RSC Advances

PAPER

View Article Online
View Journal | View Issue



Cite this: *RSC Adv.*, 2025, **15**, 37099

# Tuning i-motif folding and stability with acyclic phosphonate-linked (ZNA) cytidine residues

Marlies Derden,† Sten Reynders,† Jérôme Rihon, Xinlu Li, Frédérick Coosemans, Jef Rozenski, Piet Herdewijn, Elisabetta Groaz †* and Eveline Lescrinier †*

Received 12th August 2025
Accepted 19th September 2025

DOI: 10.1039/d5ra05934h

rsc.li/rsc-advances

i-Motifs (iMs) are non-canonical four-stranded DNA structures stabilized by hemi-protonated (C:H$^+$C) cytosine base pairs. Under non-acidic pH conditions, cytosine deprotonation disrupts these interactions, resulting in structural unfolding. To broaden the environmental conditions that support iM formation for potential biomedical applications, we investigated the effect of substituting 2′-deoxycytidine residues with either the (*R*) or (*S*) isomers of an acyclic phosphonate-linked nucleotide (ZNA) bearing a cytosine (C) or 5-fluorocytosine (5F-C) nucleobase within the 22-mer vertebrate telomeric iM. UV melting and NMR spectroscopy analyses revealed that the presence of ZNA-C in the core of the iM decreased its stability, while 5F-C altered the iM stability across different pH ranges. Molecular dynamics (MD) provided a first glimpse of the quasi-symmetrical loops 1 and 3 in the iM formed by the native vertebrate telomeric sequence. The changes in the dynamics of the loop region in the modified sequences could be correlated with their decreased stability, despite stable base pairing occurring for both (*R*) and (*S*)-ZNA modifications.

## Introduction

Cytosine-rich (C-rich) sequences in complementary DNA strands can form non-standard four-stranded structures called i-motifs (iMs) under specific conditions, both *in vitro* and *in vivo*.[1] Specifically, two parallel DNA duplexes are held together in an antiparallel orientation by intercalated, hemi-protonated C:H$^+$C base pairs.[2] This pattern leads to a compact DNA geometry, characterized by two broad major grooves and two particularly narrow minor grooves, where close phosphate–phosphate distances result in electrostatic repulsion. This effect is mitigated by the protonation of C. Destabilizing Coulomb interactions are also partially compensated by favorable hydrogen bonding between inter-stranded sugars. These sugar-to-sugar contacts, along with C:H$^+$C base pairs, facilitate the overall spatial arrangement of C-tracts as an iM.[3]

In living cells, iMs are mainly present at promoters of actively transcribing genes and within open chromatin regions, often overlapping with R-loops, with their abundance and distribution varying in a cell-type specific manner.[4] They are believed to play key roles in various DNA processes, including the regulation of genome stability, gene transcription, mRNA translation, DNA replication, telomere and centromere functions, as well as vertebrate diseases.[5] *In vitro*, the formation of iM depends not only on experimental conditions such as pH and temperature but also on the length of cytosine stretches as well as the number and types of nucleotides that link the cytosine containing residues.[6] Despite extensive research, the precise factors governing iM formation remain not fully understood. To gain deeper insight and potentially regulate iM stability and folding, various chemical modifications have been explored, including variations to the backbone and sugar moieties such as peptide nucleic acid (PNA),[7] phosphorothioate (PS),[8] 2′-fluoroarabino nucleic acid (FANA),[9] locked nucleic acid (LNA),[10] hexitol nucleic acid (HNA)[11] as well as acyclic nucleic acid analogues.[12–14] These modifications have produced varying outcomes; for instance, PNA and PS did not significantly affect iM stability, whereas LNA enhanced stability only when the modified residues were positioned at specific sites within an oligomer. Other modifications either prevented folding or led to destabilization. This variability highlights the complexity of how sugar moieties influence iM stability and emphasizes the need for a systematic investigation.

The 22 nt vertebrate telomeric fragment with the sequence 5′-d[(CCCTAA)$_3$CCCT] has been well studied *in vitro*.[15] It exhibited pH-dependent iM formation under mildly acidic conditions, dissociating at neutral pH, and demonstrated optimal stability around pH 6.0. Previous experimental studies using specific sugar modifications highlighted the importance of sugar–sugar contacts for iM stability. Depending on their position within the iM tetrameric core or loops formed by the vertebrate telomeric sequence, acyclic unlocked nucleic acids (UNA) could destabilize or stabilize the corresponding folded

*KU Leuven, Rega Institute for Medical Research, Medicinal Chemistry, Herestraat 49-Box 1041, 3000 Leuven, Belgium. E-mail: elisabetta.groaz@kuleuven.be; eveline.lescrinier@kuleuven.be*

† Equal contribution.

structure, respectively.[12] Similar effects were observed for cytidine containing acyclic threoninol nucleic acid (aTNA) residues.[13] Furthermore, the replacement of unmodified residues at selected C-core positions with aTNA allowed to modulate iM stability over different pH ranges. The introduction of a fluorine atom at the 2′-β position of 2′-deoxycytidine (dC) was found to enhance iM stability, supporting formation even at higher pH values approaching ~7.0.[16]

The pH-dependence of iM formation is an attractive feature for nanotechnology applications, particularly in the design of DNA nanodevices that use iM-forming sequences as pH-switches or components of DNA-nanomotors.[17] The pH-responsiveness of iM forming oligonucleotides holds also potential for applications in bioanalysis and targeted therapy. A pH-responsive aptamer bearing an unnatural nucleobase was designed to specifically fold through C:H$^+$C base pairing and bind target proteins in the extracellular acidic microenvironment of cancer cells to achieve selective inhibition of cancer cell migration.[18]

Understanding how different chemical modifications influence iM formation and stability is crucial for designing nucleic acid-based tools with specific functional properties. In this study, the effect of an acyclic ZNA-C modification with standard cytosine nucleobase on the iM formation ability of the vertebrate telomeric fragment 5′-d[(CCCTAA)$_3$CCCT] was investigated (Fig. 1). ZNA can exist in two enantiomeric forms, *i.e.*, (*S*)- and (*R*)-ZNA. Both isomers were considered in this work, however only the (*R*)-isomer was previously shown to be accommodated in a DNA duplex.[19] The 5-fluoro (*S*)-ZNA-C analogue was also included since a previous X-ray crystallography study suggested that formation of C:H$^+$C base pairs can be affected by a decrease of p$K_a$ of cytosine upon halogenation.[20] Moreover, the $^{19}$F in the studied modification can be exploited to probe the secondary structure of the core of the telomeric fragment, as its chemical shift is sensitive to multiple factors affecting its electronic microenvironment, including stacking, H-bonding interactions with adjacent bases, and the protonation state of the nucleobase.[21] The thermodynamic stability of the modified and unmodified sequences was studied by the UV melting method. Measurements were performed at different pH values to assess the pH-dependent stability characteristic of an iM structure. Nuclear magnetic resonance (NMR) was used to evaluate C:H$^+$C base pairing in the iM core for the modified and unmodified congeners under different pH conditions. To monitor the effect on mobility of individual atoms or residues induced by a single ZNA-C modification in the core of the iM, a molecular dynamics (MD) study was also performed.

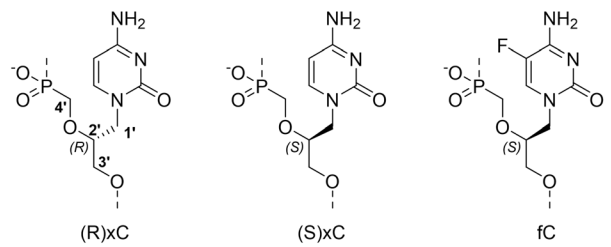


Fig. 1 Structure of the modified C residues incorporated in the vertebrate telomeric fragment.

