Nucleic Acids Research, 2018 **1**
doi: 10.1093/nar/gky121

i-Clamp phenoxazine for the fine tuning of DNA i-motif stability

Vladimir B. Tsvetkov[1,2,3], Timofei S. Zatsepin[4,5], Evgeny S. Belyaev[6], Yury I. Kostyukevich[4], George V. Shpakovski[7], Victor V. Podgorsky[1], Galina E. Pozmogova[1], Anna M. Varizhuk[1,8,*] and Andrey V. Aralov[7,*]

[1]Biophysics Department, Research and Clinical Center for Physical Chemical Medicine, Malaya Pirogovskaya str. 1a, Moscow 119435, Russia, [2]Department of Molecular Virology, FSBI Research Institute of Influenza, Ministry of Health of Russian Federation, prof. Popov str. 15/17, Saint-Petersburg, 197376, Russia, [3]Polyelectrolytes and Biomedical Polymers Laboratory, A.V. Topchiev Institute of Petrochemical Synthesis, RAS, Leninsky prospect str. 29, Moscow 119991, Russia, [4]Center for Translational Biomedicine, Skolkovo Institute of Science and Technology, 3 Nobel street, Skolkovo, Moscow 143026, Russia, [5]Chemistry Department, Lomonosov Moscow State University, Leninskie gory str. 1–3, Moscow 119992, Russia, [6]Frumkin Institute of Physical Chemistry and Electrochemistry of the Russian Academy of Science, Leninsky prospect str. 31, Moscow 119071 Russia, [7]Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Miklukho-Maklaya str. 16/10, Moscow 117997, Russia and [8]Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Vavilov str. 32, Moscow 119991, Russia

Received November 10, 2017; Revised February 01, 2018; Editorial Decision February 06, 2018; Accepted February 13, 2018

ABSTRACT

Non-canonical DNA structures are widely used for regulation of gene expression, in DNA nanotechnology and for the development of new DNA-based sensors. I-motifs (iMs) are two intercalated parallel duplexes that are held together by hemiprotonated C-C base pairs. Previously, iMs were used as an accurate sensor for intracellular pH measurements. However, iM stability is moderate, which in turn limits its *in vivo* applications. Here, we report the rational design of a new substituted phenoxazine 2′-deoxynucleotide (*i*-clamp) for iM stabilization. This residue contains a C8-aminopropyl tether that interacts with the phosphate group within the neighboring chain without compromising base pairing. We studied the influence of *i*-clamp on pH-dependent stability for intra- and intermolecular iM structures and found the optimal positions for modification. Two *i*-clamps on opposite strands provide thermal stabilization up to 10–11°C at a pH of 5.8. Thus, we developed a new modification that shows significant iM-stabilizing effect both at strongly and mildly acidic pH and increases iM transition pH values. *i*-Clamp can be used for tuning iM-based pH probes or assembling extra stable iM structures for various applications.

INTRODUCTION

Non-B DNA structures are gaining increasing attention due to their presumed biological roles (1–4), usability as aptamer scaffolds (5–6) or other types of DNA functional units (7). i-Motifs (iMs) are noncanonical C-rich DNA architectures and are composed of intercalated parallel duplexes held together by hemiprotonated C-C base pairs (8). Although the *in vivo* formation of such structures is disputable, iM-prone sites are found in telomeric regions of the human genome, centromeric regions and oncogenic promoters. These findings implicate possible regulatory functions of such sites and make them attractive potential targets for drug discovery (9–13). Stability of exogenous iMs in the complex cellular environment has been recently confirmed using in-cell NMR spectroscopy (14). iMs have diverse applications in bionanotechnology and medicinal chemistry such as pH-triggered nanomotors, hydrogels and drug carriers (15–16). Analytical chemistry makes use of iMs as probes for small molecule ligands, cations and proteins, while the most straightforward application is the development of pH-sensors (17). The sensors are typically designed for optical detection and include fluorescently labeled iMs (either terminally labeled ones (18) or those con-

*To whom correspondence should be addressed. Tel: +7 495 336 4200; Fax: +7 495 336 4200; Email: Baruh238@mail.ru
Correspondence may also be addressed to Anna M. Varizhuk. Tel: +7 499 246 4570; Fax: +7 499 246 4409; Email: annavarizhuk@gmail.com
Present address: Andrey V. Aralov, Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Miklukho-Maklaya str. 16/10, Moscow 117997, Russia.

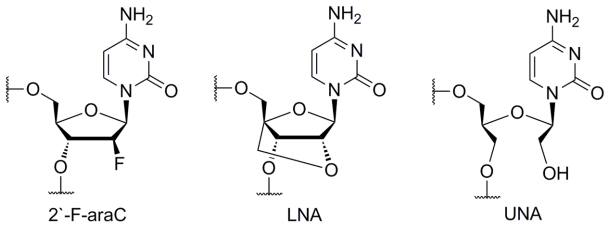


Figure 1. Structures of iM-stabilizing chemical modifications.

taining cytosine substitutions with fluorescent residues inside the core (19–20)) and iM conjugates with nanoparticles (21).

One of the first successful sensors for *in vivo* monitoring of spatial and temporal pH changes (e.g. upon endocytosis or endosome maturation) was based on an intermolecular iM (22). However, most sensors utilize intramolecular structures because of their fast folding kinetic, which is a prerequisite to rapid sensor response. Apart from relatively fast kinetics, a perfect sensor should display a high dynamic range and a linear response in a sufficiently broad pH range (23–24). Unmodified iMs are typically stable in a relatively narrow pH range (3–6.5). Outside this range, cytosines are either all protonated or all deprotonated, and hemiprotonated pairs are theoretically non-existent. However, bioinformatic studies reveal an increasing number of iMs that can sustain physiological conditions (25). This is due to the fact that the cytosine pK_a in an iM core is actually different from that of a free nucleic base (4.5) and depends on multiple factors including core size (26), loop bases (27) and additional structural elements, such as hairpins or minor groove tetrads (28). All these factors may affect both the pH transition point and cooperativity of iM folding, which open possibilities for tuning iM-based sensors (this general approach can be referred to as 'structural' (29)).

Another popular approach to modulate iM stability is chemical modification of nucleic acids. To date, several nucleotide modifications have been proposed for iM stabilization, including (i) base modifications (5-methyl-cytosine, uracil and 5-(1-propynyl)uracil as parts of 2′-deoxynucleotides (30)); (ii) sugar modifications (2′-deoxy-2′-fluororibocytidine (31), 2′-arabinocytidine (32), 2′-deoxy-2′-fluoroarabinocytidine (2′-F-araC) (33), locked (LNA) (34–35) and unlocked (UNA) (36) nucleic acids) and (iii) phosphate modifications (phosphoramidate, methylphosphonate and phosphorothioate (30,37)). Among them, LNA, 2′-F-araC and UNA appear to be the most promising for iM stabilization (Figure 1). However, each known modification has limitations. LNA stabilizes iM only at acidic pH and exhibits a position-dependent stabilization. In contrast, 2′-F-araC shows its iM-stabilizing effect at acidic and, most importantly, neutral pH without pronounced position-dependence, as well as strongly shifts pH transition values (up to 0.8). However, the stabilizing effects of both LNA and 2′-F-araC do not exceed 4°C per modification, and multiple substitutions are needed to observe such significant improvement. Regarding UNA, insertions in loops resulted in a $T_{1/2}$ increase up to 5°C, but core modifications led to drastic destabilization. Finally, chemical modifications also include nucleotide and non-nucleotide inserts containing intercalating moieties which stabilize iM structures through stacking interactions. For example, the insertion of pyrene-attached 2′-deoxyadenosine nucleotide (^{Py}A) at 5′-end led to prominent increase of T_m (+11.4°C) at pH 4.0 that was caused by the stacking interaction between ^{Py}A and the terminal hemiprotonated C–C base pair (38). However other pyrene-containing non-nucleotide inserts, uracil UNA monomer (39) and TINA (40), within the loop region were shown to destabilize iM. In turn, different porphyrin-containing nucleotide and non-nucleotide derivatives were placed into the fragment of the sequence that did not participate in iM core formation, and they provoked strong stabilization (average T_m was 53°C, whereas T_m of the unmodified sequence was <25°C) at pH 5.0 owing to porphyrin–porphyrin interactions (41). Naphthalimide derivatives were introduced into the TAA loop of iM-forming sequence and one of them showed nearly equal stabilizing effects of 6°C at different pH (5.2 and 6.2) under crowding and noncrowding conditions (42). In addition, a shift of pH transition values up to 0.5 was observed. Finally, a group of disubstituted anthraquinone linkers were introduced into central and the external TAA loops, where the best one, 2,6-substituted anthraquinone linker, increased T_m up to 8.2°C due to improved stacking interactions and hydrogen bonding in the i-motif construct (43). In this study, we examined whether G-clamp, a phenoxazine-based cytosine analog that forms extra stable pairs with guanine in duplexes (44–45), would stabilize iMs. We considered the possibilities of hemiprotonated cytosine–clamp and clamp–clamp pair formation for G-clamp and guanidino-G-clamp residues (Figure 2A). A phenoxazine nucleotide *per se* (i.e. without the clamp-specific aminoalkyl/guanidinoalkyl tether) has been reported to cause insignificant destabilization when introduced into the iM core (46). In the case of G-clamp and guanidino-G-clamp, one could expect stabilization due to additional H-bonding between the protonated clamp tether and the oxo-group of the opposing cytosine/clamp (Figure 2B). Analogous tether interactions with the oxo-group of the opposing guanine residues have been observed in duplexes with the G-clamp (44–45). Alternatively, the tether could form H-bonds with phosphodiester fragments of a neighboring strand. Analogous interactions have been observed in duplexes with the G^{8AE}-clamp (47).

We report on the validation of G-clamp pairing in iMs and discuss prospects of using clamp modifications for modulating the pH-sensitivity and thermal stability of iM-based pH probes. We also report on optimization of G-clamp geometry that yielded a new i-motif-stabilizing nucleobase analogue (*i*-clamp). To the best of our knowledge, *i*-clamp is the first nucleobase modification that shows strong iM-stabilizing effect both at strongly (pH < 4.6) and slightly acidic pH (pH > 5.6) and tends to increase iM transition pH values. We propose that i-clamp can be used for tuning iM-based pH probes or assembling extra stable iM structures for *in vivo* applications.

