Unveiling the unusual i-motif-derived architecture of a DNA aptamer exhibiting high affinity for influenza A virus

Vladimir Tsvetkov [1,2,†], Bartomeu Mir[3,†], Rugiya Alieva[4], Alexander Arutyunyan[5], Ilya Oleynikov[5], Roman Novikov[6], Elizaveta Boravleva[7], Polina Kamzeeva[8], Timofei Zatsepin[4], Andrey Aralov [8,*], Carlos González [3,*] and Elena Zavyalova [4,*]

[1]Center for Mathematical Modeling in Drug Development, Sechenov First Moscow State Medical University, Moscow 119991, Russia
[2]Department of Cell Biology, Lopukhin Federal Research and Clinical Center of Physical-Chemical Medicine, Moscow 119435, Russia
[3]Instituto de Química Física Blas Cabrera, CSIC, Madrid 28006, Spain
[4]Chemistry Department, Lomonosov Moscow State University, Moscow 119991, Russia
[5]Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Moscow 119991, Russia
[6]Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow 119991, Russia
[7]Chumakov Federal Scientific Center for Research and Development of Immunobiological Products (Institute of Poliomyelitis), Russian Academy of Sciences, Moscow 108819, Russia
[8]Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow 117997, Russia

*To whom correspondence should be addressed. Tel: +7 495 939 5418; Fax: +7 495 939 3181; Email: zlenka2006@gmail.com
Correspondence may also be addressed to Carlos González. Tel: +34 91 745 9533; Fax: +34 91 745 9533; Email: cgonzalez@iqfr.csic.es
Correspondence may also be addressed to Andrey Aralov. Tel: +7 926 606 2910; Fax: +7 926 606 2910; Email: baruh238@mail.ru
†These authors contribute equally.

Abstract

Non-canonical nucleic acid structures play significant roles in cellular processes through selective interactions with proteins. While both natural and artificial G-quadruplexes have been extensively studied, the functions of i-motifs remain less understood. This study investigates the artificial aptamer BV42, which binds strongly to influenza A virus hemagglutinin and unexpectedly retains its i-motif structure even at neutral pH. However, BV42 conformational heterogeneity hinders detailed structural analysis. Molecular dynamics simulations and chemical modifications of BV42 helped us to identify a potential binding site, allowing for aptamer redesign to eliminate the conformational diversity while retaining binding affinity. Nuclear magnetic resonance spectroscopy confirmed the i-motif/duplex junction with the three-cytosine loop nearby. This study highlights the unique structural features of the functional i-motif and its role in molecular recognition of the target.

Graphical abstract
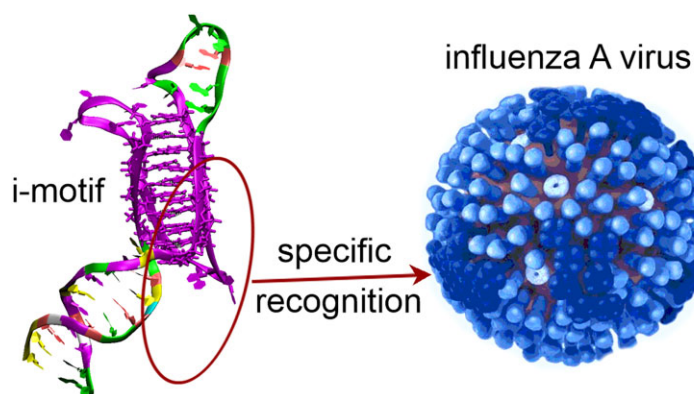
Introduction

Nucleic acid aptamers form complex 3D structure that exhibit high affinity and specificity towards various molecular targets. Aptamers are chemically synthesized and can be modified at specific sites with various functional groups enriching the alphabet of nucleic acids (1). The conformation of aptamers is susceptible to reversible thermal denaturation; also their activity can be regulated with toehold oligonucleotide strands.

