z., Kejnovska,I., Bednarova,K. and Vorlickova,M. (2019) Systematic investigation of sequence requirements for DNA i-motif formation. *Nucleic Acids Res.*, **47**, 2177–2189.
37. Fleming,A.M., Ding,Y., Rogers,R.A., Zhu,J., Zhu,J., Burton,A.D., Carlisle,C.B. and Burrows,C.J. (2017) 4n-1 Is a "Sweet Spot" in DNA i-Motif Folding of 2 '-Deoxycytidine Homopolymers. *J. Am. Chem. Soc.*, **139**, 4682–4689.
38. Martella,M., Pichiorri,F., Chikhale,R.V., Abdelhamid,M.A.S., Waller,Z.A.E. and Smith,S.S. (2022) i-Motif formation and spontaneous deletions in human cells. *Nucleic Acids Res.*, **50**, 3445–3455.
39. Rogers,R.A., Fleming,A.M. and Burrows,C.J. (2018) Rapid screen of potential i-motif forming sequences in DNA repair gene promoters. *ACS Omega*, **3**, 9630–9635.
40. Rogers,R.A., Fleming,A.M. and Burrows,C.J. (2018) Unusual isothermal hysteresis in DNA i-motif pH transitions: a study of the RAD17 promoter sequence. *Biophys. J.*, **114**, 1804–1815.
41. Kejnovska,I., Renciuk,D., Palacky,J. and Vorlickova,M. (2019) In: Yang,D. and Lin,C. (eds). *G-Quadruplex Nucleic Acids: Methods and Protocols*, Vol. **2035**, pp. 25–44.
42. Mergny,J.L. and Lacroix,L. (2003) Analysis of thermal melting curves. *Oligonucleotides*, **13**, 515–537.
43. Amato,J., D'Aria,F., Marzano,S., Iaccarino,N., Randazzo,A., Giancola,C. and Pagano,B. (2021) On the thermodynamics of folding of an i-motif DNA in solution under favorable conditions. *Phys. Chem. Chem. Phys.*, **23**, 15030–15037.
44. Rougée,M., Faucon,B., Mergny,J.L., Barcelo,F., Giovannangeli,C., Garestier,T. and Hélène,C. (1992) Kinetics and thermodynamics of triple-helix formation: effects of ionic strength and mismatches. *Biochemistry*, **31**, 9269–9278.
45. Dvorakova,Z., Renciuk,D., Kejnovska,I., Skolakova,P., Bednarova,K., Sagi,J. and Vorlickova,M. (2018) i-Motif of cytosine-rich human telomere DNA fragments containing natural base lesions. *Nucleic Acids Res.*, **46**, 1624–1634.
46. Mergny,J.L. and Lacroix,L. (1998) Kinetics and thermodynamics of i-DNA formation: phosphodiester versus modified oligodeoxynucleotides. *Nucleic Acids Res.*, **26**, 4797–4803.
47. Wu,S., Wang,X., Ye,X. and Zhang,G. (2013) pH-Induced conformational change and dimerization of DNA chains investigated by analytical ultracentrifugation. *J. Phys. Chem. B*, **117**, 11541–11547.
48. Niu,K., Zhang,X., Deng,H., Wu,F., Ren,Y., Xiang,H., Zheng,S., Liu,L., Huang,L., Zeng,B. *et al.* (2018) BmILF and i-motif structure are involved in transcriptional regulation of BmPOUM2 in Bombyx mori. *Nucleic Acids Res.*, **46**, 1710–1723.
49. Takahashi,S., Brazier,J.A. and Sugimoto,N. (2017) Topological impact of noncanonical DNA structures on Klenow fragment of DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, 9605–9610.
50. Casey,J.R., Grinstein,S. and Orlowski,J. (2010) Sensors and regulators of intracellular pH. *Nat. Rev. Mol. Cell Biol.*, **11**, 50–61.
51. Karpova,T.S., Kim,M.J., Spriet,C., Nalley,K., Stasevich,T.J., Kherrouche,Z., Heliot,L. and McNally,J.G. (2008) Concurrent fast and slow cycling of a transcriptional activator at an endogenous promoter. *Science*, **319**, 466–469.
52. Morisaki,T., Müller,W.G., Golob,N., Mazza,D. and McNally,J.G. (2014) Single-molecule analysis of transcription factor binding at transcription sites in live cells. *Nat. Commun.*, **5**, 4456.
53. Ben-Ari,Y., Brody,Y., Kinor,N., Mor,A., Tsukamoto,T., Spector,D.L., Singer,R.H. and Shav-Tal,Y. (2010) The life of an mRNA in space and time. *J. Cell Sci.*, **123**, 1761–1774.