## Materials and methods

### Synthesis of ZNA H-phosphinate building blocks

Modified (*S*)- and (*R*)-1-[2-(phosphinylmethoxy)-3-(4-monomethoxytrityl)prop-1-yl]-$N^4$-benzoylcytosine building blocks were prepared according to previously established procedures.[19,22,23] A similar protocol was used to obtain (*S*)-1-[2-(phosphinylmethoxy)-3-(4-monomethoxytrityl)prop-1-yl]-$N^4$-benzoyl-5-fluoro-cytosine from 5-fluoro-cytosine, as detailed in Scheme S1 (SI).

### Oligonucleotide synthesis

The 22 nt DNA model sequence **ON1** as well as ZNA-DNA chimeras **ON2–ON5** (Table 1) were prepared by solid-phase synthesis on an Expedite DNA synthesizer (Applied Biosystems) using a combination of an adapted H-phosphonate approach and the standard phosphoramidite method.[19,23] All oligomers were purified by HPLC and analyzed by mass spectrometry (see SI).

### UV melting studies

Oligomers were dissolved in a buffer solution containing NaCl (0.1 or 1.0 M), potassium phosphate (0.02 M, pH 7.5), and EDTA (0.1 mM). Their concentration was determined by measuring the absorbance in Milli-Q water at 260 nm and 80 °C, assuming that ZNA-C and its 5F derivative have the same extinction coefficients per base moiety in the denatured state as their natural nucleoside counterparts (A, $\varepsilon = 15\,060$; T, $\varepsilon = 8560$; C, $\varepsilon = 7100$). For all experiments, the oligonucleotide concentration was 4 μM. A Varian Cary 100 BIO spectrophotometer was used to determine the melting curves. Cuvettes were maintained at a constant temperature by water circulation through the cuvette holder. The temperature of the solution was measured with a thermistor that was directly immersed in the cuvette. Temperature control and data acquisition were carried out automatically with an IBM-compatible computer by using the Cary WinUV thermal application software. A quick heating and cooling cycle was carried out to allow for proper annealing. Denaturation curves were acquired at 260 and 270 nm (SI). The samples were then heated from 10 to 80 °C at a rate of 0.2 °C per min and were cooled again at the same speed. Experiments at different pH values ranging from pH 4.6 to 6.0 were performed. Melting temperatures ($T_m$) were calculated as the

Table 1 Composition of the synthesized oligonucleotides (modified residues are indicated in bold)

| | |
|---|---|
| **ON1** | d(CCCTAACCCTAACCCTAACCCT) |
| **ON2** | d(CCCTAAC)-(***R***)**x**C-d(CTAACCCTAACCCT) |
| **ON3** | d(CCCTAAC)-(***S***)**x**C-d(CTAACCCTAACCCT) |
| **ON4** | d(CCCTAAC)-(***S***)**f**C-d(CTAACCCTAACCCT) |
| **ON5** | d(CCCTAAC)-(***S***)**f**C-d(CTAACCCTAAC)-(***S***)**f**C-d(CT) |

midpoint of the transition using the first derivative of the experimental data; reported values are the average of two runs.

### NMR experiments

Samples for NMR experiments were dissolved in $H_2O/D_2O$ 9 : 1. Concentrations of 0.4 and 0.1 mM were used for **ON1**, while **ON2**–**ON5** were analyzed at 0.1 mM. NMR spectra were acquired on a Bruker Neo spectrometer operating at 600 Mhz with a quadruple cryoprobe ($^1$H, $^{31}$P, $^{15}$N, $^{13}$C) and processed with the Topspin software. TOCSY spectra were recorded with a standard DIPSI[24] spinlock sequence and 64 ms mixing time. NOESY spectra were acquired at 200 and 300 ms mixing times. Zero-quantum interference in two-dimensional NMR spectra was eliminated by gradients.[25] Excitation sculpting[26] was used to suppress the water signal prior to acquisition. One- and two-dimensional experiments were carried out at temperatures ranging from 0 to 40 °C.

### Molecular modeling

**Conformational sampling.** Conformational sampling was conducted on (*R*) and (*S*)-ZNA-C isomers separately, according to the accelerated methodology described by Mattelaer *et al.*[27] using the ORCA quantum mechanics software Package.[28–35] The backbone was sampled by varying the gamma and delta angles over a range of [0 to 360°] with an interval of 10°. Iterating over both axes produces a total of 1369 conformations to be evaluated. Each sampled molecule was capped at the ends of the backbone with methyl groups, where the phosphate moieties would attach in a regular oligomer. Restraints on (*R*)-ZNA were set for the $\beta$ and $\varepsilon$ angle at 180°, while the $\chi_\alpha$ angle was restrained at 155.2° (see nomenclature Fig. S3C, SI). This value was calculated as the average value of $\chi_\alpha$ in the NMR solution structures of the (*R*)-ZNA fragment (PDB ID: [9F3E]).[19] For (*S*)-ZNA the same restraints were applied for $\beta$ and $\varepsilon$, while the $\chi_\alpha$ angle was restrained at 95.2° (see nomenclature Fig. S3E, SI). A second sampling was run over the ($\chi_\alpha$, $\chi_\beta$) dyad, also over a range of [0 to 360°] with an interval of 10°. The $\beta$ and $\varepsilon$ angles were kept at the same restraint, while the ($\gamma$, $\delta$) dyad was set at (70°, 280°) and (306°, 311°) for (*R*)- and (*S*)-ZNA, respectively, which was a result of the global minimum of the previous conformational sampling experiment.

**Parametrization.** The hemicytosine ($H^+C$) charge parameters, required for the iM simulations, were retrieved from Panczyk *et al.*[36] The charges were derived by implementing the Merz–Kollman (MK) population analysis scheme,[37] written for ORCA.[38] Only an internal fragment residue of the (*S*)- and (*R*)-ZNA-C chemistry was finalized. This was carried out through a multimolecular and conformational fit of dimethylphosphonate and ZNA molecules. These were subjected to the MK population analysis scheme to derive the Electrostatic Potential (ESP) and further optimized with the restraint ESP (RESP) procedure.[39] The oxygens in the phosphate and hydrogens in the amine of the cytosine ring were equivalenced in the first step. Buried carbons (in methyl and methylene moieties) were restrained in a second step. The conformation of the following ($\gamma$, $\delta$) dyad angles were used for the (*R*)-ZNA parametrization: (70°, 60°), (70°, 280°), (310°, 80°) and (320°, 300°), while for the (*S*)-ZNA parametrization the following conformations were used (310°, 310°), (270°, 70°), (300°, 170°) and (200°, 300°). For both stereoisomers, these conformations were based on the various minima of the conformational sampling experiment (Fig. S3A and D, SI). The dimethylphosphonate data were obtained from a previous project on xylose phosphonate nucleoside parametrization.[40]

**Torsion angle fitting.** Based on the results of the conformational sampling experiment, only the $\chi_\alpha$ dihedral angle was selected for fitting (Fig. S4A, SI) along with a restrained $\chi_\beta$ value of 70°. This was performed in-house using scripted methods. The fit was produced using the AMBER's[41] interpretation of force fields (eqn (1), where $K_v$ is the force constant, $\gamma$ the phase of the function, and $n$ the multiplicity).

$$E_{\text{torsion}} = \sum_{\text{dihedrals}} K_v[1 + \cos(n\phi - \gamma)] \quad (1)$$

Ultimately, two functions were necessary to form a linear combination of the resulting profile of $\chi_\alpha$ at $\chi_\beta = 70°$. The following parameters ($K_v$, $\gamma$, $n$) were used for the parametrization of $\chi_\alpha$: [$K_v$: 2.8, $\gamma$: 355, $n$: 2.05] and [$K_v$: 4.25, $\gamma$: 230, $n$: 0.57]. A graph (Fig. S4B, SI) was provided that represents the linear combination of both analytic functions to model the behavior of this specific angle.