These unique properties make aptamers valuable tools in the fields of therapeutics and molecular diagnostics (2–6).

Initial selection of aptamers using systematic evolution of ligands by exponential enrichment (SELEX) is usually followed up with a maturation step tuning affinity and selectivity. Introduction of non-natural nucleotides make additional contacts with the target reducing dissociation constants (K_D) to 2–20 pM (7–11). Moreover, structural optimization plays a pivotal role in enhancing the fraction of favorable aptamer conformations and further lowering the K_D values of aptamer–protein complexes up to 10–100-fold (12,13).

Typically, aptamers contain stem-loop structures with bulges, non-canonical base pairs, or non-canonical secondary structures, thus creating a unique surface for target interaction (14–17). Non-canonical structures, such as triplexes and G-quadruplexes (G4s), can be key determinants for target recognition (14,17).

Here we want to focus on i-motif (iM) – an alternative non-canonical secondary structure that is formed by cytosine-rich sequences that fold into two parallel duplexes intercalated in antiparallel orientation and stabilized by hemi-protonated cytosine-cytosine$^+$ (C:C$^+$) base pairs (Figure 1A and B). First, iMs were artificial structures formed under acidic conditions by C-rich oligonucleotides. The transition pH values (pH$_i$), at which 50% of the oligonucleotide is folded into the iM structure, typically was in from 5 to 6 range (18–20). However, iMs with extended cytosine tracts (C-tracts) can remain stable even at neutral pH, with a pH$_i$ as high as 7.9 (21,22).

Evidences for iM stability under physiological conditions *in vitro* (23) and their interactions with cellular proteins (24–27) inspired genome-wide iM mapping. Data obtained with iM-specific antibody (iMab) (28) supported iM association with transcriptional regulation in living cells (29). However, iMab-based detection method requires cell membrane permeabilization and, more importantly, incubation at low temperatures, which can lead to significant overestimation of iM contribution. Indeed, in-cell nuclear magnetic resonance (NMR) studies revealed that certain iMs can form within living cells at 37°C but only at a minute fraction of genomic sites (30).

Functional iMs can be used in DNA nanotechnology, mainly as inter- or intramolecular pH sensors with an adjustable working range and intracellular localization (31–34). Introduction of iMs in aptamers allows pH-depended affinity control for bioimaging (35–39) and drug release (40–47). In addition, electrochemical sensors with iMs were used for detection of pH changes during enzymatic reactions (48). Finally, hydrogels for pH-triggered release of therapeutics have been prepared (49–51).

Musafia et al. reported several DNA aptamers that bind to the hemagglutinin protein of the influenza virus (52). These aptamers contain long C-rich motif that have the potential to fold into iMs, along with complementary sequences that form a duplex structure between 5′- and 3′-ends (Figure 1C). The leading aptamer, BV42, was shown to inhibit the binding of the influenza A virus to host cells and subsequent infection at low-nanomolar concentrations. Notably, all biological experiments were conducted in a phosphate-buffered saline (PBS) solution with neutral pH (52). Given that BV42 has a pH$_i$ of pH 7.1 (53), the necessity of the iM structure for the functioning of this aptamer was called into question.

This study aims to clarify the role of the iM structure in hemagglutinin binding and identify the putative binding site using a combination of experimental and computational techniques. To address the inherent structural heterogeneity of BV42, a novel approach was developed to limit conformational variability while preserving the binding site, resulting in an aptamer with both high affinity for hemagglutinin and structural homogeneity. Using NMR methods, BV42 structure was successfully elucidated, revealing a junction between the i-motif and duplex regions, highlighting a unique aptamer architecture critical for protein recognition.

