CHEMISTRY
A European Journal

A Journal of
ChemPubSoc Europe

Accepted Article

Title: Probing synergistic effects of DNA methylation and 2'-β-fluorination on i-motif stability

Authors: Hala Abou Assi, Yu Chen Lin, Israel Serrano, Carlos Gonzalez, and Masad J. Damha

To be cited as: *Chem. Eur. J.* 10.1002/chem.201704591

Link to VoR: http://dx.doi.org/10.1002/chem.201704591

Supported by
ACES

WILEY-VCH

10.1002/chem.201704591

WILEY-VCH

Probing synergistic effects of DNA methylation and 2'-β-fluorination on i-motif stability

Hala Abou Assi,[a] Yu Chen Lin,[a] Israel Serrano,[b] Carlos González,*[b] and Masad J. Damha*[a]

Abstract: The possible role of DNA i-motif structures in telomere biology and in the transcriptional regulation of oncogene promoter regions is supported by several recent studies. Herein we investigate the effect of four cytidine nucleosides (and combinations thereof) on i-motif structure and stability, namely 2'-deoxycytidine (dC), 2'-deoxy-5-methyl-cytidine (5-Me-dC), 2'-deoxy-2'-fluoro-arabinocytidine (2'F-araC), and 2'-deoxy-2'-fluoro-5-methyl-arabinocytidine (5-Me-2'F-araC). The base pair 5-Me-2'F-araC:2'F-araC produced i-motifs with a $pH_{1/2}$ ("pK_a") value that closely matches physiological pH (7.34 ± 0.3). NMR analysis of the most stable telomeric sequence (HJ-2) at pH 7.0 indicated that the structure is stabilized by hybrid 5-Me-dC:2'F-araC hemiprotonated base pairs and therefore highlights the significance of the interplay between base and sugar modifications on the stability of i-motif structures.

Introduction

DNA sequences containing stretches of cytosine residues can form intercalated, quadruple-helical i-motif structures under acidic conditions. i-Motifs consist of two parallel-stranded duplexes that associate head-to-tail by $C:CH^+$ base pair intercalation.[1] Due to the pH sensitivity of i-motif structures, they have found use in the design of sensors,[2] logic circuits,[3] and nanomachines.[4] The presence and biological relevance of i-motifs in eukaryotic systems within transcriptional regulatory regions and telomeres are becoming increasingly recognized.[5-11]

Numerous strategies have been developed to promote i-motif formation at physiological pH, in order to assist in identifying ligands which can selectively interact with i-motifs *in vivo*[12] and determining the role of i-motifs in transcription regulation[13] and telomere biology.[14] Recently pursued strategies include extending the length of the C-tracts;[9,11] altering the sugar, base, or phosphate moiety;[15] utilizing branched RNA to hold the intercalated duplexes together;[16] introducing the i-motif forming sequence into a supercoiled DNA plasmid;[6] and utilizing single-walled carbon nanotubes,[17] graphene quantum dots,[18] or small molecule ligands.[19] A wide range of modified sugars has been investigated, including RNA,[20-22] arabinonucleic acid (ANA),[20] locked (LNA)[23] and unlocked (UNA) nucleic acids,[24,25] 2'F-RNA,[26] and 2'-deoxy-2'-fluoro-arabinonucleic acids (2'F-ANA).[27] Of these, only 2'F-araC has been found to significantly stabilize i-motif structures at neutral pH.[27,28]