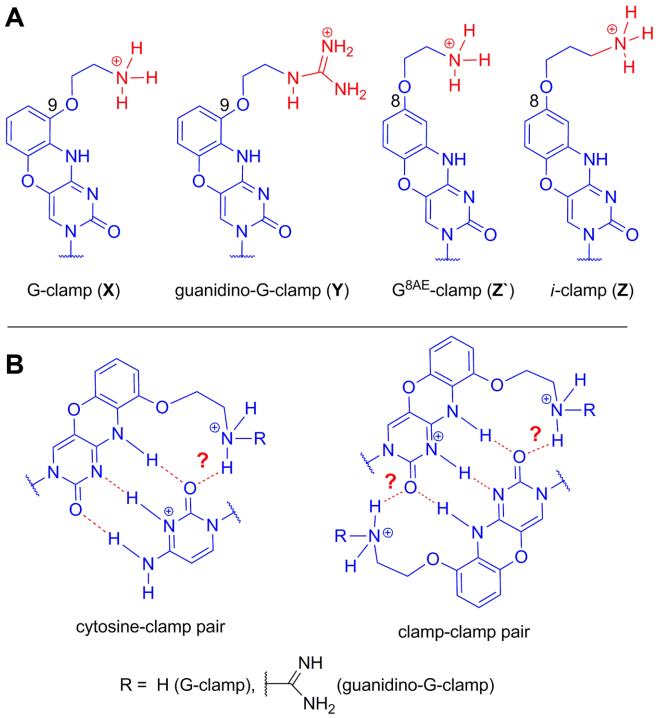


Figure 2. Structures of the clamp modifications used in this study (**A**) and putative hemiprotonated cytosine–clamp (left) and clamp–clamp (right) pairs in iMs (**B**).

MATERIALS AND METHODS

Oligonucleotide synthesis

Oligonucleotides were assembled in an ABI 3400 DNA synthesizer with the phosphoramidite method according to the manufacturer´s recommendations. Protected 2′-deoxyribonucleoside 3′-phosphoramidites, Unylinker-CPG (500Å) and *S*-ethylthio-1*H*-tetrazole were purchased from ChemGenes. Oligonucleotides were cleaved from the support and deprotected using AMA – 1:1 (v/v) conc. aq. ammonia and 40% aq. methylamine for 3 h at room temperature. Postsynthetic guanidinylation was performed according to the literature (48). Oligonucleotides were double purified by denaturing PAGE followed by RP-HPLC. The denaturing gel electrophoresis of oligonucleotides was performed in 15% PAGE containing 7 M urea in Tris-borate buffer (50 mM Tris–HCl, 50 mM boric acid, 1 mM EDTA, pH 8.3). Oligonucleotides were recovered from the gel by electroelution with Elutrap (Whatman) in Tris-borate buffer (5 mM Tris–HCl, 5 mM boric acid, pH 8.3). The HPLC purification of oligonucleotides was carried out on a 4.6 × 250 mm Jupiter C_{18} column (5 μm, Phenomenex); buffer A: 0.05 M ammonium acetate (pH 7), 5% MeCN; buffer B: 0.03 M ammonium acetate, 80% MeCN, pH 7; a gradient of B: 0–15% (1 CV), 15–50% (10 CV); a flow rate of 1 ml/min; temperature 45°C. The HPLC analysis of oligonucleotides was carried out on a 4.6 × 250 mm Jupiter C18 column (5 μm, Phenomenex); buffer A: 0.05 M ammonium acetate (pH 7), 5% MeCN; buffer B: 0.03 M ammonium acetate, 80% MeCN, pH 7; a linear gradient of B: 0–100% (8 CV); a flow rate of 1 ml/min; temperature 45°C. LC–MS analysis for oligonucleotides was performed using Agilent 1260-Bruker Maxis Impact system or Dionex Ultimate 3000 – Thermo Fisher LCQ ion trap system as described earlier with minor modifications (49). The HPLC was equipped with the 2.1 × 50 mm Jupiter C_{18} column (5 μm, Phenomenex); buffer A: 10 mM diisopropylamine, 15 mM 1,1,1,3,3,3-hexafluoroisopropanol; buffer B: 10 mM diisopropylamine, 15 mM 1,1,1,3,3,3-hexafluoroisopropanol, 80% MeCN. Salts were washed out with buffer A (4 CV) followed by a step of 100% B (2 CV) with a flow rate of 0.3 ml/min; temperature 45°C. The MS analysis of oligonucleotides was carried out in negative mode (capillary voltage: 3500 V, dry temperature: 160°C), raw spectra were deconvoluted by maximum entropy method (Bruker) or Promass software (Thermo).

UV absorption, circular dichroism and fluorescence spectroscopy

Oligonucleotides were dissolved in 10 mM sodium acetate (pH 4.0–5.2), sodium phosphate (pH 5.8–8) or Tris–HCl (pH 8.2–8.8) buffers (100 mM NaCl was added in the cases of M_0, G_0 and their derivatives) to final concentrations of 5 μM (U_0, M_0, G_0 and their derivatives) or 10 μM (T_0 and its derivatives). The samples were denatured at 95°C for 5 min and snap cooled on ice to ensure intramolecular folding (U_0 and its derivatives) or cooled slowly to obtain thermodynamically most favorable structures (T_0, M_0, G_0 and their derivatives). UV absorption spectra, CD spectra, CD melting curves and fluorescense emission spectra were recorded on a Chirascan spectrophotometer (Applied Photophysics) equipped with a thermostatted cuvette holder. Clamp fluorescence emission was registered upon excitation at 370 nm. CD spectra were recorded at 5°C. pH Transition values were determined as inflection points of fitted sigmoid curves in ΔCD^{norm} versus pH plots:

$$Y = 1 - 1/(1 + 10^{(pHi-x)*n}), \quad (1)$$

where Y is normalized ellipticity (ΔCD^{norm}) at 288 nm: $\Delta CD^{norm} = (CD^{current\ pH} - CD^{min})/(CD^{maX} - CD^{min})$, x is current pH value, pH_i is pH transition value and n is Hill coefficient.

In the case of T_1Z, the pH transition value was additionally verified using data on clamp absorbance and fluorescence emission (the experimental data were also fitted with the sigmoid function (1), Y = normalized absorbance or fluorescence intensity). Fitting was performed using DataFit 9 software (Oakdale Engineering).

Thermal difference spectra (TDS) were obtained by subtracting absorption spectra obtained at 5°C from the spectra obtained at 90°C. The melting/annealing temperatures of the i-motifs were estimated from the maximum in the first derivative of the melting curve. The ellipticity at 285 nm was registered every 1°C across the 5–90°C temperature range (heating/cooling rate = 1°C/min).

Molecular modeling

The models of the tetramolecular iMs were created using Sybyl-X software (Certara, USA) and two related structures: PDB_1YBL and PDB_2N89. The former is an A2C4

Table 1. Sequences, melting temperatures (T_m) at pH 5.8 (10 mM sodium-phosphate buffer) and transition pH values (pHi) of native and modified iMs

Code	Sequence (5′-3′)	T_m, ±1°C (ΔT)	pH_i (±0.1)
U_0	CCCATCCCATCCCATCCC	37	6.0
U_1X	**X**CCAT**X**CCATCCCATCCC	46 (+9)	6.3
U_2X	C**X**CATC**X**CATCCCATCCC	39 (+2)	5.7
U_3X	C**X**CATCCCATC**X**CATCCC	48 (+11)	5.9
U_1Y	**Y**CCAT**Y**CCATCCCATCCC	46 (+9)	6.0
U_2Y	C**Y**CATC**Y**CATCCCATCCC	42 (+5)	6.2
U_3Y	C**Y**CATCCCATC**Y**CATCCC	39 (+2)	6.1
U_1Z	**Z**CCAT**Z**CCATCCCATCCC	48 (+11)	6.9
U_2Z	C**Z**CATC**Z**CATCCCATCCC	47 (+10)	6.6
U_3Z	C**Z**CATCCCATC**Z**CATCCC	47 (+10)	5.9
T_0	TCCCCC	47	5.8
T_1X	T**X**CCCC	42 (–5)	6.0
T_2X	TCC**X**CC	–	NA
T_1Y	T**Y**CCCC	41 (-6)	6.2
T_2Y	TCC**Y**CC	–	NA
T_1Z	T**Z**CCCC	52 (+5)	6.6
T_2Z	TCC**Z**CC	33 (–14)	NA

X = G-clamp; Y = guanidino-G-clamp; Z = *i*-clamp.

tetramer (50); the latter is a T_0 analog with 2′F-araC modifications at C4 and C5 positions (33). We first combined the 3′-terminal iM core of PDB_1YBL with the 5′-terminal fragment of PDB_2N89 to obtain the unmodified iM T_0, and then performed preliminary conformation optimization. Next, we substituted cytosine residues at C2 positions for G-clamp, guanidino-G-clamp, G^{8AE}-clamp or *i*-clamp residues to obtain T_1X, T_1Y, T_1Z' and T_1Z, respectively. The models of the unimolecular iMs $U_1X/Y/Z'/Z$, U_3X, U_1Y and U_3Y were created based on the telomeric iM structure (51) and adjusted analogously to T_1X and T_1Y. Optimization at each step was performed as previously described (52). Partial charges for modified nucleotides were calculated at the same level of the theory of quantum mechanical calculations that was used to create Amber nucleic acid force field. All quantum mechanics simulations were carried out using Gaussian 09 program (53). MD simulations were performed using Amber 16 software (54) as described previously (52). The parameters for interatomic energy calculation were taken from the general AMBER force field (GAFF) for bases of modified nucleotides residues, and the OL15 force field (55–57) was used for the rest. Snapshot visualization and hydrogen bond analysis were performed using VMD (58) with a donor-acceptor distance of 3.2 Å and angle cut off of 30°. Free energies and the contributions which form them were estimated as previously described (59).

NMR spectroscopy of the iM-forming oligodeoxyribonucleotides

Oligonucleotides were dissolved in 10 mM sodium phosphate buffer (pH 5.8) to final concentrations of 0.1 mM (U_0 and its derivatives) or 0.2 mM (T_0 and its derivatives), and the samples were annealed as described in the previous subsection. 10% D_2O for each sample was added for lock signal stabilization. ^{1}H NMR common 1D spectra with WATERGATE water suppression (relaxation delay 1.2 s, 200 μs delay for homospoil/gradient recovery and 20 μs delay for binomial water suppression, 64k points digital resolution and 24 ppm sweep width in 128 transitions with 16 dummy scans) and two 2D NOESY spectra with WATERGATE water suppression (with 50 ms and 150 ms mixing time, relaxation delay 1.2 s, 2k points digital resolution, in 8 transitions with 16 dummy scans, the water suppression settings was the same as for 1D spectra) at 300 K were registered for each sample. The spectra were obtained on Bruker Avance III 500 MHz NMR spectrometer equipped with triple-channel TCI Prodigy cryoprobe.