Materials and methods

Reagents

Inorganic salts, acids, alkaline and tris were purchased from AppliChem GmbH (Darmstadt, Germany). 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide hydrochloride (EDC-HCl) from Roth (Karlsruhe, Germany) and sodium salt of N-hydroxysulfosuccinimide (s-NHS) from Chem-Impex Int'l (Wood Dale, IL, USA) were used. Glutaric aldehyde was purchased from Ruschim (Moscow, Russia). Recombinant hemagglutinin (HA1 subunit) from A/Anhui/1/2013 (H7N9) strain (ab190421, Abcam, Burlingame, CA, USA) and fetuin (Sigma-Aldrich, St. Louis, MO, USA) was used. HEPES-HCl buffer (1 M, pH 7.5) and AgNO$_3$ were purchased from AppliChem GmbH (Darmstadt, Germany) and Sigma-Aldrich (St. Louis, MO, USA), respectively.

Buffers with different pH were prepared based on conventional phosphate buffered saline (PBS, 10 mM Na$_2$HPO$_4$, 1.8 mM KH$_2$PO$_4$, 137 mM NaCl, 2.7 mM KCl pH 7.35) with addition of NaOH or HCl solutions. The target pH values were pH 6.0, 6.5, 7.0, 7.3 and 8.0. pH values were controlled with ST20 pH-meter (Ohaus Corporation, Parsippany, New Jersey, United States). All solutions were prepared using ultra-pure water produced by Millipore (Merck Millipore, Burlington, Massachusetts, United States).

Oligonucleotides and sample preparation

Oligonucleotides were synthesized using commercially available reagents by a solid-phase phosphoramidite method, followed by purification with high performance liquid chromatography. The 1,3-diaza-2-oxophenoxazine (tC^o) phosphoramidite was prepared according to the reported procedure (54). The sequences and sites of modification are provided in Supplementary Table S5. Typically, 1 or 2 μM solution of the aptamer in PBS buffer (pH 6.0, 6.5, 7.0, 7.3 or 8.0) was heated at 95°C for 5 min. The solutions were used after gradual cooling to room temperature. For the affinity experiments, the solutions were diluted with PBS buffer with the same pH. For the experiments with Ag$^+$ ions, the AgNO$_3$ solution was added to 1.3 μM BV42 aptamer solution in PBS buffer (pH 7.0, 7.3 or 8.0), the sample was incubated for 5 min at room temperature. The final concentrations of AgNO$_3$ were in the range of 0.6–20 μM.

Circular dichroism and UV spectroscopy

Around 1–2 μM aptamer solutions in PBS buffer at pH 6.0–8.0 were placed into quartz cuvettes with 1 cm path. Circular dichroism (CD) and UV spectra were acquired using Chirascan CD spectrometer (Applied Photophysics, Leatherhead, Great Britain) and MOS-500 spectrometer (BioLogic, Seyssinet-Pariset, France) equipped with a thermoelectric temperature regulator. The spectra were acquired in the wavelength range of 230–360 nm. The spectrum of the buffer was
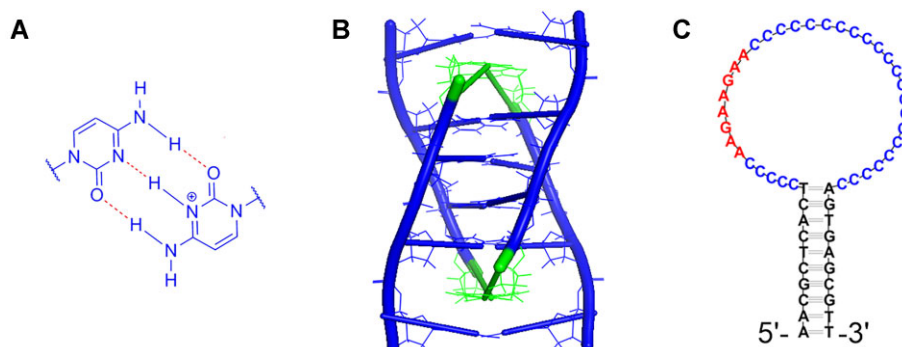
Figure 1. Non-canonical DNA structure, iM, is composed of hemi-protonated cytosine pairs (**A**) that form intercalated DNA duplexes (**B**). The strands in the duplexes are in parallel orientation; 5′-ends of DNA are colored in green. The sequence of DNA aptamer, BV42, with a long C-tract that might form the iM structure (**C**).