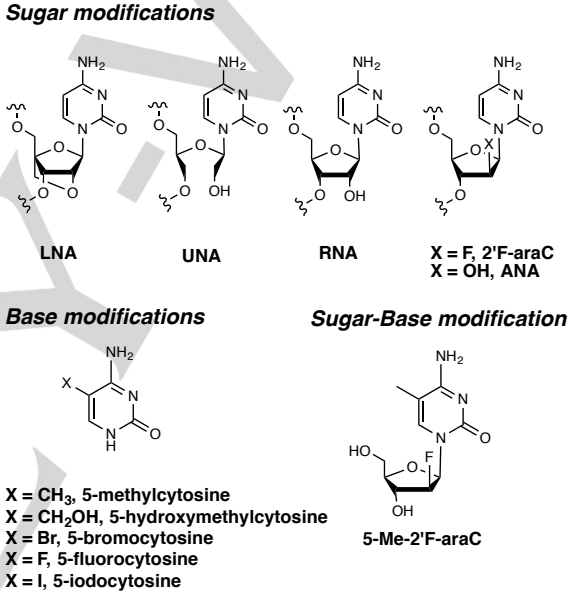


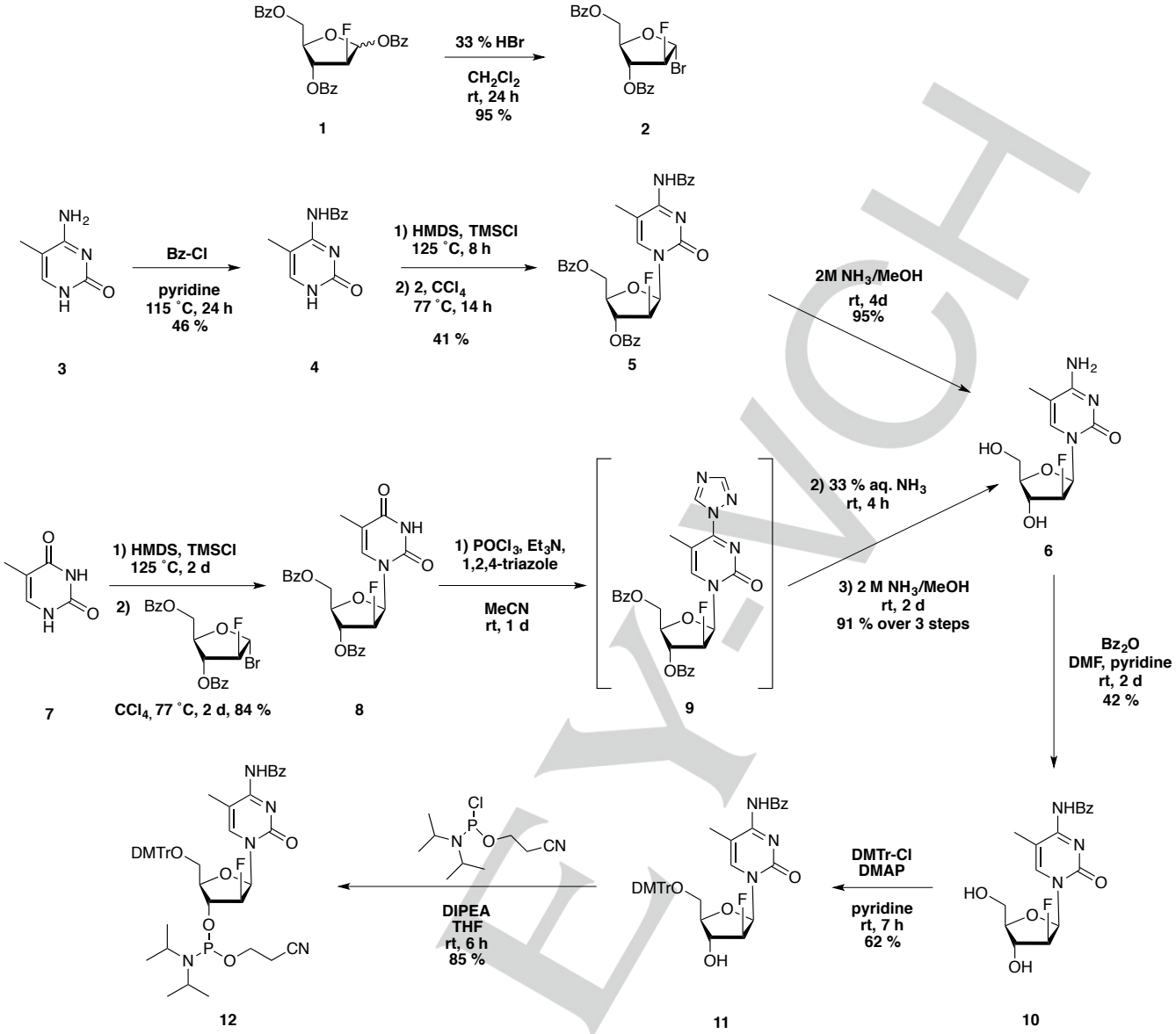
Figure 1. Base and sugar modifications previously introduced in i-motif structures.

The nucleobase modifications studied include 5-fluoro, 5-bromo, 5-iodocytosine, and 5-methylcytosine (5-Me-dC).[29] The latter stabilizes i-motif structures up to pH 6.0 ($\Delta T_{1/2}$ ≈ +1.9 °C; $\Delta pH_{1/2}$ ≈ +0.11, relative to cytosine) (Figure 1).[30] 5-Me-dC is of particular interest as C-rich regions in genomic DNA located near transcriptional start sites (TSS)[31] are susceptible to methylation and 5-hydroxymethylation (5-hMe).[32,33] In fact, 5-Me-dC and 5-hMe-dC are considered epigenetic regulators of gene expression.[34-36] Furthermore, 5-Me-dC plays a significant role in genomic imprinting, transcriptional inhibition, and gene silencing applications.[37,38] The presence of 5-Me-dC in DNA telomeric sequences affects telomere length, telomerase activity, and telomeric repeat-containing RNA (TERRA) transcription.[39,40] Therefore, investigating the effect of 5-Me-dC modification on the stability of telomeric i-motifs is of great interest.[29,30]

[a] H. Abou Assi, Y. C. Lin, Prof. M. J. Damha
Department of Chemistry, McGill University, 801 Sherbrooke St. West, Montreal, QC H3A 0B8, Canada
E-mail: masad.damha@mcgill.ca

[b] I. Serrano, Prof. C. González
Instituto de Química Física 'Rocasolano', CSIC
Serrano 119, 28006 Madrid (Spain)
E-mail: cgonzalez@iqfr.csic.es

Supporting information for this article is given via a link at the end of the document.



Scheme 1. Scheme outlining the synthesis of the modified phosphoramidite **12**.

Given the known stabilizing effect of 5-methylcytosine and 2'F-arabinose in i-motifs, an aim of this work was to synthesize 2'-deoxy-2'-fluoro-5-methyl-arabinocytidine (5-Me-2'F-araC) and study its impact on i-motif structure and stability. A second aim was to assess potential synergistic effects of three cytidine derivatives (5-Me-2'F-araC (**fC**^{m}), 2'F-araC (**fC**), and 5-Me-dC (**C**^{m})) and identify combinations of these that provide optimal stability at physiological pH.

Results and Discussion

Synthesis of 2'-deoxy-2'-fluoro-5-methyl-arabinocytidine phosphoramidite

The synthesis of the novel nucleoside analogue, 5-Me-2'F-araC, was carried out as shown in Scheme 1. The brominated sugar precursor **2** was prepared from the readily available compound **1** following previously published protocols.[41] The first synthetic route we pursued required benzoyl protection of 5-methylcytosine (**3**) to afford **4** in 46% yield. Glycosylation was accomplished by silylation of the nucleobase **4**, followed by coupling with the brominated sugar in CCl_4, resulting in an α/β-anomeric mixture. The anomeric mixture was separated by column chromatography to afford the fully protected β-nucleoside **5** in 41% yield.

To overcome the challenge of separating α/β-anomeric mixtures, an alternative synthetic route was devised. Thus, glycosylation of thymine (**7**) with the brominated sugar precursor **2** afforded the protected β-thymidine nucleoside (**8**) with anomeric purity and high yields (Scheme 1). Compound **8** was converted to 5-Me-2'F-araC (**6**) using the triazole intermediate method followed by removal of the benzoyl groups (91% yield over three steps).

Table 1. UV-melting[a] and $pH_{1/2}$[b] data of the modified oligonucleotide sequences.