Gel electrophoresis

Nondenaturing polyacrylamide gel (20%) was prepared in 20 mM Tris–AcOH buffer (pH 5.2). The oligonucleotide samples were equilibrated in sodium acetate buffer (pH 5.2) and annealed as described in the UV absorption section. A mixture of 10–100 nt single-stranded oligonucleotide fragments (Low Molecular Weight Marker, Affymetrix) was used as a control. The gel was run for 2 h at 200 V at room temperature with 20 mM Tris–AcOH buffer (pH 5.2) as a running buffer. The gel was stained with SYBR Gold (Thermo Fisher Scientific) and analyzed using Gel-Doc scanner (BioRad).

RESULTS AND DISCUSSION

Effects of G-clamp and guanidino-G-clamp insertions in i-motifs

We synthesized a series of native and modified oligodeoxyribonucleotides (ONs) that can adopt intramolecular (U_0 and its derivatives), as well as intermolecular (T_0 and its derivatives) structures (Table 1). U_0 is a fragment of human microsatellites and is frequently found in intergenic regions and introns. T_0 is a well-known model ON; its secondary structure has been characterized previously (8) (see Figure 3 for schematic representations of all iM structures). Melting curves and CD spectra of the native iMs are shown in Supplementary Figure S1 (left and right panels, respectively). The amplitudes of characteristic CD peaks (288 nm) were used to determine transition pH values (pH_i), at which 50% of the ON is folded into an iM structure (middle panel in Supplementary Figure S1). Minor hysteresis

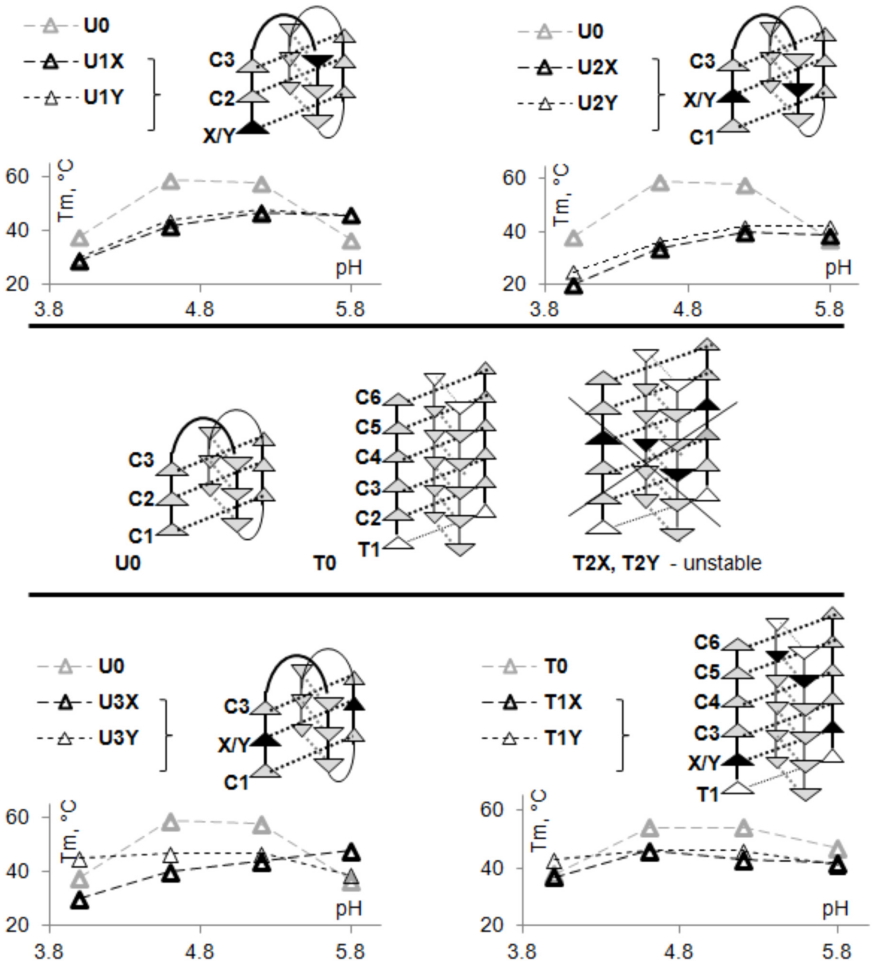


Figure 3. Schematic representations of putative iM structures and pH dependence of their thermal stabilities.

(Supplementary Figure S2A) and negligible effects of concentration on U_0 thermal stability (summarized in Supplementary Figure S2B) confirm unimolecular folding. The tetramolecular structure T_0, on the contrary, displayed pronounced hysteresis (Supplementary Figure S2A) and concentration dependence of its melting/annealing temperatures (Supplementary Figure S2B) agreed with Equation (2) for tetramolecular structures ($n = 4$):

$$1/T_{1/2} = \ln(C/2)^{n-1} * R/\Delta H - \Delta S * R/\Delta H, \qquad (2)$$

where $T_{1/2}$ is the average of annealing and melting temperatures; n is a stoichiometric coefficient, ΔH and ΔS are enthalpic and entropic contributions to free energy of iM formation.

The positions of G-clamp (**X**) and guanidino-G-clamp (**Y**) modifications suggest formation of cytosine–clamp pairs (Figure 2B; left structures) in $U_{1-2}X/Y$ iMs and clamp–clamp pairs (Figure 2B; right structures) in U_3X/Y, T_1X/Y and T_2X/Y iMs. The putative clamp–clamp pairs in T_2X and T_2Y presumed structures (Figure 3) are in close proximity, which may cause steric hindrance. Indeed, T_2X and T_2Y iMs were unstable under mildly acidic conditions (Supplementary Figure S3). Melting temperatures (T_m) of other modified structures under the same conditions are presented in Table 1; the respective CD spectra and melting curves are shown in Supplementary Figures S4 and S5.

As evident from Table 1, both modifications caused substantial iM destabilization in the tetramolecular series and zero to profound stabilization (ΔT_m average is approximately +3°C per modification) in the unimolecular series at pH 5.8. Detailed analysis of pH dependence (Supplementary Figures S4 and S5; summarized in Figure 3) revealed strong negative effects of **X** and **Y** on iM thermal stabilities at lower pH, although pH_i values slightly increased in several cases (Table 1). We attribute stabilization under mildly acidic conditions to nonspecific stacking interactions; while in general, the geometries of the G-clamp residue and its guanidino derivative seem unfavorable for efficient pairing in iMs. To clarify this matter and to gain insight into rational strategies for clamp optimization, we performed molecular modeling experiments.

Molecular modeling: classical versus new clamp variants

Modeling and molecular dynamics (MD) simulations were performed for the tetramolecular iMs with clamp–clamp pairs T_1X and T_1Y (Figure 4), unimolecular iMs with clamp–clamp pairs U_3X and U_3Y (Supplementary Figure S6) and unimolecular iMs with cytosine–clamp pairs U_1X

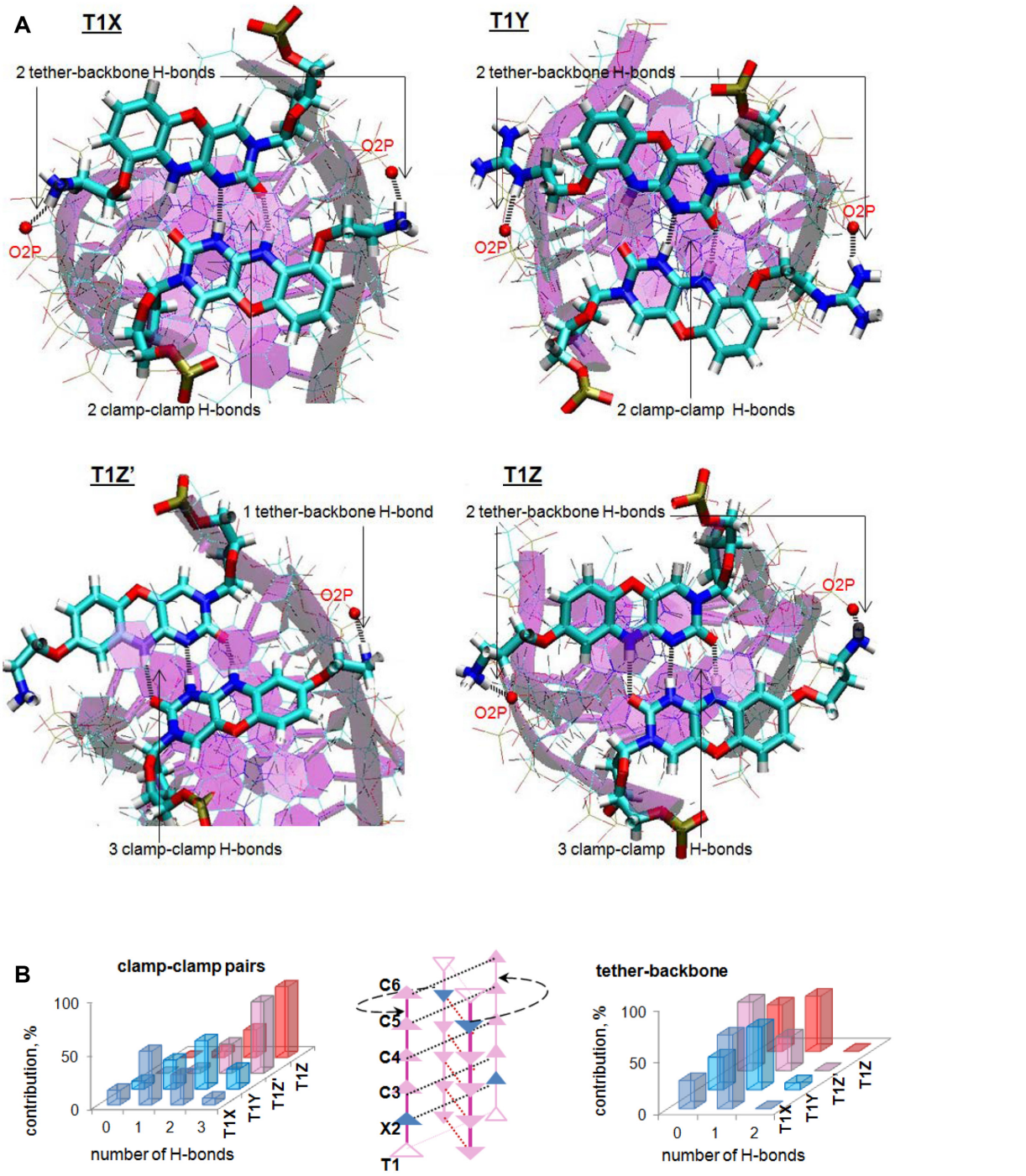


Figure 4. Molecular modeling of tetramolecular iMs with clamp–clamp pairs. The results of the molecular dynamics simulation. (**A**) 5 ns snapshots; oxygen atoms from -C5-p-C6- phosphate residues (O2P) are marked with red; black dashed lines indicate hydrogen bonds; (**B**) summary of H-bonding efficiency; contribution of snapshots with 0–3 H-bonds in the clamp–clamp pairs (left graph) or between the clamp tethers and the backbones of the neighbouring strands (right graph) is given as percentage from the total number of snapshots from the MD simulation trajectory.

and U_1Y (Supplementary Figure S7). For details on model preparation, see Materials and Methods.