**Molecular dynamics simulation.** The initial iM structure was retrieved from the RCSB,[42] PDB ID [1EL2].[43] The residue at position 8 was manually changed to an (*R*)- or (*S*)-ZNA cytosine residue to generate a modified duplex. Both models and the solution NMR structure were simulated under the same conditions. The MD simulations were run using the AMBER18 software package[41] joined with AMBERTools19, employing the Particle Mesh Ewald (PMEMD) simulation engine.[44] The DNA.OL15 force field was used, together with custom parameters to aptly model the simulated molecules.[45] The TIP3P water model was used for the explicit solvation, in a truncated octahedron box, and the charges were neutralized with sodium ions. A cut-off distance of 12 Å was used for non-bonded interactions. The minimization ran for a total of 30 000 cycles, with the first 22 500 cycles employing the steepest descent method and the last 7500 the conjugate gradient method. The SHAKE algorithm[46] was employed to allow a time step of 2 fs. An initial heating was performed for 50 ps, from 0 to 100 K with vlimit set to 15. The rest of the heating, from 100 to 300 K, ran for an additional 50 ps. Density and equilibration ran for 100 ps each, with density set at 1 g $mL^{-1}$. The Langevin thermostat[47] and the Berendsen barostat[48] were used to keep the temperature and density at a constant value. The production simulations were run for 1000 ns, with a restart every 1 ns to vary the random seed of the simulation. MD simulations were run on an NVIDIA GeForce RTX 2070 GPU through the cuda accelerated simulation engine.[49] Cpptraj was used to extract the modelled data.[50] Visualization was performed using UCSF Chimera.[51] Graphs were made using the Matplotlib python module.

## Results and discussion

### Effect of ZNA modifications on the thermal and pH stability of folded oligomers

The thermodynamic stability of folded oligomers **ON1–5** was investigated by using the UV melting method. To evaluate the pH-dependence for all synthesized oligonucleotides, measurements were performed in buffers with varying pH values (4.6 to 6.0). For the unmodified sequence **ON1**, UV melting experiments revealed a systematic increase in $T_m$ as the pH decreased within the considered pH range (Table 2). The introduction of a single ZNA-C at position 8 of **ON1** significantly reduced the stability of its folded structure, with no observable $T_m$ above pH 5.5. A slightly greater destabilizing effect was observed for the (*R*)-ZNA containing oligomer **ON2** compared to the (*S*)-ZNA counterpart **ON3**. A similar effect was previously reported when a single acyclic threoninol cytidine derivative was introduced into the core of the folded vertebrate telomer sequence, which decreased the stability of the iM and lowered the pH required to form a stable iM by approximately 0.54 units.[13] The fluorination of the nucleobase in the (*S*)-ZNA residue further lowered the pH required to obtain a detectable melting curve for **ON4** at 260 and 270 nm. Moreover, the introduction of a second modification with the aim to study an (*S*)-ZNA C:H$^+$C base pair prevented folding, even at pH 4.6.

### Effects of pH and temperature on iM structure

To obtain structural information on the folded iM structures, we performed 1D NMR experiments on the full DNA sequence **ON1** as well as modified oligonucleotides **ON2–ON5** (Fig. 2 and 3). Within the studied pH range (5.20–6.25), the $^1$H NMR spectra of the 0.4 mM **ON1** sample displayed signals in the 15–16 ppm region throughout the entire temperature range examined (0–40 °C), which are characteristic of imino protons involved in C:H$^+$C base pairs within iM structures.[1] At low temperature (0 °C), additional imino signals appeared at 11.0 and 13.3 ppm (Fig. 2). According to literature reports, the signal at 11.0 ppm can be assigned to the T10::T22 base pair,[43] while the signal at 13.3 ppm with a strong NOE interaction to A:H2 is characteristic for a Watson–Crick A::T base pair.[52] The same A:H2 signal is also in NOE contact with the C:H$^+$C imino signal at 15.4 ppm (Fig. S5, SI). The latter NOE interactions are consistent with a bimolecular iM structure for the studied sequence with A5::T16 base pairs flanking two separated iMs (Fig. 2B). The co-existence of a monomeric and dimeric iM for the studied sequence was previously determined by gel electrophoresis analysis.[53]

Table 2 $T_m$ (°C) of **ON1–5** at different pH values[a]

| | 4.60 | 5.10 | 5.25 | 5.50 | 5.75 | 6.00 |
|---|---|---|---|---|---|---|
| **ON1** | 47.0 | 44.2 | 45.1 | 38.8 | 34.4 | 30.6 |
| **ON2** | 36.4 | 33.0 | 27.9 | 26.7 | na | na |
| **ON3** | 38.5 | 34.1 | 29.7 | 28.0 | na | na |
| **ON4** | 32.9 | 29.3 | na | na | na | na |
| **ON5** | na | na | na | na | na | na |

[a] na: no observable $T_m$.

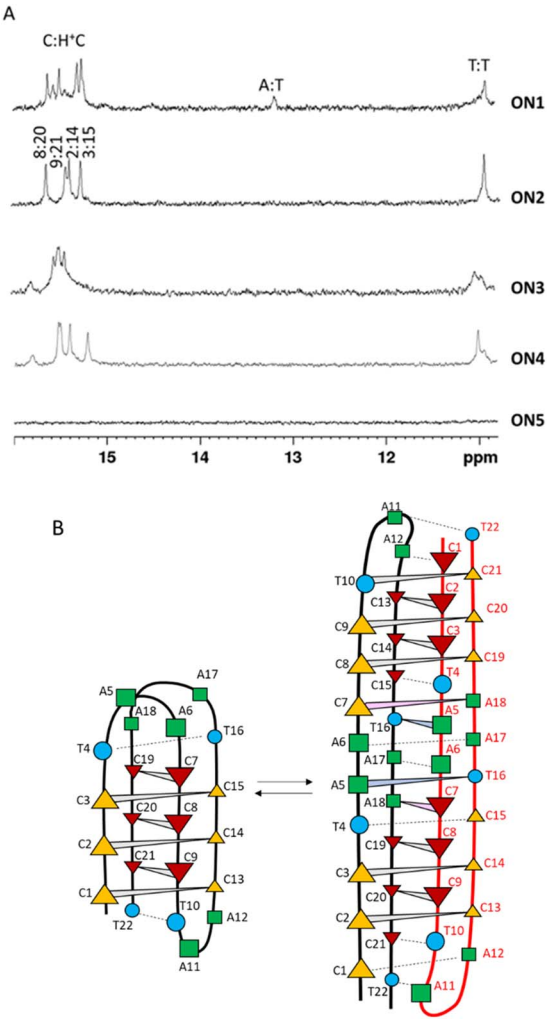


Fig. 2 (A) Imino proton signals in the 1D proton spectra of **ON1–5** at pH = 5.2 (0 °C). (B) Proposed equilibrium of the folded single-stranded iM and dimeric structure containing A5::T16 base pairs.

An equilibrium between 5′E and 3′E conformers of the monomeric iM was also observed for a sequence that lacked T22. The additional T10::T22 in **ON1** is thought to stabilize the 5′E conformation of the iM, while the presence of a Watson–Crick A::T base pair rules out a monomeric 3′E iM as the minor conformer in the studied sequence.[2] The disappearance of the minor form at lower concentration further supports its dimeric nature.