Code	Sequence (5' - 3')	$T_{1/2}$ 5.0	$\Delta T_{1/2}$ 5.0	$T_{1/2}$ 7.0	$\Delta T_{1/2}$ 7.0	$T_{1/2}$ 7.4	Oligo $pH_{1/2}$	$\Delta pH_{1/2}$	Base pairing
HT-0	CCC TAA CCC TAA CCC TAA CCC	59.2 ±0.2	-	15.2 ±0.6	-	-	6.55 ±0.3	-	C:CH$^+$
HT-2	C**fCfC** TAA C**fCfC** TAA C**fCfC** TAA C**fCfC**	74.1 ±0.5	+14.9	32.1 ±0.3	+16.9	27.3 ±0.2	7.27 ±0.4	+0.72	**fC:fC**H$^+$
HTM-1	C**C^m**CTAA C**C^m**CTAA C**C^m**CTAAC**C^m**C	64.5	+5.3	19.9	+4.7	n.d.	n.d.	n.d.	**C^m:C^m**H$^+$
HTM-2	C**C^{m}C^m** TAA C**C^{m}C^m** TAA C**C^{m}C^m** TAA C**C^{m}C^m**	64.8 ±0.2	+5.6	31.0 ±0.1	+15.8	27.6 ±0.1	6.99 ±0.3	+0.44	**C^m:C^m**H$^+$
HTM-3	**C^{m}C^{m}C** TAA **C^{m}C^{m}C** TAA **C^{m}C^{m}C** TAA **C^{m}C^{m}C**	63.6	+4.4	17.7	+2.5	n.d.	n.d.	n.d.	**C^m:C^m**H$^+$
HTM-4	**C^{m}C^{m}C^m** TAA **C^{m}C^{m}C^m** TAA **C^{m}C^{m}C^m** TAA **C^{m}C^{m}C^m**	66.2	+7.0	22.7	+7.5	n.d.	n.d.	n.d.	**C^m:C^m**H$^+$
HTFM	C**fC^{m}fC^m** TAA C**fC^{m}fC^m** TAA C**fC^{m}fC^m** TAA C**fC^{m}fC^m**	70.4 ±0.5	+11.2	30.3 ±0.2	+15.1	25.2 ±0.2	7.13 ±0.3	+0.58	**fC^m:fC^m**H$^+$
HJ-1	C**fCC^m** TAA C**C^{m}fC** TAA C**fCC^m** TAA C**C^{m}fC**	69.9 ±0.2	+10.7	31.2 ±0.2	+16.0	27.9 ±0.1	6.53 ±0.3	-0.02	**fC:fC**H$^+$ **C^m:C^m**H$^+$
HJ-2	C**C^{m}C^m** TAA C**fCfC** TAA C**fCfC** TAA C**C^{m}C^m**	75.5 ±0.2	+16.3	35.0 ±0.2	+19.8	32.2 ±0.0	7.17 ±0.2	+0.62	**fC:C^m**H$^+$
HJ-3	C**fC^{m}fC^m** TAA C**fCfC** TAA C**fCfC** TAA C**fC^{m}fC^m**	74.5 ±0.3	+15.3	33.0 ±0.1	+17.8	29.5 ±0.1	7.34 ±0.3	+0.79	**fC^m:fC**H$^+$
HJ-4	C**fC^{m}fC^m** TAA CCC TAA CCC TAA C**fC^{m}fC^m**	68.8 ±0.2	+9.6	28.7 ±0.6	+13.5	25.3 ±0.3	6.70 ±0.1	+0.15	**fC^m**:CH$^+$

[a] Oligonucleotide concentration: 4 μM (single strands). $T_{1/2}$ data were calculated from UV-visible spectroscopy thermal denaturation profiles where $T_{1/2}$ corresponds to the midpoint of the dissociation transition obtained at 0.5 °C/min. $\Delta T_{1/2}$ values were calculated for the main melting transitions relative to the respective unmodified strands. (-) indicates sequences for which no melting transition was detected, or an absence of $\Delta T_{1/2}$ and (n.d.) implies not determined. [b] The pK_a of the i-motif structures is referred to as $pH_{1/2}$, defined as the pH at which 50% of the structures are folded into i-motif. The $pH_{1/2}$ values for i-motif structures were calculated from the plot of molar ellipticity at 284 nm at different pH values in 10 mM NaP_i buffer at 5 °C. Oligonucleotide concentration: 50 μM (single strands). CD *versus* pH data were fit to a standard titration model involving a single protonation event using $CD_{obs} = CD_{high} + (CD_{low} - CD_{high})/(1 + 10^{pH-pKa})$. Nucleotide codes: **fC**: 2'F-araC, **C^m**: 5-Me-dC, **fC^m**: 5-Me-2'F-araC; and C, T, and A: dC/A/T).

The NOESY NMR spectra of **6** showed a clear correlation between H-1' and H-4' of the sugar ring and between H-5' of the sugar and H-6 of the nucleobase, as expected for the β-anomer (Figure S1). Selective benzoylation of the N^4 position gave **10**, which upon 5'-dimethoxytritylation and 3'-phosphitylation afforded the desired 5-Me-2'F-araC phosphoramidite derivative **12** as a mixture of diastereomers (^{31}P NMR, Figure S2). The long-range ^{19}F-^{31}P coupling was consistent with the "W"-like conformation adopted by the F2'-C2'-C3'-O3'-P3' framework in a C2'-*endo* arabinose pucker.[42]

Effect of nucleobase and sugar modifications on i-motif formation

A library of eleven oligonucleotides (ONs) was synthesized; their sequences and properties are listed in Tables 1. i-Motif formation and stability were investigated through circular dichroism (CD), melting temperature ($T_{1/2}$) experiments, and in the case of HJ-2, structural determination *via* nuclear magnetic resonance (NMR). The $T_{1/2}$ and $pH_{1/2}$ values of the studied ONs were very close to the physiological values. We will refer to T_m as $T_{1/2}$ and pH_T as $pH_{1/2}$ since the kinetics of i-motif folding/unfolding is slower than that of normal duplexes and therefore the melting transition observed is for a system that is not at equilibrium.

First, 5-Me-dC was incorporated in the human telomeric sequence at the same positions as previously reported for 2'F-araC: four incorporations in HTM-1, eight in HTM-2 and HTM-3, and twelve in HTM-4 (Table 1).[27] Recently, Xu *et al.* incorporated one, two (forming a base pair), and three consecutive 5-Me-dC units in a 24-nt human telomeric repeat.[30] Their studies revealed that one or two 5-Me-dC inserts stabilize DNA i-motif structures, while three consecutive modifications were destabilizing. Interestingly, in our studies, HTM-4, which contains four stretches of three consecutive 5-Me-dC modifications, was more stable than HT-0 at pH 5.0 ($\Delta T_{1/2}$ = +7.0 °C, Table 1) and at pH 7.0 ($\Delta T_{1/2}$ = +7.5 °C). The highest $T_{1/2}$ values at pH 7.0 were obtained for sequence HTM-2 ($T_{1/2}$ = 31.0 °C, Figure S3) with eight 5-Me-dC incorporations.