All attempts to incorporate the aminoethyl/guanidinoethyl tether of the phenoxazine nucleotide inside the iM core resulted in van der Waals repulsion between the tether and the backbone of the neighboring strand. Thus, although the protonated tether could participate in H-bonding with an opposing base in isolated parallel duplexes, in intercalated duplexes, such bonding is strongly constrained by the neighboring strand. Instead, the tether tends to interact with the neighboring strand backbone according to our data.

Importantly, the tether-backbone interactions gradually dragged the clamp residues apart in our MD simulations,

which lead to the disruption of the clamp–clamp (top panels in Figure 4a and Supplementary Figure S6) and cytosine–clamp (Supplementary Figure S7) pairing. We concluded that the tether should be shifted to another position of the phenoxazine moiety (i.e. C8) and/or elongated to enable efficient tether-backbone interactions with simultaneous base pairing. This conclusion encouraged us to study respective clamp analogs in additional modeling experiments.

First we tested the recently reported G^{8AE}-clamp (47). Its 2-aminoethyl tether is shifted to position 8 of the phenoxazine ring (G^{8AE}-clamp; Figure 2A). The modification was favorable but insufficient for iM stabilization; H-bonding in the pairs was restored, but the tether-backbone interactions were partially lost (see T_1Z' in Figure 4, U_3Z' in Supplementary Figure S6 and U_1Z' in Supplementary Figure S8). Therefore, we replaced the aminoethyl group with the aminopropyl one (*i*-clamp; Figure 2A). The resulting structures were stable throughout the simulation (see T_1Z in Figure 4, U_3Z in Supplementary Figure S6 and U_1Z in Supplementary Figure S8) with maintenance of both base pairing and tether-backbone binding. H-bonding efficacy in classical and new clamp–clamp pairs and tether-backbone interaction efficacy are summarized in Figure 4B (tetramolecular iMs) and Supplementary Figure S6 (unimolecular iMs).

As illustrated by the left graph in Figure 4b, G-clamp residues in T_1X formed loose pairs that were held together by a single H-bond in 51% MD simulation snapshots and two H-bonds in 29% snapshots. Guanidino-G-clamp pairs in T_1Y were also relatively loose with a single H-bond in 27% snapshots and 2 H-bonds in 46% snapshots. G^{8AE}-clamp and *i*-clamp, on the contrary, formed stable pairs in T_1Z' and T_1Z, respectively, with three H-bonds in ∼67% snapshots (these modifications appear to be equally beneficial in terms of pairing efficiency). In terms of tether-backbone interaction efficiency (right graph in Figure 4B), *i*-clamp (54% T_1Z snapshots with one bond versus 46% snapshots with zero bonds) was clearly superior to G^{8AE}-clamp (33% T_1Z' snapshots with one bond versus 67% with zero bonds). The distributions of H-bonding probabilities for classical and modified clamp pairs in $U_3X/Y/Z'/Z$ agreed with those in $T_1X/Y/Z'/Z$, except that *i*-clamp was superior to G^{8AE}-clamp in both pairing efficiency and tether-backbone interactions (Supplementary Figure S6, bottom graphs). The analogous trend was observed for iMs with cytosine–clamp pairs ($U_1X/Y/Z'/Z$). G-clamp formed 1–2 H-bonds with opposing cytosines in most of the snapshots (see 'X6 + C16' and 'X1 + C11' H-bonding time plots in Supplementary Figure S7), and guanidino-G-clamp–cytosine pairs were mostly held together by 2 H-bonds ('Y6 +C16' and 'Y1 + C11' plots in Supplementary Figure S7), while *i*-clamp and G^{8AE}-clamp tended to form 2–3 (usually 3) H-bonds with opposing cytosines ('Z6/Z6' + C16' and 'Z1/Z'1 + C11' plots in Supplementary Figure S8). At the same time, the 8-aminopropyl tether of *i*-clamp bound to the neighboring strand backbone much more efficiently than the 8-aminoethyl one (see 'Z6 8AP tether + DNA backbone' and 'Z6' 8AE tether + DNA backbone' plots in Supplementary Figure S8).

It can be concluded that, according to our modeling results, *i*-clamp would be a good candidate for iM stabilization.

Positional effects and interaction patterns: detailed *in silico* analysis

As evident from Table 1, the effects of G-clamp and guanidino-G-clamp inserts are position-dependent. In the tetramolecular series, terminal C-track modifications (T_1X and T_1Y) were tolerated, while middle-chain modifications (T_2X and T_2Y) lead to complete iM destruction. In the unimolecular series, shifting the modifications from C-track termini (U_1X and U_1Y) to middle-track positions (U_2X and U_2Y) had negative effects on thermal stabilities ($T_m^{U1X} = 46 \pm 1°C$ versus $T_m^{U2X} = 39 \pm 1°C$; $T_m^{U1Y} = 46 \pm 1°C$ versus $T_m^{U2Y} = 42 \pm 1°C$, pH 5.8). We compared U_1X/Y and U_2X/Y models to elucidate the observed position dependency and performed additional modeling experiments with their Z' and Z analogs to clarify whether similar effects can be expected for the new clamp variants. Major MD simulation results for $U_1X/Y/Z'/Z$ and $U_2X/Y/Z'/Z$ are shown in Supplementary Figures S7–S9.

Interestingly, in-depth investigation of the $U_1X/Y/Z'/Z$ models revealed a specific role of clamp residues adjacent to loops: they participated in stacking interactions with loop bases. This was most apparent for U_1Y and U_1Z (Supplementary Figure S10). In U_1Y MD simulation, the phenoxazine residue of Y6 tended to stack with Thy-5 via either oxazine or benzene ring. We considered all stacking variants, i.e. estimated center of mass (COM) distances between Thy and all three phenoxazine rings and selected the minimal one for each snapshot. The conventional approach suggests using an overall phenoxazine COM, but it would be non-representative in our case because of the transient stacking mode. The COM^{min} distance time plot in the left panel of Supplementary Figure S10 illustrates Y6 stacking with Thy-5 after ∼4 ns (COM distance < 4 Å). Co-planarity of the thymine and phenoxazine residues is confirmed by the Φ time plot, where Φ is the angle between the normals to thymine and phenoxazine planes (Φ ≈ 10° after 4 ns). Analogously, Z6 stacking with Ade-14 throughout the U_1Z simulation is illustrated by the respective COM distance (≤4 Å) and Φ (≤18°) time plots in the right panel of Supplementary Figure S10.

Figure 5A shows comparison of free energies of $U_1X/Y/Z'/Z$ and $U_2X/Y/Z'/Z$ calculated based on MD simulations. In all cases terminal and adjacent to loop modifications (the U_1 series) were energetically more favorable than middle-core modifications (the U_2 series). Analysis of relative contributions of various forces to $U_{1–2}$ energy difference (Figure 5B) revealed the primary role of the torsion energy (dihedrals). This can be interpreted as follows: clamp residues tend to form H-bonds with opposing cytosines and/or stack with the neighbouring ones, which 'fixes' them in particular planes. Middle-core residues have minimal conformational mobility, whereas terminal and loop-adjacent ones retain certain freedom. Positively charged clamp tethers attract negatively charged backbone phosphates of the neighboring strands, which requires deviation of the latter strands from the optimal backbone geometry and accounts for the dihedral stress (the stress in the middle of the core cannot be released).

The fact that the tethers interact predominantly with phosphate residues is evident from Supplementary Table S2.