In the pH range between 5.20 and 5.50, the $^1$H NMR spectra of **ON2–3** at 0 °C contained one sharp signal at 11.0 ppm and four signals between 15 and 16 ppm due to C:H$^+$C base pairs (Fig. 2A). The signal at 13.3 ppm, which was previously attributed to A::T base pairing in the dimeric structure of the full DNA sequence **ON1**, was absent in the spectra of the modified sequences at 0.1 mM. No imino signals were observable for **ON2** and **ON3** above 30 and 40 °C at pH 5.5, while for **ON1** imino

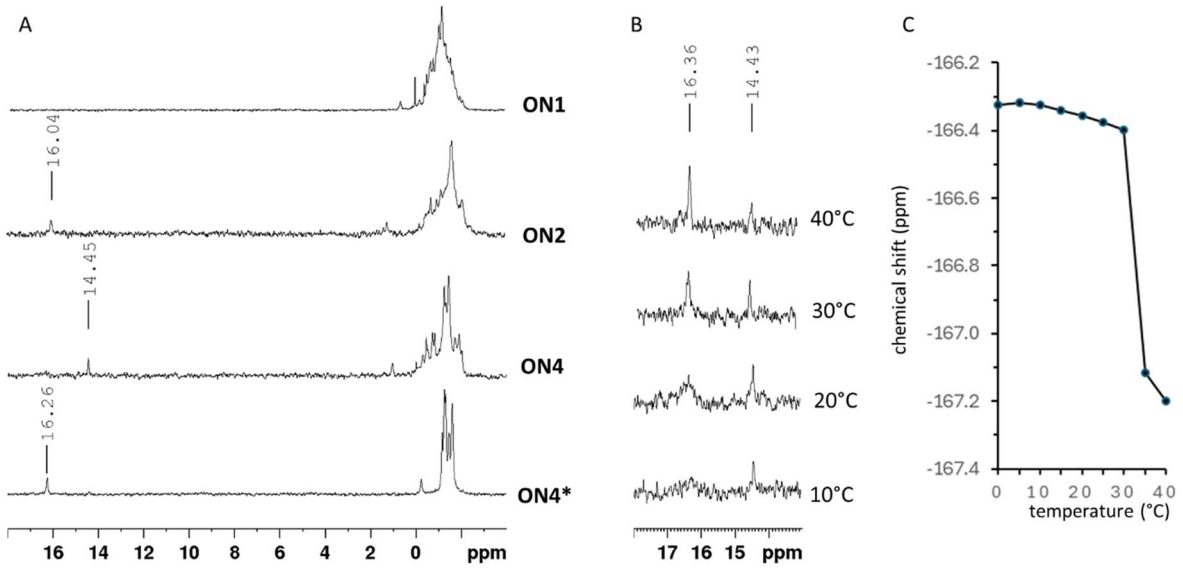


Fig. 3 (A) $^{31}$P spectra of **ON1**, **ON2**, and **ON4** at 10 °C in 90% $H_2O$/10% $D_2O$. Top 3 spectra were recorded at pH 5.2, bottom spectrum labelled as **ON4*** at pH 5.5. (B) Temperature dependence of the phosphonate moiety in **ON3** (90% $H_2O$/10% $D_2O$, pH 5.2). (C) Temperature dependence of $^{19}$F in **ON4** at pH 5.2 in 90% $H_2O$/10% $D_2O$.

signals were still detectable at 40 °C throughout the full pH range up to 6.3 (Fig. S6, SI).

A significant loss of iM stability was also previously observed in circular dichroism (CD) spectroscopy experiments upon introduction of aTNA[13] and UNA[12] monomers into the tetrameric region of an identical DNA sequence. For **ON4** containing a single modified 5-F-cytosine (*S*)-ZNA nucleoside at position 8, four sharp C:H$^+$C imino signals were only visible at 0 °C (Fig. 2A) and at pH 5.20 or lower. The fluorine substituent in 5-fluorocytidine is known to substantially lower the apparent p$K_a$ of the N3 nitrogen by 1.7–1.8 units compared to cytidine,[54] leading to a lower pH needed for stable iM formation in solution. A computational study by Yang *et al.* predicted that 5F-dC can form iMs, as the base-pairing energies (BPEs) of proton-bound heterodimers between cytosine and 5F-dC surpass those of canonical G:C base pairs, despite being lower than the BPEs of unmodified C:H$^+$C base pairs.[55]

A second 5F-C (*S*)-ZNA modification (**ON5**) was introduced to study the effect of a modified base pair within the iM structure, however imino signals were absent, even at low pH and temperature, suggesting that no stable iM was formed in contrast to what is reported for aTNA.[13]

Subsequently, $^{31}$P NMR was also performed to examine potential backbone changes upon introduction of ZNA modifications. The $^{31}$P signals arising from the phosphonate linkage in **ON2–ON5** were shifted downfield compared to those in the traditional phosphodiester linkages of **ON1**, reflecting a significantly different electronic environment (Fig. 3A). A phosphonate signal at 16.04 ppm was observed for (*R*)-ZNA-C in **ON2**, which was close to the reported values of 16.83 ppm for (*R*)-ZNA-C in a duplex and 16.75 ppm for (*S*)-ZNA-C in a hairpin.[19] In the folded iMs of **ON3** and **ON4** at 10 °C, the $^{31}$P resonance of (*S*)-ZNA-C and (*S*)-ZNA-5F-C were shifted downfield to 14.45 and 14.42 ppm, respectively (10 °C, pH 5.2). Upon increasing the temperature, a second signal arose in the $^{31}$P spectra of **ON3** and **ON4** at 16.36 and 16.26 ppm, respectively, while the original signal gradually decreased (Fig. 3B). This is characteristic for the co-existence of two conformational states that exchange slowly on the NMR timescale. Since the rising signal above 16 ppm matched the phosphonate chemical shift observed at pH 5.5 in the absence of iM formation, it was assigned to the unfolded sequence. For **ON2**, broadening of the phosphonate signal, rather than appearance of separate signals, was observed upon increasing the temperature to 40 °C, indicating that, for this oligonucleotide, the exchange rate between folded and unfolded structures approached the difference in chemical shifts between both states, which is considerably smaller than in **ON3** and **ON4**.

In contrast to the phosphonate signal in the iM backbone of **ON4**, the $^{19}$F signal of its modified nucleobase in the iM core at pH 5.2 exhibited a progressive chemical shift change with increasing temperatures (Fig. 3C). At 0 °C, **ON4** displayed a signal at −166.32 ppm corresponding to an iM-like structure, based on the $^1$H NMR spectrum showing characteristic imino proton signals at 14–16 ppm. Increasing the temperature led to a significant chemical shift change between 30 and 35 °C towards −167.20 ppm at 40 °C, approaching the $^{19}$F chemical shift of −167.26 reported in the literature for 5F-dC. This chemical shift change was comparable to the approximately 3.0 ppm downfield deshielding effect reported for 5F-U in a Watson–Crick geometry relative to a random coil context.[56]

## Characteristic NOE interactions within the ZNA modified iM structure

Unambiguous peak assignment was only possible for the well-resolved C:H$^+$C signals in the spectrum of **ON2** containing an

(*R*)-ZNA residue with a canonical nucleobase. This was achieved starting from the typical interresidue interactions of H1′ across the narrow grooves of the iM (Fig. 4).[43] As observed in the corresponding ZNA modified DNA chimeric duplex,[19] the C1′ proton signals of (*R*)-ZNA-C arose close to 3.4 ppm, rather than in the characteristic anomeric region where the H1′ protons of deoxyriboses typically resonate.