CD experiments at pH 5.0 confirmed i-motif formation, exhibiting characteristic positive and negative CD bands at around 285 nm and 255-260 nm, respectively (Figure S4). At pH 7.0, the strands with modified cytosine inserts retained the characteristic i-motif signature. By contrast, the control strand (HT-0) no longer showed the characteristic positive i-motif signature at pH 7.0, as expected. This stabilization has been ascribed to the hydrophobicity and bulkiness of the 5-methyl group, which consequently leads to reduced flexibility of the

methylated i-motif structures.[30] Moreover, Yang *et al*. showed that 5-methylcytosine stabilizes DNA i-motif conformations by increasing the base-pairing energies (BPEs) relative to cytosine.[43] From an electronic perspective, this increase in BPE is expected since the electron-donating methyl group stabilizes the positive charge, strengthens the base-pairing interaction, and therefore stabilizes the methylated i-motif structure.

Based on these results, 5-Me-2'F-araC was incorporated at the same positions as HTM-2 to provide sequence HTFM (Table 1). HTFM was more stable than HT-0 ($T_{1/2}$ = 70.4 °C *vs*. 59.2 °C, pH 5.0) and HTM-2 ($T_{1/2}$ = 64.8 °C) at all studied pHs; however, HTFM exhibited a lower $T_{1/2}$ value relative to the 2'F-araC modified sequence (HT-2). HTFM still maintained significant stability at neutral pH ($T_{1/2}$ 30.2 *vs.* 32.1 °C). At pH 7.0 and 7.4, HT-2, HTM-2, and HTFM were at least 15 °C more stable than the control sequence (Table 1 and Figure S3).

Next, we investigated the effect of combining 5-Me-dC and 2'F-araC modifications in the same sequence (not in the same nucleotide). It is interesting to note that while HJ-1 and HJ-2 each comprise four 5-Me-dC inserts and four 2'F-araC inserts, HJ-2 exhibited significantly higher stability at all studied values of pH. Based on the NMR studies described below, we hypothesize that the positional effect observed in these sequences arises from the nature of their base pairs (**fC**:**fC**H^+ / $\mathbf{C}^m$:$\mathbf{C}^m H^+$ in HJ-1 *versus* **fC**:$\mathbf{C}^m H^+$ in HJ-2, Table 1). HJ-3, containing four 2'F-araC and four 5-Me-2'F-araC inserts, exhibited very similar stability to HT-2 and HJ-2 at pH 5.0. At pH 7.0 and pH 7.4, the stability trend observed was HJ-2 > HJ-3 > HT-2. Sequence HJ-4 (containing only four 5-Me-2'F-araC modifications) exhibited the lowest $T_{1/2}$ value; nevertheless, HJ-4 was significantly more stable than the control sequence HT-0 in the pH 5.0-7.4 range.

Effect of chemical modifications on the $pH_{1/2}$ of i-motif structures

Next we were interested in determining the pK_a values of the newly synthesized nucleoside **6** and the $pH_{1/2}$ for some of the ON sequences in order to investigate the effect of combined modifications on the $pH_{1/2}$ of i-motif structures. The pK_a values of 5-Me-2'F-araC (pK_a = 4.2) and the other cytidine derivatives (dC (4.4), 5-Me-dC (4.6), and 2'F-araC (3.9)) were determined by monitoring the change in absorbance at 280 nm as a function of pH (Figure S5 and Table S1). As expected, 2'F-araC exhibited the lowest pK_a value (3.9) due to the electronegative nature of the fluorine atom and 5-Me-dC exhibited the highest pK_a value (4.6) due to the presence of the electron-donating methyl group. The $pH_{1/2}$ values for i-motif structures were calculated from the plots of molar ellipticity at 284 nm at different pH values (Figure S6). The $pH_{1/2}$ data of the i-motif structures generally correlated well with $T_{1/2}$ data (Table 1; Figure S7). With the exception of HJ-1, all of the modified sequences exhibited a higher $pH_{1/2}$ value with respect to HT-0. Consistent with the $T_{1/2}$ data, the $pH_{1/2}$ value of HJ-2 was greater than that of HJ-1 (7.17 *versus* 6.53, Table 1). HJ-3 exhibited the highest $pH_{1/2}$ value of 7.34, which falls between the intracellular (7.2) and extracellular (7.4) pH values in normal differentiated adult cells and the intracellular pH value in cancerous cells (7.4).[44]

Structural determination via NMR

To obtain further insight on the stabilizing effect of the 2'F-araC / 5-Me-dC combination in i-motif structures, an NMR study of HJ-2 was undertaken. ^{1}H-NMR melting temperature experiments confirmed the high stability of HJ-2, with cytosine imino and amino protons observed at 15-16 and 8.5-10.5 ppm, respectively, at exceptionally high temperatures at neutral pH (45 °C, Figure 2A). Based on the NOE assignments (Figures 2B and S8), the imino proton signals observed at 45 °C correspond to the **fC**:$\mathbf{C}^m H^+$ base pairs confirming that the stability of this specific structure, HJ-2, is due to 2'F-araC base pairing with 5-Me-dC. The presence of 5-Me-dC and 2'F-araC substitutions improved the chemical shift dispersion in comparison with the unmodified sequence. The two-dimensional NMR spectra of HJ-2 exhibited all the NOE patterns characteristic of an i-motif structure (Figure S8 and S9).

Resonances of cytosine nucleobases were identified by their H5-H6 TOCSY cross-peaks and H5-amino NOEs. A similar pattern was observed for 5-Me-dC (i.e. Me5-H6 TOCSY cross-peak and Me5-amino NOEs). Nucleobase spin systems were connected with their own sugars through H6-H1' and H6-H2'/H2'' NOE cross-peaks. 2'F-araC residues were distinguished from dC by the characteristic splitting in the H1' resonances of the fluorinated sugar, due to the strong ^{1}H-^{19}F coupling. Hemiprotonated base pairs were identified by following intra-nucleotide H6-H5/Me5 → intra-nucleotide H5/Me5-amino → inter-strand amino-imino → intra-nucleotide imino-amino → intra-nucleotide amino-H5/Me5 → intra-nucleotide H5/Me5-H6 pathways in the NOESY spectra (Figure S9B). Four pathways were found connecting 5-Me-dC and 2'F-araC residues, and two connecting dC residues. This shows the formation of four **fC**:$\mathbf{C}^m H^+$ and two C:CH^+ base pairs (Figures 2B, S8, and S9).