subtracted as the baseline. The samples were heated with a mean ramp of $1.0°C/min$. The melting experiments were conducted in the temperature range of 10–85°C. The melting temperatures were derived from the temperature dependencies of molar CD intensity at the wavelength of 285 nm and UV absorption at 260 nm. The curves were approximated with the Boltzmann function, $y = A2 + (A1 − A2) / (1 + exp ((x-x0)/dx))$ using OriginPro software (OriginLab, Northampton, MA, USA), and the parameter x0 corresponded to the value of the melting temperature. The fitting results are shown in Supplementary Materials.

NMR

^{1}H NMR spectra of BV42 were recorded using Bruker AVANCE III HD 300 and Bruker AVANCE III HD 400 spectrometers with 300.1 and 400.1 MHz, respectively (Bruker, MA, USA). [1]H chemical shifts were referenced relative to external sodium 2,2-dimethyl-2-silapentane-5-sulfonate. ^{1}H NMR spectra of samples in H_2O (90%) / D_2O (10%) buffer (phosphate-citrate buffer with 100 mM Na_2HPO_4 and 50 mM citric acid at pH 4.7 or 6.0) were recorded in the temperature range from 5°C to 80°C using the WATERGATE-W5 pulsed-field gradient pulse sequence (zggpw5 from Bruker library) for complete H_2O-signal clear suppression. A relaxation delay of 2.0 sec and an acquisition time of 1.5 s were used for all experiments. Starting solutions of DNA aptamer were annealed in all cases for 5 min at 95°C and cooled at 5°C temperature. All samples were diluted by D_2O to the final content of 10%.

Samples for NMR experiments were dissolved in 9:1 H_2O/D_2O (10 mM sodium phosphate buffer). Experiments were carried out in neutral or slightly acidic conditions (pH 6–7). The pH was adjusted by adding aliquots of concentrated solution of either DCl or NaOD. NMR spectra were acquired on Bruker spectrometers operating at 600 or 800 MHz, equipped with cryoprobes and processed with TOPSPIN software (Bruker, MA, USA). NOESY (nuclear Overhauser effect spectroscopy) spectra in 9:1 H_2O/D_2O were acquired with mixing times of 100 and 200 ms. TOCSY spectra were recorded with 80 ms mixing time.

Fluorescence melting experiments

Around 2 μM tC°-modified aptamer solutions in PBS buffer at pH 6.0–8.0 were placed in a 96-well black non-binding microplate with clear bottom (Greiner Bio-One GmbH, Frick-

enhausen, Germany). The fluorescence spectra were obtained using CLARIOstar (BMG Labtech, Ortenberg, Germany). The excitation wavelength was 370 nm, the emission wavelengths were in the range of 410–550 nm. The melting experiment was performed by gradual increase in temperature from 20°C to 65°C with a mean ramp of 0.7°C/min. The curves were approximated with the Boltzmann function, $y = A2 + (A1 − A2) / (1 + exp ((x-x0)/dx))$ using OriginPro software (OriginLab, Northampton, MA, USA), and the parameter x0 corresponded to the value of the melting temperature.

Viruses

Influenza viruses and allantoic fluids were provided by the Chumakov Federal Scientific Center for Research and Development of Immune and Biological Products of the Russian Academy of Sciences. Influenza A viruses (strains A/chicken/Rostock/45/1934, H7N1 and A/duck/Moscow/4524/2011, H3N2), as well as Newcastle disease virus (NDV) were studied. Viral stocks were propagated in the allantoic cavity of 10-day-old embryonated specific pathogen-free chicken eggs. The eggs were incubated at 37°C for 48 h after infection, cooled to 4°C, kept for 16 h and then harvested. The study design was approved by the Ethics Committee of the Chumakov Institute of Poliomyelitis and Viral Encephalitides, Moscow, Russia (Approval #4 from 2 December 2014). Viruses were inactivated by the addition of 0.05% (v/v) glutaric aldehyde, preserved by the addition of 0.03% (w/v) NaN_3 and stored at + 4°C.