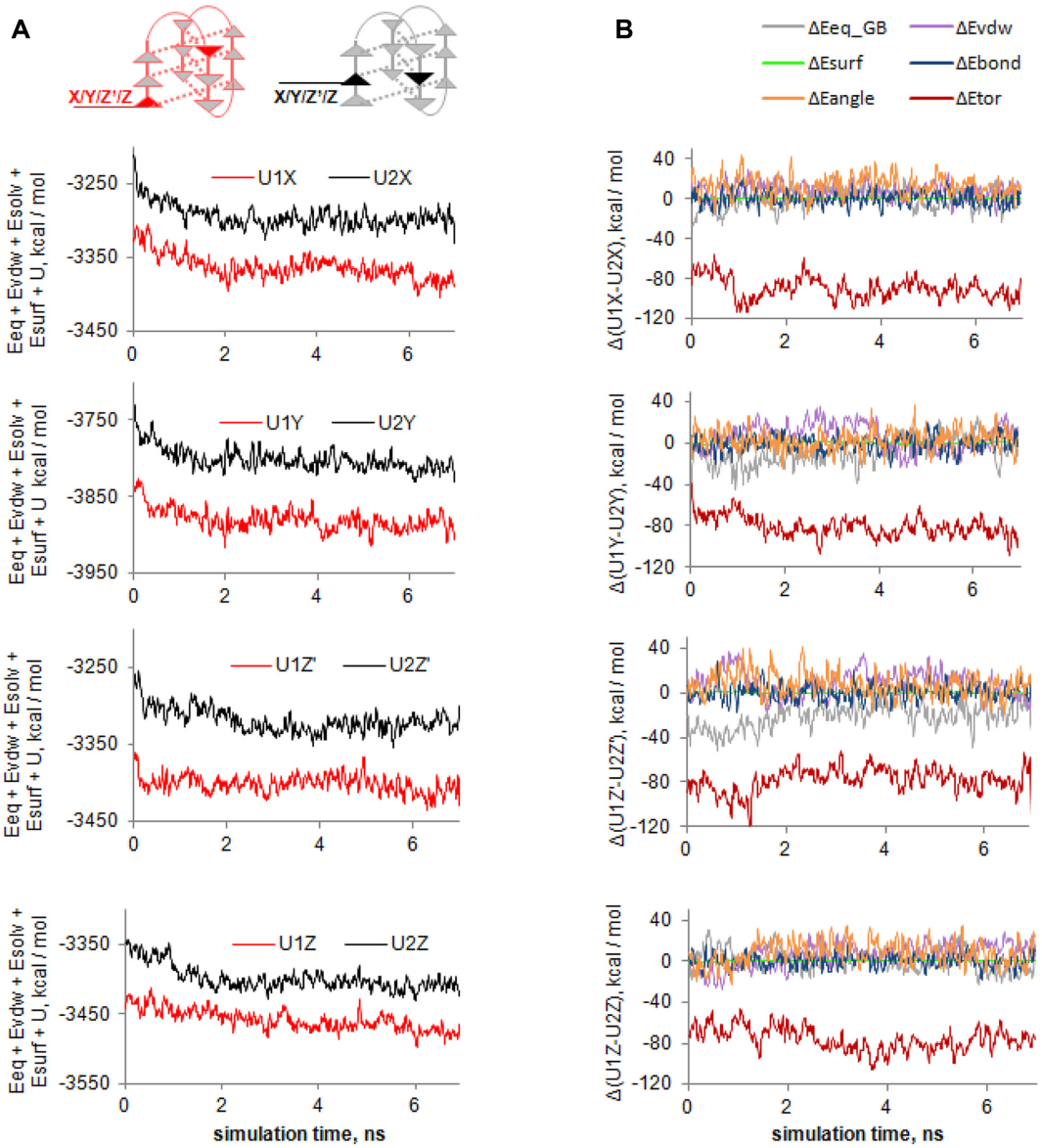
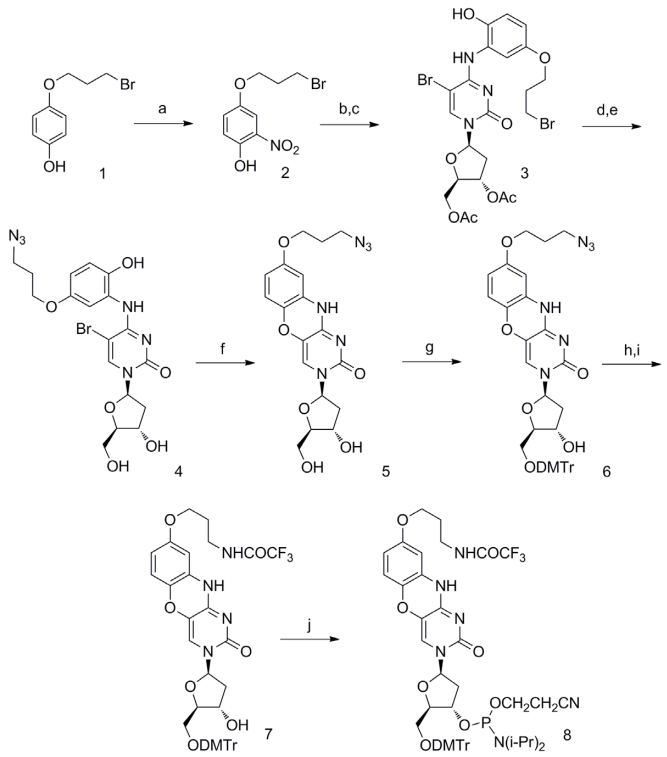


Figure 5. Free energies of U_0 derivatives with terminal and middle-chain cytosine–clamp pairs. (**A**) Comparison of total free energies. (**B**) Contribution of particular factors to the difference (Δ) between $U_1X/Y/Z'/Z$ and $U_2X/Y/Z'/Z$ energies. Eeq + EGB—electrostatic energy of the molecule in the gas phase and the polar component of solvation energy computed using the Generalized Born (GB) method; Evdw—van der Waals energy; Esurf—non-polar component of solvation energy; Ubond, Uangle and Utor—bond stretching, angle bending and torsion stress energies. U = Ubond + Uangle + Utor. The energy plots were smoothed using the moving average method (span = 5).

The data in the table also illustrate the benefits of the 8AP-tether (*i*-clamp) over the 9AE one (G-clamp). The probabilities of tether-phosphate bonds, as well as H-bonds 4-$NH_2^{cytosine}$ – $O2^{phenoxazine}$, H-$N3^{+cytosine}$ – $N1^{phenoxazine}$ and H-$N10^{phenoxazine}$ – $O2^{cytosine}$ in hemiprotonated cytosine–clamp pairs (calculated as percentage of the respective snapshots) increase from $U_{1,2}X$ to $U_{1,2}Z$. The left part of the table can be regarded as a detalization of the graphs in Supplementary Figures S7 and S8. Once again, our molecular modeling results suggest that *i*-clamp is a leading clamp variant for iMs, and terminal modifications would be likely superior to middle-core ones.

Synthesis of the *i*-clamp phosphoramidite

i-Clamp is a new structure, and its incorporation into ONs required synthesis of the respective phosphoramidite. For the synthesis of *i*-clamp phosphoramidite **8**, we have applied a previously developed scheme with several modifications (Scheme 1) (47). 4-(3-Bromopropoxy)phenol **1** was treated with 56% aqueous nitric acid solution at 0°C yielding para-substituted nitrophenol **2**. Compound **2** was catalytically hydrogenated and reacted with 3′,5′-*O*-acetyl-4-*N*-(1,2,4-triazol-1-yl)-5-bromo-2′-deoxycytidine in the presence of DIPEA to give product **3**. The derivative **4** was prepared by treatment of **3** with sodium azide in DMF followed by the removal of acetyl groups with ammonia in aqueous



Scheme 1. Synthesis of the *i*-clamp phosphoramidite. (**A**) HNO_3, benzene, 48%; (**B**) H_2, Pd/C, MeOH; (**C**) 3′,5′-*O*-acetyl-4-*N*-(1,2,4-triazol-1-yl)-5-bromo-2′-deoxycytidine, DIPEA, CH_2Cl_2, 58% for 2 steps; (**D**) NaN_3, DMF, (**E**) NH_3, MeOH, H_2O, 74% for 2 steps; (**F**) TEA, C_2H_5OH, 42%; (**G**) DMTr-Cl, pyridine, 77%; (**H**) H_2, Pd/C, TEA, MeOH; (**I**) ethyl trifluoroacetate, MeOH, 74% for 2 steps; (**J**) $NCCH_2CH_2OP(Cl)N$-*i*Pr_2, DIPEA, CH_2Cl_2, 77%.

methanol. The cyclization was performed by refluxing **4** in the mixture of TEA/C_2H_5OH. Standard 5′-dimethoxytrityl protection of phenoxazine derivative **5** was followed by reduction of the azido group and trifluoroacetyl protection. Phosphitylation of **7** afforded the target phosphoramidite **8** that was used in common solid phase synthesis of modified ONs T_1Z, T_2Z, and $U_{1–3}Z$ (Table 1). After standard AMA deprotection, modified ONs were purified by RP HPLC and characterized by LC-MS (Supplementary Table S1).

Effects of *i*-clamp insertions in i-motifs

The iMs with *i*-clamp insertions were characterized analogously to the iMs with G-clamp and guanidino-G-clamp insertions by melting curves and CD spectra (Figure 6A, Supplementary Figure S11). The new modification enhanced thermal stabilities of the structures at mildly acidic pH values (Table 1) in the unimolecular series (ΔT_m average is approximately +5°C per modification) and, in contrast to G-clamp and its guanidino derivative, for tetramolecular T_1Z (ΔT_m = +5°C). Similar to the T_2X and T_2Y cases, we attribute the negative effect of *i*-clamp in T_2Z to steric hindrance. Data on pH dependencies of T_1Z and $U_{1–3}Z$ melting temperatures are summarized in Figure 6A. Position dependency of *i*-clamp effects in U_0 derivatives is in line with molecular modeling data (U_1Z is superior to U_2Z).

The formation of T_1Z, U_1Z and U_3Z iMs at pH 5.8 was confirmed by NMR spectroscopy data (Figure 6B, Supplementary Figures S12 and S13). Two groups of peaks near 15 ppm in the ^{1}H-NMR spectrum (Supplementary Figure S13) of U_3Z point to the coexistence of two iM forms. This assumption agrees with reduced cooperativity of U_3Z melting under mildly acidic conditions (pH 5.2–5.8) compared with a pH of 4.0 (Supplementary Figure S11). To get additional insight into this matter, we performed electrophoretic mobility assay at pH 5.2 (Supplementary Figure S14A). The results should be interpreted with some care, considering that traditional running buffers, including the Tris–acetate one used herein, are not perfect for mild acidic conditions.

At 10 μM concentration U_0 and its derivatives were mostly monomeric with electrophoretic mobilities between those of 15-mer and 20-mer ssDNA controls. However, in the case of U_3Z, a presumably tetrameric admixture with electrophoretic mobility close to that of a 70-mer ssDNA control was clearly observed. Increasing oligonucleotide concentration to 40 μM resulted in the appearance of additional isomers (dimers, as suggested by the respective band shifts) in all samples. At a high concentration (150 μM), comparable to that used in the NMR experiments (100–200 μM), the dimers dominated. In the case of U_3Z, two bands with close shifts were observed in the 'dimer' range. Those bands could refer to 3′E and 5′E conformers, which typically differ in chemical shifts of imino-proton NMR signals (24), or isomers with different loop arrangements, i.e. loops spanning over minor- or major-grooves (Supplementary Figure S14B). The band shifts of T_0 and T_1Z agree roughly with tetramolecular structures (Supplementary Figure S14A).

Inter- *vs* intramolecular folding of T_1Z and U_1Z was further verified by analyzing concentration effects on melting and annealing temperatures at pH 5.2. The results are analogous to those obtained previously for T_0 and U_0 (Supplementary Figure S2). U_1Z and T_1Z exhibited moderate and pronounced hysteresis, respectively (Supplementary Figure S2A). Concentration dependence of $T_{1/2}$ (Supplementary Figure S2B) agreed with Equation (2) at $n = 4$ for T_1Z and was mostly within the experimental error for U_1Z: increasing U_1Z concentration from 5.5 to 75 μM enhanced the iM stability slightly, while further increase had a negligible effect.

We conclude that the T_1Z is likely tetramolecular, while U_1Z is likely intramolecular at low (≤10 μM) concentrations but prone to dimerization at high concentrations. Putative schemes of U_1Z dimers are shown in Supplementary Figure S14b. Interestingly, they suggest clamp–clamp pairing, as opposed to cytosine–clamp pairing in monomers, though shifted isomers cannot be excluded.