H1'-H1' NOE cross-peaks along the minor grooves were found between 2'F-araCs on one side and between 5-Me-dCs on the other side (Figure 2B); therefore, all 2'F-araCs occupy one minor groove and base pair with the 5-Me-dCs. These connectivities, together with some weak sequential NOEs, allowed the unambiguous assignment of the residues in the central core of the i-motif. Hence, the structure of HJ-2 is stabilized by four hemiprotonated **fC**:$\mathbf{C}^m H^+$ base pairs flanked by two C:CH^+ base pairs (Figure 2). The experimental NOEs strongly indicate that the structure of HJ-2 is very similar to that of the native structure (HT-0), with both adopting a 5'E configuration.[5] Based on this similarity, a model structure of HJ-2 was built from the structure of the human telomeric C-rich strand reported by Phan and co-workers (PDB 1ELN).[5]

10.1002/chem.201704591

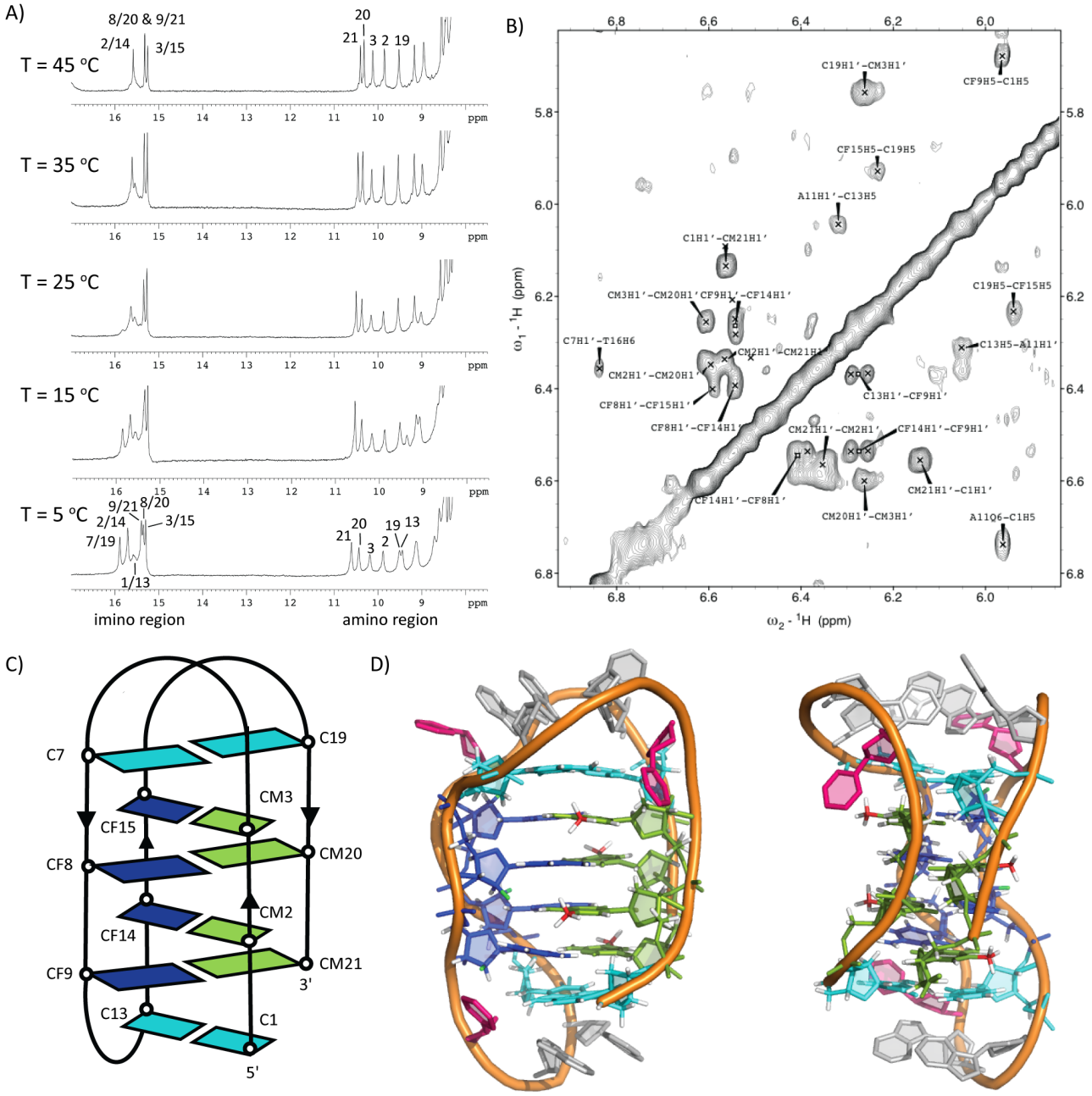


Figure 2. A) ^{1}H-NMR melting experiments of HJ-2 at pH 7.0. The observed imino and amino signals are labeled based on the assignment of the NOESY spectra (Figure S8). B) H1'-H1' region of the NOESY spectra of HJ-2 (mixing time 150 ms, T = 5 °C, pH 7.0). H1'-H1' connectivities are characteristic of i-motif structures. C) Schematic representation of HJ-2 structure with 5'E intercalation topology. D) Two views of the solution structure of HJ-2: the major groove side (left) and the 5-Me-dC minor groove side (right). Color code: dC in cyan; 5-Me-dC in green (5-Methyl group in red); 2'F-araC in blue (2'F atoms in green); dT in magenta; and dA in grey.

After performing the appropriate mutations, the coordinates were submitted to a restrained molecular dynamics calculation with the AMBER package. A number of experimental distance constraints involving protons of the central i-motif core were included. The final model is shown in Figure 2.

In spite of the overall similarity with the unmodified structure, some differences are observed: 2'F-araC sugars adopt a South conformation instead of the North conformation usually found in the native i-motif (Figure 3). Similar to our recently reported observations in the tetrameric structure of d(TCCfCfCC),[27] the arabinose 2'-fluorine is well accommodated within the major groove of the i-motif. Also, HJ-2 exhibited very similar pattern of favourable electrostatic contacts that were previously obtained in the 2'F-araC tetrameric i-motif structure (Figure 3).[27] The electronegative fluorine and the positively polarized H2' alleviate the unfavorable electrostatic interactions leading to closer

sequential (H2'-O4', H2'-O5') and inter-strand distances (H2'-O2) compared to the unmodified structure.