The activity of the viruses was determined by a hemagglutination assay according to Killian et al. ([55]). 50 μl of solutions of the viruses diluted two times step-by-step in PBS was placed into a V-shaped 96-well microtiter plate in a volume of 50 μl. Then, 50 μl of 0.5% chicken red blood cells (RBCs) in PBS were added to the well. The plate was kept in the refrigerator at 4°C for 1 h. Then, the hemagglutination titer was estimated as the maximal dilution of the virus that did not cause the precipitation of RBCs; this well contained 1 HAU (hemagglutination unit) of the virus in the probe. Influenza A virus had hemagglutination activity of 80 000 HAU/mL, and NDV had hemagglutination activity of 2000 HAU/mL.

The concentrations of viruses were determined with Nanoparticle Tracking Analysis (NTA). NTA was conducted using a ZetaView® PMX420-QUATT instrument (Particle Metrix GmbH, Germany), while the data were analyzed by ZetaView NTA software. Operating instructions of the manu-

facturer were followed before calibrating the instrument with a known concentration of 100 nm polystyrene nanoparticles (Applied Microspheres B.V., the Netherlands). The standards were suspended in particle-free water, whereas the investigated samples were diluted 1:100 with PBS at pH 7.3. Particles were counted and size-distributed at 10 cycles of 11 frames per cycle under sensitivity of 65 and a shutter value of 100. Influenza A virus contained $6 \cdot 10^9$ VP/mL (viral particle per mL), and NDV contained $4 \cdot 10^{10}$ VP/mL.

Affinity assays with influenza A virus

To study the affinity to influenza viruses, 5′-biotinylated aptamers were immobilized onto streptavidin-coated biosensors (Octet SAX biosensors, ForteBIO, Menlo Park, CA, USA) by incubating the sensor in 1 μM aptamer solution in PBS (pH 6.0 or 8.0) for 5 min. Viral particles of A/chicken/Rostock/45/1934 (H7N1) strain or NDV were prepared in the PBS (pH 6.0 or 8.0) by diluting the virus 5, 10 or 20 times. After signal stabilization in the PBS at pH 6.0–8.0, the association step was conducted with virus solution in the same PBS buffer. Then the dissociation step was conducted in the same PBS buffer. Both the association and the dissociation stages were monitored for 200 s. Regeneration of the sensors was conducted by incubation in PBS buffer at pH 8.0 for 5 min. The design of the experiment provided zero reference signal, so the sensorgrams were processed without subtraction with baseline correction only. Each experiment was performed twice.

Sensorgram processing was performed using OriginPro software (OriginLab, Northampton, MA, USA) using the Langmuir 1:1 binding model recommended for affinity estimations ([56]). Values of the association (k_{ass}) and dissociation (k_{diss}) rate constants were calculated; then, equilibrium dissociation constants K_D were calculated as the ratio of the two rate constants ($K_D = k_{diss} / k_{ass}$).

Alternatively, binding of aptamers to influenza A viruses was studied using BV42 aptamer modified with cyanine 3 (Cy3) at 3′-end. The influenza A virus (H3N1 subtype) in allantoic fluid was filtered through track membrane with 40 nm pores using Elveflow microfluidic device equipped with MFS flow sensor (Elveflow, Paris, France). The membrane was washed with three volumes of PBS (pH 6.0 or 8.0). Then the membrane was placed into a tube with one volume of PBS (pH 6.0 or 8.0). This procedure eliminates the impurities and changes pH in the solution. The activity of the virus was determined with hemagglutination assay.