Figure 6c summarizes pH sensitivities of *i*-clamp absorbance and fluorescence in T_1Z. The observed quenching of *i*-clamp fluorescence and the bathochromic shift of its absorption band under acidic conditions (Figure 6c) are in line with the reported data for related dC analogues (46). The transition pH value calculated based on *i*-clamp absorbance and fluorescence changes (pH_i = 6.7±0.1, middle panel in Figure 6c) agrees with our estimates based on the T_1Z CD data (6.6±0.1,middle panel in Supplementary Figure S11). Overall, *i*-clamp insertions caused profound

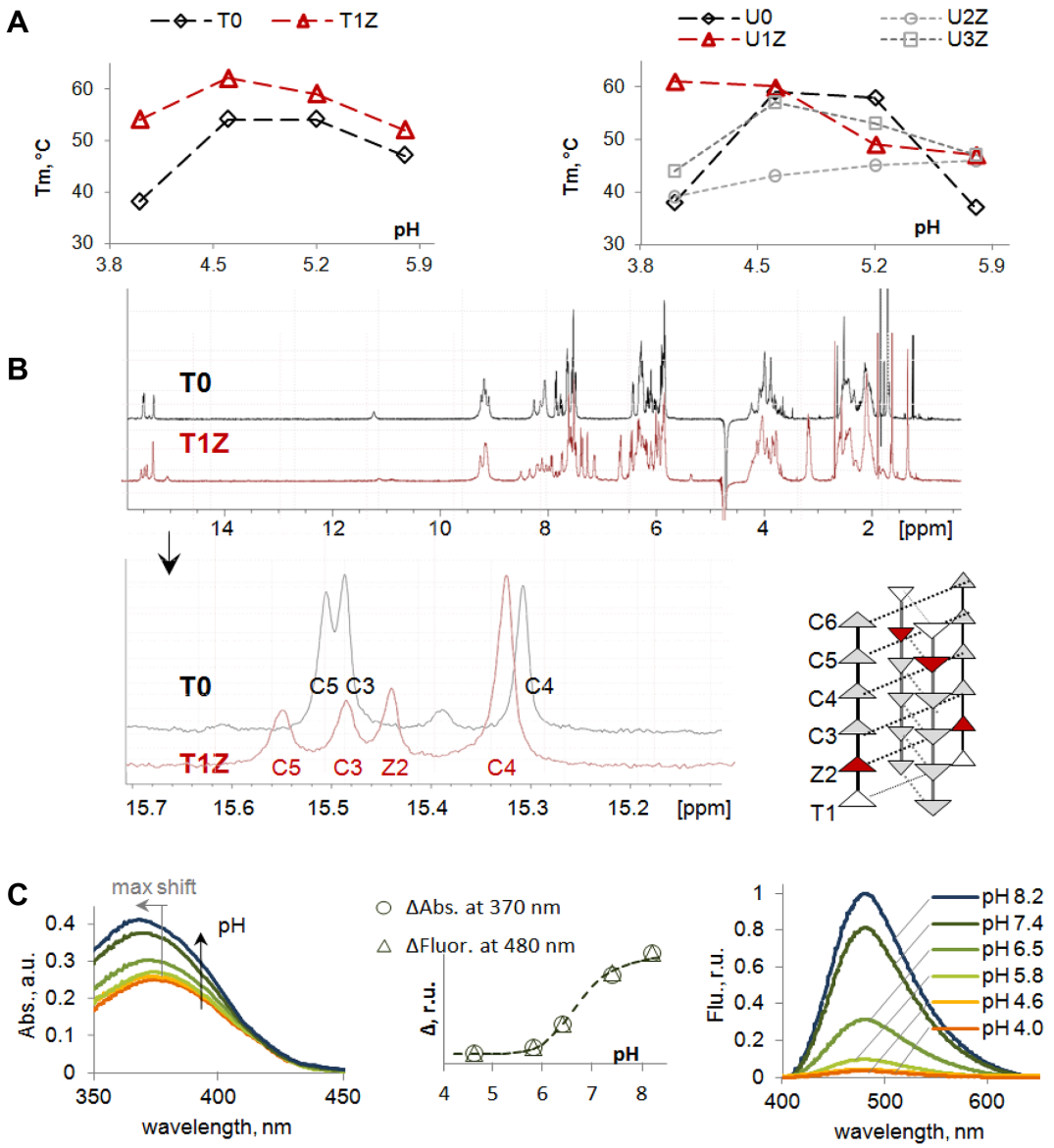


Figure 6. Characterization of iMs with *i*-clamp insertions. (**A**) pH-dependence of iM thermal stabilities; (**B**) ^{1}H NMR spectra (pH 5.8); (**C**) pH-dependence of *i*-clamp absorbance and fluorescence in T_1Z.

iM stabilization under both strongly acidic (pH 4.0) and mildly acidic (5.8) conditions (Figure 6a) and substantially increased transition pH values (ΔpH_i T_1Z = +0.8; ΔpH_i U_1Z = +0.9; Table 1). In this regard, we questioned whether *i*-clamp could be used to obtain stable iMs under near-physiological conditions.

Modified iMs at physiological pH

For studies under near-physiological conditions, we synthesized an additional set of ONs (Table 2) that included known native iMs, M_0 and G_0, and their derivatives with *i*-clamp insertions. M_0 is a model iM, and G_0 is genomic sequence; both iMs have been reported to be stable at a pH of 7.4 (25). The results of iM characterization by optical methods are summarized in Figures 7 and 8.

As evident from Figure 7A, all the modified iMs were folded at pH 7.4 in the presence of 100 mM NaCl ((see characteristic positive CD bands at 288 nm and negative TDS bands at 295 nm), whereas the native model ON M_0 was mostly unfolded (no significant iM signatures in TDS and CD spectra). Thermal stabilities of the modified iMs at pH 7.4 (Figure 7B) depended on the positions of *i*-clamp residues. Melting and annealing temperatures (T_m and T_a) are given in Table 2 (hysteresis could be attributed to relatively slow folding kinetics as compared to temperature ramping rate). Middle-chain modifications were generally unfavourable, 3′-terminal modifications had minor negative effects, and 5′-terminal modifications were stabilizing (especially, in the case of the 5′-terminal clamp–clamp pair). Similar effects were observed in the genomic iM series (Supplementary Figure S15).

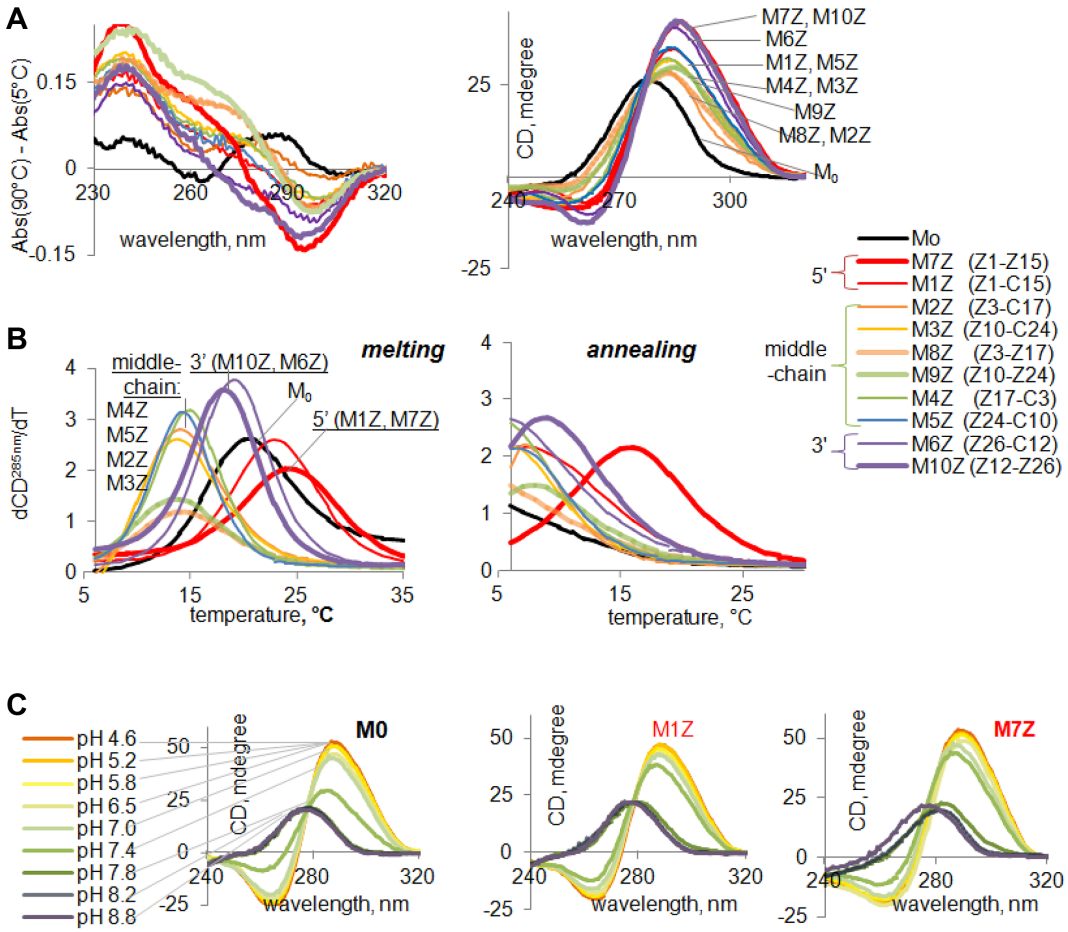


Figure 7. Characterization of model iMs with *i*-clamp insertions that are stable near-physiological conditions. (**A**) TDS and CD spectra, pH 7.4. (**B**) Melting and annealing curves (first derivatives), pH 7.4. The modified pairs are specified in the legends; (A) and (B) have joint legends. c: pH-Dependencies of M_0, M_1Z and M_7Z CD.