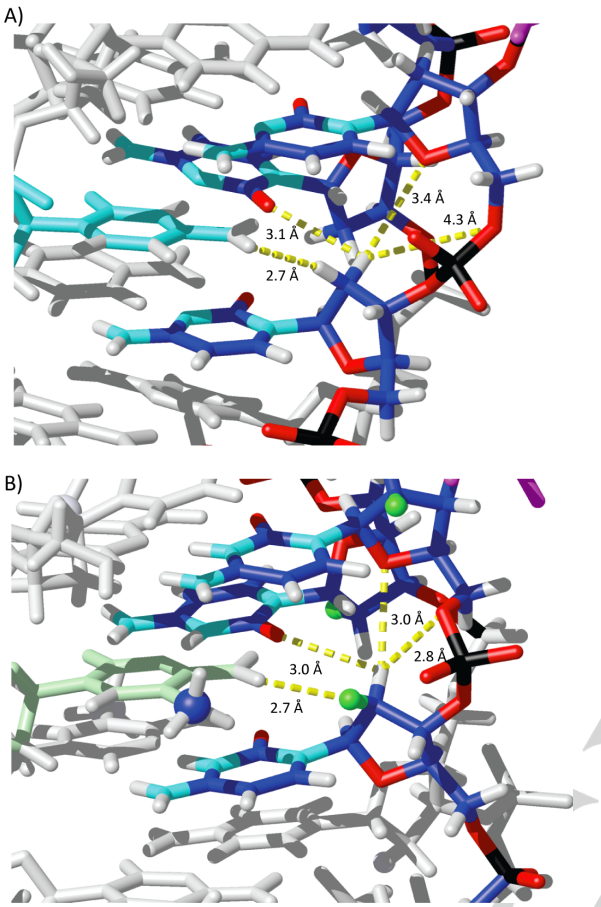


Figure 3: Details of sugar contacts in A) unmodified telomeric i-motif (PDB 1ELN) and B) 2'F-araC and 5-Me-dC substituted i-motif HJ-2. Fluorine atom is shown in green and close distances corresponding to favorable electrostatic interactions are shown in yellow.

Conclusions

In conclusion, 5-Me-2'F-araC, a new cytidine modification for use in i-motif studies was synthesized in very good yields. This nucleoside stabilizes i-motif structures when incorporated alone or in combination with other cytidine derivatives. Certain cytidine combinations afforded i-motif structures with $pH_{1/2}$ values close to physiological pH, stressing the impact of base pairing on i-motif stability. Based on the $T_{1/2}$ data obtained at pH 7.4, the following stabilization trend was observed: **fC**:$\mathbf{C}^{m}H^{+}$ > $\mathbf{fC}^{m}$:**fC**H^{+} > **fC**:**fC**H^{+} ≈ $\mathbf{C}^{m}$:$\mathbf{C}^{m}H^{+}$ > $\mathbf{fC}^{m}$:$\mathbf{fC}^{m}H^{+}$ ≈ $\mathbf{fC}^{m}$:CH^{+} > C:CH^{+}. Our studies also revealed that consecutive 5-Me-dC units stabilize i-motif structures at near physiological pH. By carefully selecting the position of the various cytidine modifications studied here, it may be possible to fine-tune the pH and temperature-dependent responses of nanodevices constructed from i-motifs.[4]

This study offers a library of chemically modified oligonucleotides that may be used to probe the biological significance of i-motifs that are prevalent in telomeric and promoter regions of the human genome. Furthermore, they may find use in the discovery and development of specific ligands or antibodies that target native i-motifs *in vivo*.

Experimental Section

General Experimental. Unless indicated otherwise, reactions were performed in oven-dried flasks with dry solvents under Argon. Solvents used were dried by MB-SPS 800 Auto (from MBraun). Proton, fluorine, and phosphorous nuclear magnetic resonance spectra (^{1}H NMR, ^{19}F NMR, ^{31}P NMR) were acquired on the Varian Mercury 400 MHz and the Bruker AV500 spectrometer. Carbon magnetic resonance spectra (^{13}C NMR) were acquired on the Bruker AV500. Data were processed by the TopSpin or MestReNova software. Chemical shifts (δ) are reported in parts per million (ppm), calibrated to the residual protium in the deuterated solvent. Coupling constants (J) are reported in Hz. Multiplicities are reported using the following abbreviations: *s* = singlet, *d* = doublet, *t* = triplet, *q* = quartet, *m* = multiplet (where the range of chemical shift is given). Analytical thin-plate chromatography was performed on pre-coated 200 μm layer thick silica gel TLA-R10011B-323 plates (Silicycle). Purifications by column chromatography were performed on silica gel (40-63 μm, 230-400 mesh). High-resolution mass spectra (HRMS) were obtained on a mass spectrometer under electron spray ionization (ESI) conditions. Detailed experimental procedures for all synthesized compounds along with NMR and HRMS characterization are clearly presented in the supplementary information.

Oligonucleotide Synthesis and Purification. Oligonucleotide synthesis was performed at 1 μmol scale on the Unylinker CPG solid support (from ChemGenes) on an ABI 340 DNA Synthesizer (from Applied Biosystems). Phosphoramidites thymidine (dT), deoxyadenosine (*N*-Bz) (dA), deoxycytidine (*N*-acetyl) (dC), and 5-methyl-2'-deoxycytidine (5-MeC) were used at 0.1 M concentration in acetonitrile and coupled for 200 s. 2'F-araC was used at 0.13 M concentration and coupled for 600 s. 2'F-5-M-araC was coupled for 1200 s. After the completion of synthesis, the CPG was transferred to a 1.5 mL screw-cap eppendorf where 1 mL NH_4OH was added and placed on a shaker for 48 hours at room temperature. The solution was centrifuged then decanted from the cleaved CPG. Samples were vented for 1 h, chilled in dry ice, and evaporated to dryness, where they were then purified by ion exchange HPLC on Agilent 1200 Series Instrument (from Agilent Technologies) using a Protein-Pak DEAE 5PW column (7.5 x 75 mm). The buffer system consisted of water (solution A) and 1 M aqueous lithium perchlorate (solution B) at a flow rate of 1 mL/min. The gradient was 0-60% solution B over 42 minutes at 60 °C. Samples were desalted using the NAP-25 desalting columns according to manufacturer's procedure. The extinction coefficients of 2'F-araC and 5-Me-2'F-araC were approximated to that of the unmodified sequence. Unless otherwise indicated, MilliQ water and autoclaved consumables were used. Masses were verified by ESI-MS (Table S3).