Next, strong intraresidue anomeric to aromatic NOE interactions allowed to assign the non-exchangeable nucleobase protons in cytosine residues, and subsequently their amino protons involved in hydrogen bonding within the iM core. From the latter assignment, the imino signals could also be identified in the NOESY spectrum. For UNA-C, destabilization of the iM was attributed to disruption of interactions within tetrameric core.[12] In contrast, **ON2** adopted a stable iM at low temperature with characteristic NOE interactions across its narrow grooves, suggesting that the ZNA monomer did not cause substantial alterations in the tetrameric core of the iM structure. As described for **ON1**,[43] loop signals in loops 1 and 3 were broadened due to conformational exchange and showed overlap due to symmetry within the folded structure, preventing the assignment of loop 1 (T4-A6) and loop 3 (T16-A18).

A similar assignment procedure was also carried out for (*S*)-ZNA C containing **ON3** (Fig. S7, SI). The four sharp imino signals above 15 ppm could be assigned to the four C:H$^+$C base pairs in the central core of the iM, including (*S*)-ZNA-C8::C20. In contrast to **ON2**, the interresidue interactions of C15:H1′ to *x*C8:H1′ across the narrow groove in the iM were not detected. Unfortunately, a strong overlap of imino signals for C9::C21 and C2::C14 base pairs prevented their unambiguous assignment. Broad signals in the imino proton region did not yield observable NOE contacts; this could arise from the closing base pairs C7::C19 and C1::C13, which were absent in the spectrum of **ON2**.

### Structural basis of iM destabilization by ZNA

An in-depth structural MD study was performed to gain further insight into the origin of the decreased iM stabilization due to the ZNA modification, as observed in UV melting and NMR experiments. After establishing an appropriate forcefield for (*R*)- and (*S*)-ZNA according to procedures described in the methods section, MD simulations were performed on the original iM of **ON1** as well as modified **ON2** and **ON3** sequences for 1000 ns (Fig. 5). Root-mean-square displacement (RMSD) plots for the resulting trajectories exhibited stable motions of the models throughout the entire simulation (Fig. 6, left panels), indicating that all simulations predicted stable iM structures when built of hemi-protonated C:H$^+$C pairs. Native and modified sequences led to slightly different stable states of the iM structure relative to the starting solution NMR structure of **ON1**. The modified sequences **ON2** and **ON3** displayed stronger fluctuations throughout the MD trajectory, suggesting that the ZNA-C residue slightly loosened the folded state of the iM, which corresponded to the observed decrease in $T_m$. However, analysis of hydrogen bonds between C8 and C20 residues revealed stable base pairing in both iMs (Table 3) which is in line with the sharp imino signal for the C:H$^+$C pair involving the modified residues observed in the NMR spectra.

The mobility of individual atoms or residues is reflected by the root-mean-square fluctuation (RMSF) of atomic positions relative to the reference structure. From the RMSF plot for individual atoms throughout the full simulation relative to the starting NMR structure of **ON1** (Fig. 6, right), it is clear that the core of the iM contains the least mobile elements with C:H$^+$C pairs exhibiting comparable low fluctuations from the reference state. Their limited mobility reflects a stable structure for the stacked hemi-protonated cytosine pairs, including the pair with the ZNA modification at position 8 in **ON2** and **ON3.**

Most TAA loop segments exhibited significant repositioning throughout the entire MD trajectory with strong peaks in the RMSF plot, which were due to the flapping motions of nucleobases in the loop residues. The largest differences in fluctuations between simulations were observed throughout the last 100 ns in the equilibrated phase MD trajectory.

As previously reported for the iM of **ON1**,[36,57] T10 and adenines A11 and A12 in loop 2 displayed strong flexibility, while the position of residues in loops 1 and 3 was much more stable. This phenomenon was attributed to the stabilizing effect of mutual interactions between loops 1 and 3 in the DNA iM. The RMSF plots of the modified sequences over the last 100 ns in the MD trajectory revealed a different pattern for **ON2**. Specifically, residue A12 in loop 2 demonstrated significantly reduced RMSF, while flapping motions for the nucleobases of T4 and A5 in loop 1 persisted in the last 100 ns of the simulation. Previous conformational dynamics unraveled that the stability of the studied iM is mainly associated with motion in interacting loops 1 and 3, where temperature and loop sequence can affect the dynamics and stability of the iM.[57] Increased RMSFs over the last 100 ns of the MD trajectory for the atoms of T4 and A5 reflected their increasing mobility within loop 1 according to the order **ON1** ≪ **ON3** < **ON2** (Fig. 6). The analysis of hydrogen bond distances in the last 200 ns of MD simulations for the of A6::A18 base pair flanking the core of the iM (Table 3) revealed increased average distances between atoms involved in hydrogen bonding for the modified oligomers compared to **ON1**, especially for **ON3**, indicative of increased

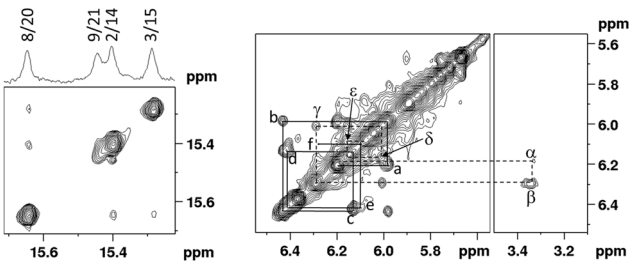


Fig. 4 Imino proton (left) and H1′, H1′-region (right) of the NOESY spectrum of **ON2** in 90% $H_2O$/10% $D_2O$ at 0 °C and pH 5.18 with mixing time of 300 ms. The NOESY walk of typical H1′–H1′cross peaks is indicated. Labelled peaks indicate a cross peak of adjacent H1′–H1′ protons with a: T22-C1, b: C1-C21, c: C21-C2, d: C2-C20, e: C20-C3, f: C3-C19, α: C15-*x*C8, β: *x*C8-C14, γ:C14-C9, δ: C9-C13 and ε: C13-T10.

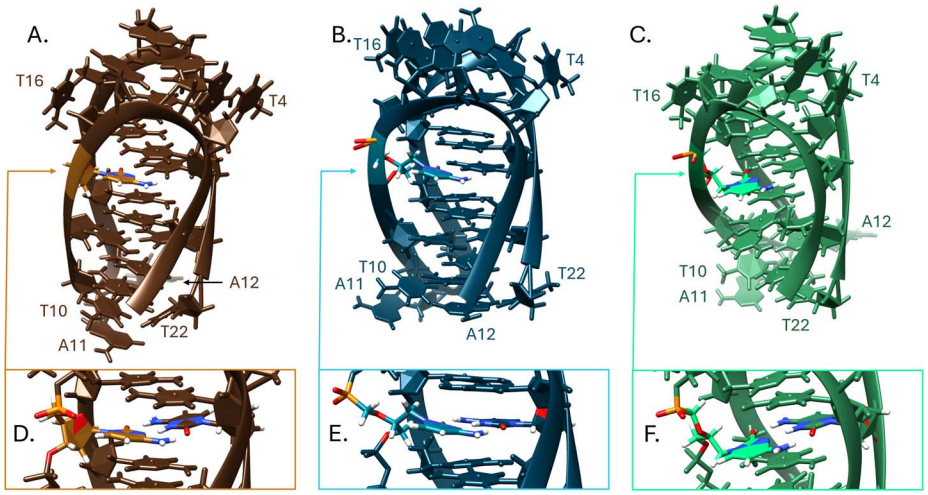


Fig. 5 (A) Full model view of **ON1** (last frame in MD). (B) Full model view of **ON2** (last frame in MD), showing the modified iM structure with an (*R*)-ZNA-C residue at position 8. (C) Full model view of **ON3** (last frame in MD), showing the modified iM structure with an (*S*)-ZNA-C residue at position 8. (D) Close-up view of the DNA ribose chemistry at position 8. (E) Close-up view of the modified (*R*)-ZNA-C at position 8. (F) Close-up view of the modified (*S*)-ZNA-C at position 8.