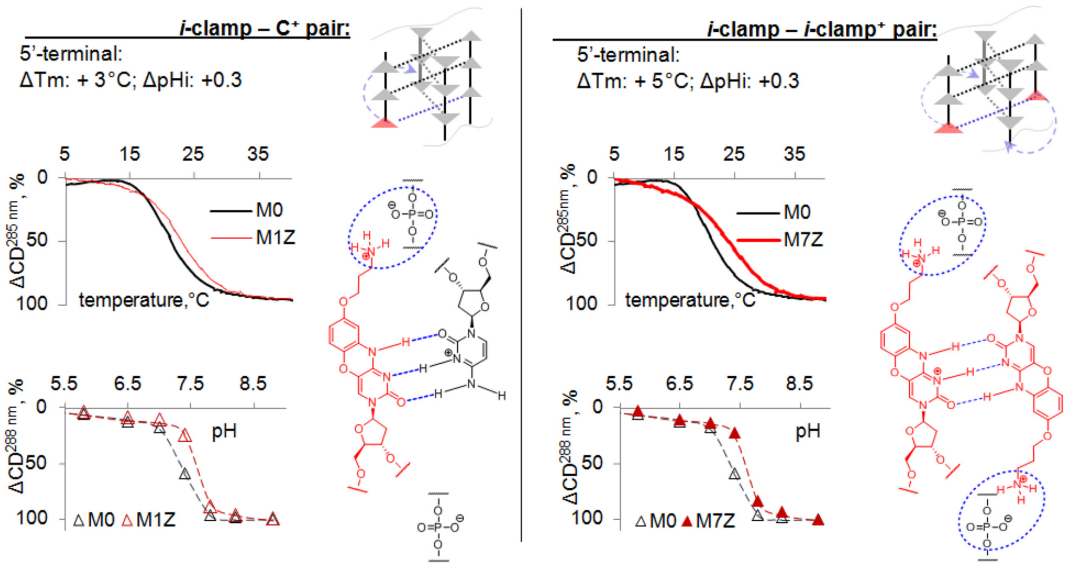


Figure 8. Summary of effects of terminal *i*-clamp insertions in M_0. Melting curves at pH 7.4 (top graphs), CD-based evaluation of transition pH (bottom graphs) and schematic representations of the presumed *i*-clamp interactions in iMs.

Table 2. Sequences and thermal stabilities of native and modified iMs under near-physiological conditions (10 mM sodium-phosphate, pH 7.4, 100 mM NaCl)

Code	Sequence (5′-3′)	T_m, ±1°C (ΔT)	T_a, ±1°C (ΔT)
M_0	CCCCCTTCCCCCTTCCCCCTTCCCCC	20	≤7
M_1Z	**Z**CCCCTTCCCCCTTCCCCCTTCCCCC	23 (+3)	10 (+3)
M_2Z	CC**Z**CCTTCCCCCTTCCCCCTTCCCCC	14 (−6)	≤4 (−3)
M_3Z	CCCCCTTCC**Z**CCTTCCCCCTTCCCCC	13 (−7)	≤4 (−3)
M_4Z	CCCCCTTCCCCCTTCC**Z**CCTTCCCCC	15 (−5)	≤5 (−2)
M_5Z	CCCCCTTCCCCCTTCCCCCTTCC**Z**CC	14 (−6)	≤7 (0)
M_6Z	CCCCCTTCCCCCTTCCCCCTTCCCC**Z**	19 (−1)	≤4 (−3)
M_7Z	**Z**CCCCTTCCCCCTT**Z**CCCCTTCCCCC	25 (+5)	16 (+9)
M_8Z	CC**Z**CCTTCCCCCTTCC**Z**CCTTCCCCC	14 (−6)	≤5 (−2)
M_9Z	CCCCCTTCC**Z**CCTTCCCCCTTCC**Z**CC	13 (−7)	≤7 (0)
$M_{10}Z$	CCCCCTTCCCC**Z**TTCCCCCTTCCCC**Z**	18 (−2)	9 (−1)
G_0	CCCCCCTCCCCCCCTCCCCCCTCCCCCC	26	9
G_1Z	**Z**CCCCCTCCCCCCCTCCCCCCTCCCCCC	28 (+2)	13 (+4)
G_2Z	CCCCCCTCCC**Z**CCCTCCCCCCTCCCCCC	24 (−2)	9 (0)
G_3Z	CCCCCCTCCCCCCCTCCCCCCTCCCCC**Z**	20 (−6)	10 (+1)

Z = *i*-clamp.
T_m – melting temperature, T_a – annealing temperature.

A detailed analysis of pH effects was performed for M_0 and its 5′-modified derivatives M_1Z and M_7Z (Figures 7C and 8). The modified iMs demonstrated slightly sharper transition at higher pH values: $pH_i^{M0} = 7.3 \pm 0.1$; $pH_i^{M1Z, M7Z} = 7.6 \pm 0.1$. Positive effects of the 5′-teminal M_0 modification are summarized in Figure 8. The position dependency of *i*-clamp effects in M_0 and G_0 derivatives is in line with our modelling results obtained for U_0 derivatives. Preference of terminal substitutions over middle-chain ones was shown to be a general trend for both classical and new clamp variants. The fact that 5′-terminal modifications are superior to 3′-terminal ones is probably due to domination of 5′E conformers.

Figure 8 shows presumed *i*-clamp–cytosine$^+$ and *i*-clamp–*i*-clamp$^+$ pairs supported by $U_{1-3}Z$ and T_1Z modeling data. Considering that C-tracks of M_0 and G_0 derivatives are relatively long, certain misalignment in the iM cores cannot be excluded. Thus, it is theoretically possible that 5′-terminal *i*-clamp residues in M_1Z, M_7Z and G_1Z stabilize the iMs through stacking interactions rather than H-bonding. However, the previously described 5′-terminal tether-lacking G-clamp analog (phenothiazine), that reportedly did not participate in base pairing, had a negative effect on iM stability (ΔT_m up to –3°C) (19). Principally different effects (stabilization by 5′-*i*-clamp versus destabilization by phenothiazine) support different interaction patterns and argue indirectly against misalignment in M_1Z, M_7Z and G_1Z.

CONCLUSION

We have shown that G-clamp and guanidino-G-clamp residues incorporated into iM cores could participate in non-canonical base pairing and/or form additional contacts with backbone phosphates via their tethers. The geometries of the G-clamp and its known derivative are suboptimal for simultaneous pairing and tether-backbone interactions according to our molecular modeling data, which explain mostly negative effects of G-clamp insertions on iM thermal stabilities. We have designed the new *i*-clamp analogue with an optimized geometry. Incorporation of the new analogue into several model structures resulted in substantial stabilization of the iM core (ΔT_m average is approximately +5°C per modification) under acidic conditions and, in some cases, in an increase of transition pH values (ΔpH_i up to +0.9). This analogue is a good candidate for the fine tuning of iM properties at an acidic pH, which could be used for iM-based nanodevices for *in vitro* and *in vivo* applications.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

NMR spectra of the low molecular weight compounds were recorded in the joint center of FMI IPCE RAS. We want to thank Mrs. V. Farzan and Ms. K. Nagornova for oligonucleotide synthesis and purification.

FUNDING

Russian Science Foundation [14-25-00013]. Funding for open access charge: Russian Science Foundation [14-25-00013].
Conflict of interest statement. None declared.

REFERENCES

1. Saini,N., Zhang,Y., Usdin,K. and Lobachev,K.S. (2013) When secondary comes first—the importance of non-canonical DNA structures. *Biochimie*, **95**, 117–123.
2. Hänsel-Hertsch,R., Di Antonio,M. and Balasubramanian,S. (2017) DNA G-quadruplexes in the human genome: detection, functions and therapeutic potential. *Nat. Rev. Mol. Cell Biol.*, **18**, 279–284.
3. Kshirsagar,R., Khan,K., Joshi,M.V., Hosur,R.V. and Muniyappa,K. (2017) Probing the potential role of non-B DNA structures at yeast meiosis-specific DNA double-strand breaks. *Biophys. J.*, **112**, 2056–2074.
4. Bacolla,A., Tainer,J.A., Vasquez,K.M. and Cooper,D.N. (2016) Translocation and deletion breakpoints in cancer genomes are associated with potential non-B DNA-forming sequence. *Nucleic Acids Res.*, **44**, 5673–5688.