UV Melting Studies. UV thermal denaturation data were obtained on the Varian CARY 300 UV-Vis spectrophotometer (from Agilent Technologies) equipped with a Peltier temperature controller. The concentration of oligonucleotides used was 4 μM prepared in 10 mM sodium phosphate buffer (pH 5.0, pH 7.0, and pH 7.4). Concentrations were determined after quantitating the samples by UV absorbance at λ = 260 nm. Samples were heated at 90 °C for 15 minutes, slowly cooled to room temperature, then stored at 5 °C overnight before measurements were performed. Denaturation curves were acquired at 265 nm at a ramp of 0.5 °C/min. Samples were kept under a nitrogen flow at temperatures below 12 °C.

The optimal melting temperature ($T_{1/2}$) values were calculated using the first derivative method.

Circular Dichroism Studies. CD studies were performed at 5 ℃ on the JASCO J-810 spectropolarimeter (from JASCO Analytical Instruments) using a 1 mm path length cuvette. Temperature was maintained using the Peltier unit within the instrument. Spectra were recorded from 350-230 nm at a scan rate of 100 nm/min and a response time of 2.0 s with three acquisitions recorded for each spectrum. The spectra were normalized by subtraction of the background scan with the corresponding buffer. Data were smoothed using the means-movement function in the JASCO graphing software. The concentration of oligonucleotides used was 50 μM prepared in 10 mM sodium phosphate buffer (pH 5.0 and pH 7.0).

Nuclear Magnetic Resonance. Samples for NMR experiments were dissolved 9:1 H_2O/D_2O (10 mM sodium phosphate buffer. All NMR spectra were acquired in Bruker Avance spectrometers operating at 600 and 800 MHz, equipped with cryoprobes and processed with the TOPSPIN software. NOESY[45] spectra in were acquired with mixing times of 150 and 250 ms. TOCSY[46] spectra were recorded with the standard MLEV-17 spin-lock sequence and a mixing time of 80 ms. Water suppression was achieved by including a WATERGATE[47] module in the pulse sequence prior to acquisition. The spectral analysis program SPARKY was used for semiautomatic assignment of the NOESY cross-peaks and quantitative evaluation of the NOE intensities.

Structural calculations. Restrained molecular dynamics calculations were carried out with the SANDER module of the package AMBER 12.0.[48] Coordinates from the human telomeric i-motif (pdb code 1ELN) were utilized as starting points for the AMBER refinement, consisting of an annealing protocol in vacuo, followed by trajectories of 500 ps each in which explicit solvent molecules were included. The Particle Mesh Ewald method was used to evaluate long-range electrostatic interactions. The specific protocols for these calculations have been described elsewhere.[49] The BSC1 force field was used to describe the DNA,[50] and the TIP3P model was used to simulate water molecules. Suitable parameters for the 2'F-araC derivatives were extracted from Noy *et al*.[51] Analysis of the representative structures was carried out with MOLMOL program.[52]

Keywords: i-motif • telomeric DNA • 5-methylcytosine • epigenetic modifications • 2'F-ANA