fraying of this base pair within the interacting loops 1 and 3. Increased dynamics in loops 1 and 3 of **ON2** and **ON3** was supported by the disappearance of the imino proton for the flanking C7::C19 base pair in the NMR spectrum of **ON2**, while it was observable for the reported d(CCCTAA*5mC*CCTAACC-C*U*AACCCT) sequence.[43]

Despite low RMSF values for atoms in hemi-protonated cytosine pairs, the incorporation of a ZNA-C residue had a clear impact on the puckering behavior of the sugar ring in its complement (residue 20) as well as in residues 2 and 3 with nucleobases stacking with the ZNA-nucleobase. For each of these residues, fluctuation between multiple ribose ring conformations was observed during the last 100 ns of the MD simulation (Fig. S8–S10, SI), which corresponded to a slightly increased RMSF at the sugar phosphate backbone for these residues in **ON2** and **ON3** relative to **ON1**. The effect on sugar

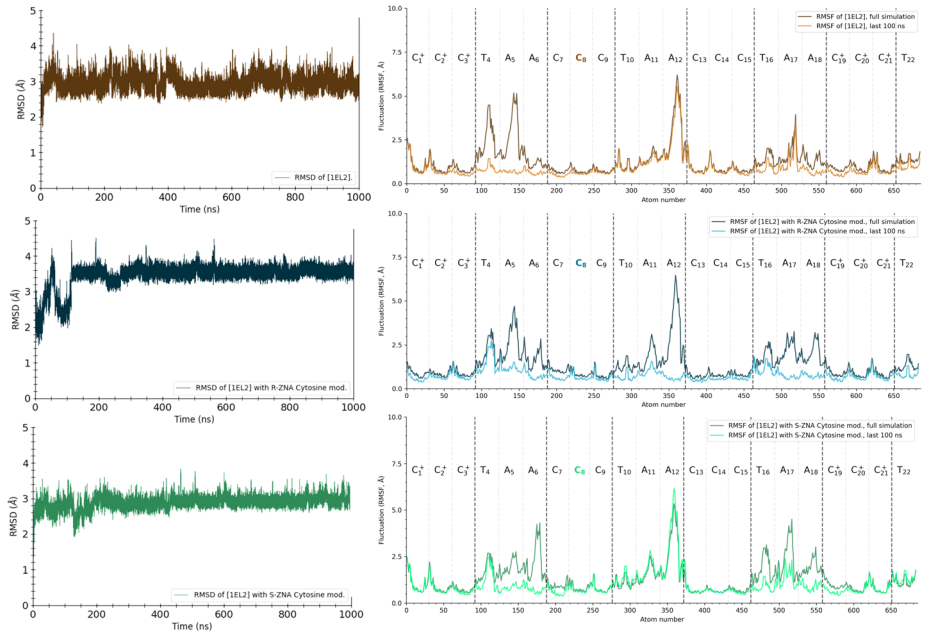


Fig. 6 (left): Root-mean-square displacement (RMSD) of iM spatial structure for **ON1** (top), **ON2** (middle), and **ON3** (bottom) during 1000 ns MD simulation. (right): Root-mean-square fluctuation (RMSF) of atomic positions within the iM structure obtained for **ON1** (top), **ON2** (middle), and **ON3** (bottom) over the full simulation time (dark color) and last 100 ns of the simulation (light color). The letters above denote the residues corresponding to the given atomic numbers. The $C^+$ letters denote the protonated cytosines. Atoms at the intersection between two residues (dashed lines in the RMSF plots) correspond to the sugar phosphate backbone, with atoms of nucleobases in between.

Table 3 Average atom distances (Å) in the last 200 ns of MD simulations on **ON1**, **ON2**, and **ON3** for a selected set of atoms. Standard deviations are reported in parentheses

| Atoms | ON1 | ON2 | ON3 |
|---|---|---|---|
| C8:O2-C$^+$20:H41 | 1.812 (±0.098) | 1.850 (±0.107) | 1.848 (±0.104) |
| C8:N3-C$^+$20:H3 | 1.940 (±0.092) | 1.933 (±0.091) | 1.916 (±0.085) |
| C8:H41-C$^+$20:O2 | 1.980 (±0.159) | 1.958 (±0.150) | 1.929 (±0.135) |
| A6:N7-A18:H62 | 2.042 (±0.143) | 3.188 (±1.714) | 5.631 (±0.193) |
| A6:H62-A18:N7 | 2.035 (±0.139) | 2.947 (±1.566) | 3.625 (±0.186) |
| T10:O4-T22:H3 | 2.046 (±0.551) | 8.843 (±0.374) | 2.335 (±1.480) |
| T10:H3-T22:O4 | 2.045 (±0.516) | 8.933 (±0.386) | 2.253 (±1.274) |
| A12:N7-T22:H3 | 15.243 (±2.159) | 2.035 (±0.162) | 15.438 (±2.022) |
| A12:H62-T22:O2 | 19.848 (±2.748) | 1.964 (±0.170) | 20.051 (±2.488) |

puckering in residue 3 was passed on to the first residue in loop 1 (T4) through the backbone linkage. In residue 9, immediately downstream of ZNA-C, the dominant sugar puckering shifted to the C2′-endo (South) conformation, while in DNA iM structures, the deoxyribose sugars primarily adopt the atypical C3′-endo (North) conformation, which is also found in A-type double stranded helices. The altered sugar conformation of residue 9 led to repositioning of loop residue T10, resulting in the loss of contact with T22. Instead of the sheared T10::T22 seen in **ON1** and **ON3**, T22 moved within hydrogen bonding distance of A12 to form a reverse Hoogsteen base pair in the structure of **ON2**, resulting in reduced RMSF for A12 upon incorporation of the ZNA-C modification (Fig. 6). Unfortunately, the imino signal at 11 ppm did not exhibit any NOE interactions. Therefore, NMR could not be used to obtain data on the base pairing pattern of T22.

At the other end of the iM, an A6::A18 Hoogsteen:Hoogsteen base pair was consistently observed between interacting loops 1 and 3 in all simulations. Nucleobases in the A6::A18 base pair were sandwiched between the C:H$^+$C base pair formed by C7 and C19 on one side and A17 on the other side. The increased RMSF for atoms in the nucleobase of A17 is caused by *syn*/*anti* flipping during the stabilized phase of the MD trajectories with transient loss of stacking interactions (Fig. 6). A similar RMSF pattern in loops 1 and 3 was also reported for comparative MD simulations on **ON1** using different forcefields.[36] However, these results contrast with reported NMR data for d(CCCTAA*5mC*CCTAACCC*U*AACCCT),[43] which describe *syn*/*anti* flipping for A18 in an A6::A18 base pair as an explanation for the observed conformational exchange between two conformers of interacting loops 1 and 3. In this NMR study, two base substitutions (italicized) were introduced in the vertebrate telomere sequence to break the quasi-symmetry between loops 1 and 3 that prevented full assignment in the NMR spectra of the natural sequence **ON1**. It is unclear to what extent these modifications affect the structure and dynamics of both loops, however authors suggested that tight packing of loops 1 and 3 might favor the configuration with the non-methylated U16 inside and T4 outside. In the equilibrated phase of all MD simulations, both T4 and T16 were pointing outside (Fig. 7). In agreement with NMR data, this configuration has a quasi-symmetrical structure for loops 1 and 3 with residues C7/C19, A6/A18, and T4/T16 that are replaced by their correspondent upon 180° rotation around the quasi-symmetry axis.