The 96-well transparent plates with flat bottom (MedPolymer, Saint-Petersburg, Russia) were incubated with 200 μL of 10 μg/mL fetuin solution for 12 h at room temperature. Then the wells were washed with water two times and dried in air. Around 10 μL of influenza A virus (H3N1 subtype, 640 HAU/mL, pH 6.0 or 8.0) was added in fetuin-treated plate and incubated for 30 min at room temperature. Each well was washed three times with 100 μL of PBS (pH 6.0 or 8.0) with 0.01% Tween-20. Next, 100 μL of BV42-Cy3 aptamer in a concentration of 0, 12.5, 25, 50, 100 or 200 nM in PBS (pH 6.0 or 8.0) with 0.01% Tween-20 was incubated with fetuin-virus complex for 30 min at room temperature. Each well was washed three times with 100 μL of PBS pH 6.0 or 8.0 with 0.01% Tween-20. Then 100 μL of water was added to each well. The fluorescence was registered using CLARIOstar (BMG Labtech, Ortenberg, Germany) with

excitation wavelength of 545 nm and emission wavelength of 620 nm.

A competition between BV42 and CapTr4C3 aptamers was studied in the following way. After virus binding to fetuin, unlabeled CapTr4C3 aptamer was added in a concentration 0, 50, 100, 200 or 400 nM in PBS (pH 6.0 or 8.0) with 0.01% Tween-20. After 30-min incubation, 100 nM solution of BV42-Cy3 aptamer was added for 30 min at room temperature. Each well was washed three times with 100 μL of PBS (pH 6.0 or 8.0) with 0.01% Tween-20. Then 100 μL of water was added to each well. The fluorescence was registered using CLARIOstar (BMG Labtech, Ortenberg, Germany) with excitation wavelength of 545 nm and emission wavelength of 620 nm.

The inhibitory activities of BV42 and CapTr4C3 were estimated using the modified version of hemagglutination assay. RBCs were purified; the solution was centrifuged at 2400 rpm for 5 min; then the supernatant was discarded and the same volume of PBS (pH 6.0 or 8.0) was added. The procedure was repeated 3 times. The aptamers were folded in PBS (pH 6.0 or 8.0). 50 μL of 2 HAU of influenza A virus in the buffer was mixed with 50 μL of aptamer in a concentration of 32, 64, 128, 256, 512, 1024 or 2048 nM in the same buffer in U-shaped 96-well plate. After 10 min incubation, the 50 μL of RBC in the same buffer was added. The mixture was kept at + 4°C for 1–2 h until RBC dots are formed in the control wells (without influenza A virus). Then the absorbance at 550 nm was recoded using CLARIOstar reader (BMG Labtech, Ortenberg, Germany). Three independent repeats were conducted.

Electrophoretic mobility shift assay was performed with BV42 and CapTr4C3 aptamers labeled with cyanine-3. The aptamers were folded in PBS (pH 6.0 or 8.0) at 1 μM concentration. Then the solutions were diluted with the same buffer to 100 nM concentration. The aptamers were incubated with influenza A virus in PBS (pH 6.0 or 8.0), respectively. The final concentration of the aptamer was 10 nM; the final concentrations of influenza A virus was $3.4 \cdot 10^9$ VP/mL, $3 \cdot 10^9$ VP/mL, $2.6 \cdot 10^9$ VP/mL, $2.3 \cdot 10^9$ VP/mL, $1.5 \cdot 10^9$ VP/mL, $7.5 \cdot 10^8$ VP/mL and 0 VP/mL. The incubation time was 15 min. Then, 6x loading dye (cat. #R0611, Fermentas, Hanover, MD, USA) was added. The sample was separated in 12% polyacrylamide gel (1:20 N,N-methylenebis(acrylamide) : acrylamide). The reagents for polyacrylamide gel were purchased from Helicon (Moscow, Russia). The samples at pH 6.0 were separated in a gel prepared in citrate-phosphate buffer (pH 6.0). The samples at pH 8.0 were separated in a gel prepared in tris-borate-EDTA buffer (pH 8.3). The gels were scanned with Typhoon (GE Healthcare, Chicago, IL, USA) and processed with ImageMaster 2D Platinum 7.0 software (GE Healthcare, Chicago, IL, USA).