[1] K. Gehring, J. L. Leroy, M. Gueron, *Nature* **1993**, *363*, 561-565.
[2] L. H. Lu, M. Wang, L. J. Liu, C. Y. Wong, C. H. Leung, D. L. Ma, *Chem. Commun.* **2015**, *51*, 9953-9956.
[3] Y. H. Shi, H. X. Sun, J. F. Xiang, H. B. Chen, Q. F. Yang, A. J. Guan, Q. Li, L. J. Yu, Y. L. Tang, *Chem. Commun.* **2014**, *50*, 15385-15388.
[4] Y. C. Dong, Z. Q. Yang, D. S. Liu, *Acc. Chem. Res.* **2014**, *47*, 1853-1860.
[5] A. T. Phan, M. Gueron, J. L. Leroy, *J. Mol. Biol.* **2000**, *299*, 123-144.
[6] D. Sun, L. H. Hurley, *J. Med. Chem.* **2009**, *52*, 2863-2874.
[7] H. J. Kang, S. Kendrick, S. M. Hecht, L. H. Hurley, *J. Am. Chem. Soc.* **2014**, *136*, 4172-4185.
[8] Y. X. Cui, D. M. Kong, C. Ghimire, C. X. Xu, H. B. Mao, *Biochemistry* **2016**, *55*, 2291-2299.
[9] E. P. Wright, J. L. Huppert, Z. A. E. Waller, *Nucleic Acids Res.* **2017**, *45*, 2951-2959.
[10] S. Takahashi, J. A. Brazier, N. Sugimoto, *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 9605-9610.
[11] A. M. Fleming, Y. Ding, R. A. Rogers, J. Zhu, J. Zhu, A. D. Burton, C. B. Carlisle, C. J. Burrows, *J. Am. Chem. Soc.* **2017**, *139*, 4682-4689.
[12] H. A. Day, P. Pavlou, Z. A. Waller, *Bioorg. Med. Chem.* **2014**, *22*, 4407-4418.
[13] T. A. Brooks, S. Kendrick, L. Hurley, *FEBS J.* **2010**, *277*, 3459-3469.
[14] Y. Chen, K. Qu, C. Zhao, L. Wu, J. Ren, J. Wang, X. Qu, *Nat. Commun.* **2012**, *3*, 1074.
[15] S. Benabou, A. Aviñó, R. Eritja, C. González, R. Gargallo, *Rsc Adv.* **2014**, *4*, 26956.
[16] S. Robidoux, R. Klinck, K. Gehring, M. J. Damha, *J. Biomol. Struct. Dyn.* **1997**, *15*, 517-527.
[17] X. Li, Y. Peng, J. Ren, X. Qu, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 19658-19663.
[18] X. Chen, X. Zhou, T. Han, J. Wu, J. Zhang, S. Guo, *ACS Nano* **2013**, *7*, 531-537.
[19] S. Kendrick, H. J. Kang, M. P. Alam, M. M. Madathil, P. Agrawal, V. Gokhale, D. Z. Yang, S. M. Hecht, L. H. Hurley, *J. Am. Chem. Soc.* **2014**, *136*, 4161-4171.
[20] S. Robidoux, M. J. Damha, *J. Biomol. Struct. Dyn.* **1997**, *15*, 529-535.
[21] D. Collin, K. Gehring, *J. Am. Chem. Soc.* **1998**, *120*, 4069-4072.
[22] K. Snoussi, S. Nonin-Lecomte, J. L. Leroy, *J. Mol. Biol.* **2001**, *309*, 139-153.
[23] N. Kumar, J. T. Nielsen, S. Maiti, M. Petersen, *Angew. Chem.* **2007**, *46*, 9220-9222.
[24] A. Pasternak, J. Wengel, *Bioorg. Med. Chem. Lett.* **2011**, *21*, 752-755.
[25] P. Perlíková, K. K. Karlsen, E. B. Pedersen, J. Wengel, *Chembiochem* **2014**, *15*, 146-156.
[26] C. P. Fenna, V. J. Wilkinson, J. R. Arnold, R. Cosstick, J. Fisher, *Chem. Commun.* **2008**, 3567-3569.
[27] H. Abou Assi, R. W. V Harkness, N. Martin-Pintado, C. J. Wilds, R. Campos-Olivas, A. K. Mittermaier, C. Gonzalez, M. J. Damha, *Nucleic Acids Res.* **2016**, *44*, 4998-5009.
[28] H. Abou Assi, R. El-Khoury, C. Gonzalez, M. J. Damha, *Nucleic Acids Res.* **2017**,doi.org/10.1093/nar/gkx838.
[29] L. Lannes, S. Halder, Y. Krishnan, H. Schwalbe, *Chembiochem* **2015**, *16*, 1647-1656.
[30] B. C. Xu, G. Devi, F. W. Shao, *Org. Biomol. Chem.* **2015**, *13*, 5646-5651.
[31] S. Balasubramanian, L. H. Hurley, S. Neidle, *Nat. Rev. Drug Discov.* **2011**, *10*, 261-275.
[32] T. A. Brooks, L. H. Hurley, *Genes Cancer* **2010**, *1*, 641-649.
[33] J. Dai, E. Hatzakis, L. H. Hurley, D. Yang, *Plos One* **2010**, *5*, e11647.
[34] S. G. Jin, S. Kadam, G. P. Pfeifer, *Nucleic Acids Res.* **2010**, *38*, e125.
[35] O. Yildirim, R. Li, J. H. Hung, P. B. Chen, X. Dong, L. S. Ee, Z. Weng, O. J. Rando, T. G. Fazzio, *Cell* **2011**, *147*, 1498-1510.
[36] T. F. Kraus, D. Globisch, M. Wagner, S. Eigenbrod, D. Widmann, M. Munzel, M. Muller, T. Pfaffeneder, B. Hackner, W. Feiden, U. Schuller, T. Carell, H. A. Kretzschmar, *Int. J. Cancer* **2012**, *131*, 1577-1590.
[37] A. Bird, *Genes & Dev.* **2002**, *16*, 6-21.
[38] A. Bird, *Nature* **2007**, *447*, 396-398.
[39] S. Gonzalo, I. Jaco, M. F. Fraga, T. P. Chen, E. Li, M. Esteller, M. A. Blasco, *Nat. Cell Biol.* **2006**, *8*, 416-U466.
[40] L. J. Ng, J. E. Cropley, H. A. Pickett, R. R. Reddel, C. M. Suter, *Nucleic Acids Res.* **2009**, *37*, 1152-1159.
[41] M. I. Elzagheid, E. Viazovkina, M. J. Damha, *Current protocols in nucleic acid chemistry / edited by Serge L. Beaucage ... [et al.]* **2002**, *Chapter 1*, Unit 1 7.
[42] C. J. Wilds, M. J. Damha, *Nucleic Acids Res.* **2000**, *28*, 3625-3635.
[43] B. Yang, R. R. Wu, M. T. Rodgers, *Anal. Chem.* **2013**, *85*, 11000-11006.
[44] B. A. Webb, M. Chimenti, M. P. Jacobson, D. L. Barber, *Nat. Rev. Cancer* **2011**, *11*, 671-677.
[45] A. Kumar, R. R. Ernst, K. Wuthrich, *Biochem. Biophys. Res. Commun.* **1980**, *95*, 1-6.
[46] A. Bax, D. G. Davis, *J. Magn. Reson.* **1985**, *65*, 355-360.
[47] M. Piotto, V. Saudek, V. Sklenar, *J. Biomol. NMR* **1992**, *2*, 661-665.
[48] D.A. Case, T.A. Darden, T.E. Cheatham, III, C.L. Simmerling, *et al*. AMBER 12, University of California, San Francisco, **2012**.
[49] R. Soliva, V. Monaco, I. Gomez-Pinto, N. J. Meeuwenoord, G. A. Van der Marel, J. H. Van Boom, C. Gonzalez, M. Orozco, *Nucleic Acids Res.* **2001**, *29*, 2973-2985.
[50] I. Ivani, P. D. Dans, A. Noy, A. Perez, I. Faustino, A. Hospital, J. Walther, P. Andrio, R. Goni, A. Balaceanu, G. Portella, F. Battistini, J. L. Gelpi, C. Gonzalez, M. Vendruscolo, C. A. Laughton, S. A. Harris, D. A. Case, M. Orozco, *Nat. Methods* **2016**, *13*, 55-58.
[51] A. Noy, F. J. Luque, M. Orozco, *J. Am. Chem. Soc.* **2008**, *130*, 3486-3496.
[52] R. Koradi, M. Billeter, K. Wuthrich, *J. Mol. Graph.* **1996**, *14*, 51.

10.1002/chem.201704591

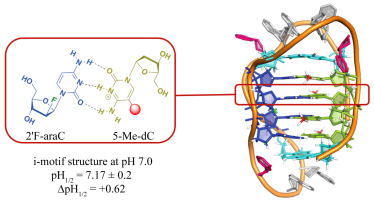
WILEY-VCH

Entry for the Table of Contents (Please choose one layout)

Layout 1:

FULL PAPER

The interplay between sugar and base modifications, in particular 5-Me-dC:2'F-araC base-pairing, leads to very stable i-motif structures at neutral pH.



Hala Abou Assi, Yu Chen Lin, Israel Serrano, Carlos González, and Masad J. Damha**

Page No. – Page No.

Probing synergistic effects of DNA methylation and 2'-β-fluorination on i-motif stability

Layout 2:

FULL PAPER

((Insert TOC Graphic here; max. widh: 11.5 cm; max. height: 2.5 cm))

Page No. – Page No.

Text for Table of Contents