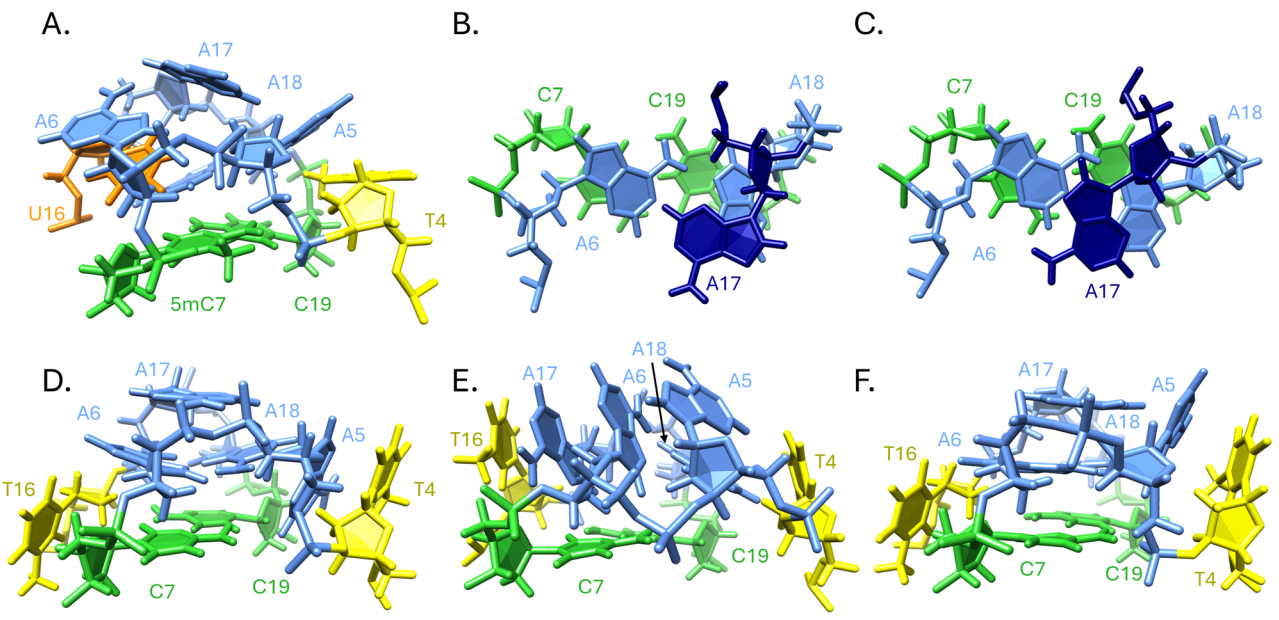


Fig. 7 (A) Comparison of loop 1 (T4-A6) and loop 3 (T16-A18) contacts in the solution NMR structure of d(CCCTAA*5mC*CCTAACCC*U*AACCCT)[43] (1el2.pdb). (D) Simulated **ON1** (frame at 1000 ns MD). (E) Simulated **ON2** (frame at 1000 ns MD). (F) Simulated **ON3** (frame at 1000 ns MD). Residues are color coded by type (C: green, T: yellow, U: orange, A: blue). Across all three simulated models (**ON1**–**ON3**), nucleobase A17 occurs in both *syn*- and *anti*-conformation. Representative frames from model **ON3** in top view on the A6::A18, sandwiched between C7::C19 and A17 are shown with A17 in *syn* at 910 ns (B) and *anti* at 995 ns (C). Nucleobase A17 is colored dark blue.

## Conclusions

In this work, we employed a combination of UV melting studies, NMR spectroscopy, and MD simulations to evaluate the structural effect of replacing a phosphodiester-deoxysugar residue within an iM with phosphonate-linked acyclic pseudosugar building blocks, namely (*R*) and (*S*) ZNA cytosine as well as (*S*)-ZNA 5-fluorocytosine monomers. Similar chemical shifts and NOE interactions in the NMR spectra of the 22-mer vertebrate telomeric DNA containing either a single (*R*)- or (*S*)-ZNA-C modified nucleotide in the iM tetrameric core confirmed a pH-dependent iM structure formation based on hemiprotonated C:H$^+$C base pairs. However, a significant decrease in thermal stability was observed by UV melting and NMR experiments. MD and NMR studies on modified oligomers demonstrated that a single acyclic ZNA-C residue fits in the tetrameric core of the iM but significantly impacts its dynamics in loop 1, likely contributing to the observed decrease in iM stability, with (*R*)-ZNA causing a greater destabilization compared to (*S*)-ZNA.

The presence of a 5-fluorine atom at the nucleobase moiety of (*S*)-ZNA-C further lowered the pH required to form a stable iM, while preserving the NOE interactions identified for the (*S*)-ZNA-C containing iM. This observation suggests the possibility of modulating the pH stability of iMs by introducing substituents at the C5-position of cytidine, enabling the design of iM structures suited for different pH ranges. Furthermore, the 5F-C nucleobase allowed to readily monitor the temperature-dependent folding of the corresponding iM by $^{19}$F NMR. The ability of 5F-C to provide distinct $^{19}$F spectral signatures *in vitro* is noteworthy, as this could be exploited for future applications such as in-cell NMR analysis of iM structures, especially since fluorine is nearly absent in biological systems, resulting in minimal background fluorine signal.[58]

The observed decrease in thermal stability upon ZNA incorporation could provide a useful means to finely tune iM behavior, potentially allowing to control their structural dynamics under diverse *in vitro* conditions; however further studies such as a systematic investigation of sequence context are needed to elucidate the broader applicability of this approach. Interestingly, since ZNA was demonstrated to be highly stable against enzymatic degradation,[23] it would be worthwhile to investigate whether this modification could enable the design of nuclease resistant iM structures.

## Author contributions

EL, EG, PH (conception and design), MD, SR, JR, XL, FC, JR, (investigation, synthesis, characterization, data collection, analysis of the data), EL, EG, MD, SR (drafting and writing of the paper), EL, EG, PH, MD, SR (revisiting it critically for intellectual content), EL, EG, PH (funding, supervision and resources), EL, EG, PH (final approval of the version to be published). All authors agree to be accountable for all aspects of the work.

## Conflicts of interest

There are no conflicts to declare.

## Data availability

Data for this article, including NMR and modeling data, are available *via* Figshare at **https://figshare.com/s/97546c036e9a979726f1**; DOI: **https://doi.org/10.6084/m9.figshare.29879990**.

Supplementary information: experimental procedures, NMR spectra, and other data supporting the results or analyses. See DOI: **https://doi.org/10.1039/d5ra05934h**.

## Acknowledgements

This research was supported by project grants of KU Leuven Research fund [No. C14/19/102] and Research Foundation-Flanders (FWO) [No. G085321N] as well as the FWO medium-scale research infrastructure grant (I008520N) entitled "An NMR facility tailored for biomedical research". We express our gratitude to Luc Baudemprez for his excellent technical assistance. Sten Reynders is an SB PhD fellow at the Research Foundation-Flanders (FWO) [No. SB/1S32925N]. Xinlu Li wishes to acknowledge the China Scholarship Council (CSC, grant number 201806350235).