5. Gelinas,A.D., Davies,D.R. and Janjic,N. (2016) Embracing proteins: structural themes in aptamer-protein complexes. *Curr. Opin. Struct. Biol.*, **36**, 122–132.
6. Varizhuk,A., Ilyinsky,N., Smirnov,I. and Pozmogova,G. (2016) G4 aptamers: trends in structural design. *Mini Rev. Med. Chem.*, **16**, 1321–1329.
7. Cao,Y., Gao,S., Yan,Y., Bruist,M.F., Wang,B. and Guo,X. (2017) Assembly of supramolecular DNA complexes containing both G-quadruplexes and i-motifs by enhancing the G-repeat-bearing capacity of i-motifs. *Nucleic Acids Res.*, **45**, 26–38.
8. Gehring,K., Leroy,J.L. and Gueron,M. (1993) A tetrameric DNA structure with protonated cytosine-cytosine base pairs. *Nature*, **363**, 561–565.
9. 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.
10. Mir,B., Serrano,I., Buitrago,D., Orozco,M., Escaja,N. and González,C. (2017) Prevalent sequences in the human genome can form mini i-motif structures at physiological pH. *J. Am. Chem. Soc.*, **139**, 13985–13988.
11. Takahashi,S., Brazier,J.A. and Sugimoto,N. (2017) Topological impact of noncanonical DNA structures on Klenow fragment of DNA polymerase. *PNAS*, **114**, 9605–9610.
12. Kendrick,S., Kang,H.J., Alam,M.P., Madathil,M.M., Agrawal,P., Gokhale,V., Yang,D.Z., Hecht,S.M. and Hurley,L.H. (2014) The dynamic character of the BCL2 promoter i-motif provides a mechanism for modulation of gene expression by compounds that bind selectively to the alternative DNA hairpin structure. *J. Am. Chem. Soc.*, **136**, 4161–4171.
13. Benabou,S., Avino,A., Eritja,R., Gonzalez,C. and Gargallo,R. (2014) Fundamental aspects of the nucleic acid i-motif structures. *RSC Adv.*, **4**, 26956–26980.
14. 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 stability of DNA i-motifs in the nuclei of living mammalian cells. *Angew. Chem. Int. Ed.*, **130**, 2187–2191.
15. Heinen,L., Heuser,T., Steinschulte,A. and Walther,A. (2017) Antagonistic enzymes in a biocatalytic pH feedback system program autonomous DNA hydrogel life cycles. *Nano Lett.*, **17**, 4989–4995.
16. Shi,L., Peng,P., Du,Y. and Li,T. (2017) Programmable i-motif DNA folding topology for a pH-switched reversible molecular sensing device. *Nucleic Acids Res.*, **45**, 4306–4314.
17. Alba,J.J., Sadurni,A. and Gargallo,R. (2016) Nucleic acid i-motif structures in analytical chemistry. *Crit. Rev. Anal. Chem.*, **46**, 443–454.
18. Dembska,A. and Juskowiak,B. (2015) Pyrene functionalized molecular beacon with pH-sensitive i-motif in a loop. *Spectrochim. Acta A Mol. Biomol. Spectrosc.*, **150**, 928–933.
19. Bielecka,P. and Juskowiak,B. (2015) Fluorescent sensor for pH monitoring based on an i-motif-switching aptamer containing a tricyclic cytosine analogue (tC). *Molecules*, **20**, 18511–18525.
20. Mata,G. and Luedtke,N.W. (2015) Fluorescent probe for proton-coupled DNA folding revealing slow exchange of i-motif and duplex structures. *J. Am. Chem. Soc.*, **137**, 699–707.
21. Sharma,J., Chhabra,R., Yan,H. and Liu,Y. (2007) pH-Driven conformational switch of "i-motif" DNA for the reversible assembly of gold nanoparticles. *Chem. Commun. (Camb.)*, 477–479.
22. Modi,S., Swetha,M.G., Goswami,D., Gupta,G.D., Mayor,S. and Krishnan,Y. (2009) A DNA nanomachine that maps spatial and temporal pH changes in living cells. *Nat. Nanotech.*, **4**, 325–330.
23. Chen,C., Li,M., Xing,Y.Z., Li,Y.M., Joedecke,C.C., Jin,J., Yang,Z.Q. and Liu,D.S. (2012) Study of pH-induced folding and unfolding kinetics of the DNA i-motif by stopped-flow circular dichroism. *Langmuir*, **28**, 17743–17748.
24. Lieblein,A.L., Buck,J., Schlepckow,K., Furtig,B. and Schwalbe,H. (2012) Time-resolved NMR spectroscopic studies of DNA i-motif folding reveal kinetic partitioning. *Angew. Chem. Int. Edit.*, **51**, 250–253.
25. 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.
26. Fleming,A.M., Ding,Y., Rogers,R.A., 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.
27. Fujii,T. and Sugimoto,N. (2015) Loop nucleotides impact the stability of intrastrand i-motif structures at neutral pH. *Phys. Chem. Chem. Phys.*, **17**, 16719–16722.
28. Day,H.A., Pavlos,P. and Waller,Z.A.E. (2014) i-Motif DNA: structure, stability and targeting with ligands. *Bioorg. Med. Chem.*, **22**, 4407–4418.
29. Tatarinova,O., Tsvetkov,V., Basmanov,D., Barinov,N., Smirnov,I., Timofeev,E., Kaluzhny,D., Chuvilin,A., Klinov,D., Varizhuk,A. *et al.* (2014) Comparison of the 'chemical' and 'structural' approaches to the optimization of the thrombin-binding aptamer. *PLoS ONE*, **9**, e89383.
30. Lacroix,L. and Mergny,J.L. (2000) Chemical modification of pyrimidine TFOs: effect on i-motif and triple helix formation. *Arch. Biochem. Biophys.*, **381**, 153–163.
31. Fenna,C.P., Wilkinson,V.J., Arnold,J.R., Cosstick,R. and Fisher,J. (2008) The effect of 2′-fluorine substitutions on DNA i-motif conformation and stability. *Chem. Commun. (Camb.)*, 3567–3569.
32. Robidoux,S. and Damha,M.J. (1997) D-2-deoxyribose and D-arabinose, but not D-ribose, stabilize the cytosine tetrad (i-DNA) structure. *J. Biomol. Struct. Dyn.*, **15**, 529–535.
33. Assi,H.A., Harkness,R.W.t., Martin-Pintado,N., Wilds,C.J., Campos-Olivas,R., Mittermaier,A.K., Gonzalez,C. and Damha,M.J. (2016) Stabilization of i-motif structures by 2′-β-fluorination of DNA. *Nucleic Acids Res.*, **44**, 4998–5009.
34. Kumar,N., Petersen,M. and Maiti,S. (2009) Tunable c-MYC LNA i-motif. *Chem. Commun. (Camb.)*, 1532–1534.
35. Kumar,N., Nielsen,J.T., Maiti,S. and Petersen,M. (2007) i-Motif formation with locked nucleic acid (LNA). *Angew. Chem. Int. Ed. Engl.*, **46**, 9220–9222.
36. Pasternak,A. and Wengel,J. (2011) Modulation of i-motif thermodynamic stability by the introduction of UNA (unlocked nucleic acid) monomers. *Bioorg. Med. Chem. Lett.*, **21**, 752–755.
37. Mergny,J.L. and Lacroix,L. (1998) Kinetics and thermodynamics of i-DNA formation: phosphodiester versus modified oligodeoxynucleotides. *Nucleic Acids Res.*, **26**, 4797–4803.
38. Lee,I.J., Yi,J.W. and Kim,B.H. (2009) Probe for i-motif structure and G-rich strands using end-stacking ability. *Chem. Commun.*, **36**, 5383–5385.
39. Perlíková,P., Karlsen,K.K., Pedersen,E.B. and Wengel,J. (2014) Unlocked nucleic acids with a pyrene-modified uracil: synthesis, hybridization studies, fluorescent properties and i-motif stability. *ChemBioChem*, **15**, 146–156.
40. El-Sayed,A.A., Pedersen,E.B. and Khaireldin,N.A. (2012) Studing the influence of the pyrene intercalator TINA on the stability of DNA i-motifs. *Nucleosides Nucleotides Nucleic Acids*, **31**, 872–879.
41. Stephenson,A.W.I., Partridge,A.C. and Filichev,V.V. (2011) Synthesis of β-pyrrolic modified porphyrins and their incorporation into DNA. *Chem. Eur. J.* **17**, 6227–6238.
42. El-Sayed,A.A., Pedersen,E.B. and Khaireldin,N.Y. (2016) Thermal stability of modified i-motif oligonucleotides with naphthalimide intercalating nucleic acids. *Helv. Chim. Acta*, **99**, 14–19.
43. Gouda,A.S., Amine,M.S. and Pedersen,E.B. (2017) Improved i-motif thermal stability by insertion of anthraquinone monomers. *Org. Biomol. Chem.*, **15**, 6613–6621.
44. Lin,K.-Y. and Matteucci,M.D. (1998) A cytosine analogue capable of clamp-Like binding to a guanine in helical nucleic acids. *J. Am. Chem. Soc.*, **120**, 8531–8532.
45. Ortega,J.A., Blas,J.R., Orozco,M., Grandas,A., Pedroso,E. and Robles,J. (2007) Binding affinities of oligonucleotides and PNAs containing phenoxazine and G-clamp cytosine analogues are unusually sequence-dependent. *Org. Lett.*, **9**, 4503–4506.
46. Reilly,S.M., Lyons,D.F., Wingate,S.E., Wright,R.T., Correia,J.J., Jameson,D.M. and Wadkins,R.M. (2014) Folding and hydrodynamics of a DNA i-motif from the c-MYC promoter determined by fluorescent cytidine analogs. *Biophys. J.*, **107**, 1703–1711.
47. Varizhuk,A.M., Zatsepin,T.S., Golovin,A.V., Belyaev,E.S., Kostyukevich,Y.I., Dedkov,V.G., Shipulin,G.A., Shpakovski,G.V. and Aralov,A.V. (2017) Synthesis of oligonucleotides containing novel G-clamp analogue with C8-tethered group in phenoxazine ring:

Implication to qPCR detection of the low-copy Kemerovo virus dsRNA. *Bioorg. Med. Chem.*, **25**, 3597–3605.
48. Wilds,C.J., Maier,M.A., Manoharan,M. and Egli,M. (2003) Structural basis for recognition of guanosine by a synthetic cytosine analogue: guanidinium G-clamp. *Helv. Chim. Acta.*, **86**, 966–978.
49. Apffel,A., Chakel,J.A., Fischer,S., Lichtenwalter,K and Hancock,W.S. (1997) Analysis of oligonucleotides by HPLC-electrospray ionization mass spectrometry. *Anal. Chem.*, **69**, 1320–1325.
50. Esmaili,N. and Leroy,J.L. (2005) i-Motif solution structure and dynamics of the d(AACCCC) and d(CCCCAA) tetrahymena telomeric repeats. *Nucleic Acids Res.*, **33**, 213–224.
51. Phan,A.T., Gueron,M. and Leroy,J.L. (2000) The solution structure and internal motions of a fragment of the cytidine-rich strand of the human telomere. *J. Mol. Biol.*, **299**, 123–144.
52. Protopopova,A.D., Tsvetkov,V.B., Varizhuk,A.M., Barinov,N.A., Podgorsky,V.V., Klinov,D.V. and Pozmogova,G.E. (2018) The structural diversity of C-rich DNA aggregates: unusual self-assembly of beetle-like nanostructures. *Phys. Chem. Chem. Phys.*, **20**, 3543–3553.
53. Frisch,M.J., Trucks,G.W., Schlegel,H.B., Scuseria,G.E., Robb,M.A., Cheeseman,J.R., Scalmani,G., Barone,V., Mennucci,B., Petersson,G.A. *et al.* (2009) *Gaussian 09, Revision A.1*, Gaussian, Inc., Wallingford.
54. Case,D.A., Babin,V., Berryman,J.T., Betz,R.M., Cai,Q., Cerutti,D.S., Cheatham,T.E. III, Darden,T.A., Duke,R.E., Gohlke,H. *et al.* (2014) *AMBER 14*. University of California, San Francisco.
55. Krepl,M., Zgarbová,M., Stadlbauer,P., Otyepka,M., Banas,P., Koca,J., Cheatham,T.E. III and Sponer,J. (2012) Reference simulations of noncanonical nucleic acids with different chi variants of the AMBER force field:Quadruplex DNA, quadruplex RNA, and Z-DNA. *J. Chem. Theory Comput.*, **8**, 2506–2520.
56. Zgarbová,M., Luque,F.J., Sponer,J., Cheatham,T.E. III, Otyepka,M. and Jurečka,P. (2013) Toward improved description of DNA backbone: Revisiting epsilon and zeta torsion force field parameters. *J. Chem. Theory Comput.*, **9**, 2339–2354.
57. Zgarbová,M., Sponer,J., Otyepka,M., Cheatham,T.E. III, Galindo-Murillo,R. and Jurečka,P. (2015) Refinement of the sugar-phosphate backbone torsion beta for AMBER force fields improves the description of Z- and B-DNA. *J. Chem. Theory Comput.*, **12**, 5723–5736.
58. Humphrey,W., Dalke,A. and Schulten,K. (1996) VMD: Visual molecular dynamics. *J. Mol. Graph. Model.*, **14**, 33–38.
59. Tsvetkov,V.B., Varizhuk,A.M., Pozmogova,G.E., Smirnov,I.P., Kolganova,N.A. and Timofeev,E.N. (2015) A universal base in a specific role: tuning up a thrombin aptamer with 5-nitroindole. *Sci. Rep.*, **5**, 16337.