## References

1 K. Gehring, J. Leroy and M. Guéron, *Nature*, 1993, **363**, 561–565.
2 A. Lieblein, J. Buck, K. Schlepckow, B. Fürtig and H. Schwalbe, *Angew Chem. Int. Ed. Engl.*, 2012, **51**, 250–253.
3 Z. A. E. Waller, *i-Motif Nucleic Acids in Handbook of Chemical Biology of Nucleic Acids*, Springer Nature, Singapore, 2023.
4 I. Zanin, o. E. Ruggier, G. Nicoletto, S. Lago, I. Maurizio, I. Gallina and S. Richter, *Nucleic Acids Res.*, 2023, **51**, 8309–8321.
5 S. Tao, Y. Run, D. Monchaud and W. Zhang, *Trends Genet.*, 2024, **40**(20), 4786–4804.
6 P. Školáková, D. Renčiuk, J. Palacký, D. Krafčík, Z. Dvořáková, I. Kejnovská, K. Bednářová and M. Vorlíčková, *Nucleic Acids Res.*, 2019, **47**, 2177–2189.
7 Y. Krishnan-Ghosh, E. Stephens and S. Balasubramanian, *Chem. Commun.*, 2005, 5278–5280, DOI: **10.1039/B510405J**.
8 J.-L. Mergny and L. Lacroix, *Nucleic Acids Res.*, 1998, **26**, 4797–4803.
9 H. A. Assi, R. W. t. Harkness, N. Martin-Pintado, C. J. Wilds, R. Campos-Olivas, A. K. Mittermaier, C. González and M. J. Damha, *Nucleic Acids Res.*, 2016, **44**, 4998–5009.
10 N. Kumar, J. T. Nielsen, S. Maiti and M. Petersen, *Angew. Chem., Int. Ed.*, 2007, **46**, 9220–9222.
11 M. Ghezzo, L. Grigoletto, R. Rigo, P. Herdewijn, E. Groaz and C. Sissi, *Biochimie*, 2023, **214**, 112–122.
12 A. Pasternak and J. Wengel, *Bioorg. Med. Chem. Lett.*, 2011, **21**, 752–755.

13 S. Pérez-Rentero, R. Gargallo, C. González and R. Eritja, *RSC Adv.*, 2015, **5**, 63278–63281.
14 V. Kumar, T. J. D. Nguyen, J. Palmfeldt and K. V. Gothelf, *Org. Biomol. Chem.*, 2019, **17**, 7655–7659.
15 M. Abdelhamid and Z. Waller, *Front. Chem.*, 2020, **8**, 40.
16 H. Assi, R. Harkness, N. Martin-Pintado, C. Wilds, R. Campos-Olivas, A. Mittermaier, C. González and M. Damha, *Nucleic Acids Res.*, 2016, **44**, 4998–5009.
17 D. Yuanchen, Y. Zhongqiang and L. Dongsheng, *Acc. Chem. Res.*, 2014, **47**, 1853–1860.
18 Y. Chen, K. Morihiro, Y. Nemoto, A. Ichimura, R. Ueki, S. Sando and A. Okamoto, *Chem. Sci.*, 2024, **15**, 17097–17102.
19 X. Li, E. Groaz, P. Herdewijn and E. Lescrinier, *Chem.–Eur. J.*, 2024, **30**, e202401254.
20 T. Sunami, J. Kondo, I. Hirao, K. Watanabe, K. Miura and A. Takénaka, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2004, **60**, 90–96.
21 H. Chen, S. Viel, F. Ziarelli and L. Peng, *Chem. Soc. Rev.*, 2013, **42**, 7971–7982.
22 M. Luo, E. Groaz, S. De Jonghe, R. Snoeck, G. Andrei and P. Herdewijn, *ACS Med. Chem. Lett.*, 2018, **9**, 381–385.
23 M. Luo, E. Groaz, M. Froeyen, V. Pezo, F. Jaziri, P. Leonczak, G. Schepers, J. Rozenski, P. Marlière and P. Herdewijn, *J. Am. Chem. Soc.*, 2019, **141**, 10844–10851.
24 A. Shaka, C. Lee and A. Pines, *J. Magn. Reson.*, 1988, **77**, 274–293.
25 M. Thrippleton and J. Keeler, *Angew Chem. Int. Ed. Engl.*, 2003, **42**, 3938–3941.
26 T. Hwang and A. Shaka, *J. Magn. Reson., Ser. A*, 1995, **112**, 275–279.
27 C. Mattelaer, H. Mattelaer, J. Rihon, M. Froeyen and E. Lescrinier, *J. Chem. Theory Comput.*, 2021, **17**, 3814–3823.
28 F. Neese, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2022, **12**(5), e1606.
29 R. Krishnan, J. Binkley, R. Seeger and P. JA, *J. Chem. Phys.*, 1980, **72**, 650–654.
30 M. Frisch, J. Pople and J. S. Binkley, *J. Chem. Phys.*, 1984, **80**, 3265–3269.
31 A. Hellweg, C. Hättig, S. Höfener and W. Klopper, *Theor. Chem. Acc.*, 2007, **117**, 587–597.
32 R. Sure and S. Grimme, *J. Comput. Chem.*, 2013, **34**, 1672–1685.
33 H. Kruse and S. Grimme, *J. Chem. Phys.*, 2012, 136.
34 F. Neese, *J. Comput. Chem.*, 2003, **24**, 1740–1747.
35 F. Weigend, *J. Comput. Chem.*, 2007, **29**, 167–175.
36 T. Panczyk, K. Nieszporek and P. Wolski, *Molecules*, 2022, **27**(15), 4915.
37 U. Singh and P. Kollman, *J. Comput. Chem.*, 1984, **5**, 129–145.
38 J. Rihon, C. Mattelaer, R. Montalvão, M. Froeyen, V. Pinheiro and E. Lescrinier, *Nucleic Acids Res.*, 2024, **52**, 2836–2847.
39 C. Bayly, P. Cieplak, W. Cornell and P. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269–10280.
40 Y. Xu, E. Groaz, J. Rihon, P. Herdewijn and E. Lescrinier, *Eur. J. Med. Chem.*, 2023, **255**, 115379.
41 D. A. Case, T. E. Cheatham III, T. Darden, H. Gohlke, R. Luo, K. M. Merz Jr, A. Onufriev, C. Simmerling, B. Wang and R. J. Woods, *J. Comput. Chem.*, 2005, **26**, 1668–1688.
42 H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Res.*, 2000, **28**, 235–242.
43 A. Phan, M. Guéron and J. Leroy, *J. Mol. Biol.*, 2000, **299**, 123–144.
44 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.
45 M. Zgarbová, M. Otyepka, J. Sponer, A. Mládek, P. Banáš, T. Cheatham and P. Jurečka, *J. Chem. Theory Comput.*, 2011, **7**, 2886–2902.
46 V. Kräutler, W. van Gunsteren and P. Hünenberger, *J. Comput. Chem.*, 2001, **22**, 501–508.
47 R. Davidchack, R. Handel and M. Tretyakov, *J. Chem. Phys.*, 2009, **130**, 234101.
48 H. Berendsen, J. Postma, W. van Gunsteren, A. DiNola and J. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.
49 R. Salomon-Ferrer, A. Götz, D. Poole, S. Le Grand and R. Walker, *J. Chem. Theory Comput.*, 2013, **9**, 3878–3888.
50 D. R. Roe and T. E. Cheatham, *J. Chem. Theory Comput.*, 2013, **9**, 3084–3095.
51 E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605–1612.
52 S. S. Wijmenga and B. N. M. van Buuren, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1998, **32**, 287–387.
53 J. Leroy, M. Guéron, J. Mergny and C. Hélène, *Nucleic Acids Res.*, 1994, **22**, 1600–1606.
54 I. Wempen, R. Duschinsky, L. Kaplan and J. Fox, *J. Am. Chem. Soc.*, 1961, **83**, 4755–4766.
55 B. Yang and M. Rodgers, *J. Am. Chem. Soc.*, 2014, **136**, 282–290.
56 A. Solodinin, A. Gautrais, S. Ollivier and H. Yan, *ACS Omega*, 2019, **4**, 19716–19722.
57 M. Mondal, D. Bhattacharyya and Y. Gao, *Phys. Chem. Chem. Phys.*, 2019, **21**, 21549–21560.
58 F. Theillet, *Chem. Rev.*, 2022, **122**, 9497–9570.