Affinity assays with recombinant hemagglutinin

The affinity of the aptamers to recombinant hemagglutinin and influenza viruses was estimated using biolayer interferometry (Blitz, ForteBIO, Menlo Park, CA, USA) at 20°C. Samples were placed into black 0.5 mL tubes (Sigma-Aldrich, St. Louis, MO, USA) in a 220 μL volume.

To study the affinity to recombinant hemagglutinin, biosensors intended for the amine coupling reaction (Octet AR2G biosensors, ForteBio, Menlo Park, CA, USA) were activated for 5 min in a solution of 200 mM EDC-HCl and 100 mM

s-NHS. Then, hemagglutinin was loaded from a 10 μL drop of 15 μg/mL solution in PBS at pH 6.0 for 10 min. The sensor was washed and blocked with 100 mM HEPES-HCl buffer at pH 7.5 for 3 min. After signal stabilization in the PBS (pH 6.0 or 8.0), the association step was conducted with aptamer solution in the concentration 1000, 500, 250 or 125 nM in the same PBS buffer. Then, the dissociation step was conducted in the same PBS buffer. Both the association and the dissociation stages were monitored for 200 s. Regeneration of the sensors was conducted by incubation in PBS at pH 8.0 for 5 min. The design of the experiment provided zero reference signal, so the sensorgrams were processed without subtraction with baseline correction only. Each experiment was performed twice.

Sensorgram processing was performed using OriginPro software (OriginLab, Northampton, MA, USA) using the Langmuir 1:1 binding model recommended for affinity estimations (56). Values of the association (k_{ass}) and dissociation (k_{diss}) rate constants were calculated from the exponential fit of the curves with equation *Signal = A × (1 − exp(−B t))*, where *Signal* is a shift (nm), *A* is a plateau value and $B = k_{ass} C_{apt} + k_{off}$, C_{apt} is the aptamer concentration at the stage. Then, equilibrium dissociation constants K_D were calculated as the ratio of the two rate constants ($K_D = k_{diss} / k_{ass}$).

Molecular modeling

The aptamer models were built based on the strategy developed by Tsvetkov V.B. and described in the work (57) using Sybil-X software (Certara, USA). At the first stage, duplex and iM models were created. The iM's core comprising ten cytosine-cytosine pairs was obtained based on the reported tetramolecular iM models (PDB 1YBL). Then they were joined to each other, then loops were added to the iM. At each stage of docking, molecular model optimization was carried out using the SYBYL X and the Powell method with the following settings: partial charges and parameters for interatomic interactions were from the Amber7ff02 force field, a non-bonded cut-off distance was set to 8 Å, a distance-dependent dielectric function was applied, the number of iterations was equal to 500, the simplex method was used in the initial optimization and the energy gradient convergence criterion of 0.05 kcal/mol/Å was used. Partial charges on 1,3-diaza-2-oxophenoxazine atoms were calculated according to the following scheme. First, calculation of electron density distribution was performed by DFT using the hybrid meta functional M06-2X (58) and 6–311G+(d,p) Pople basis sets. Then, the Merz–Singh–Kollman scheme was applied to the electron density distribution obtained to calculate the grid for the electrostatic potential fitting with the following parameters: (6/41 = 10) – the number of surfaces around the atoms and (6/42 = 17) – the density of test points on these surfaces. The Restrained ElectroStatic Potential method was applied to the fitting of the grid obtained in the previous step to calculate partial atomic charges. All quantum mechanics simulations were performed using Gaussian 16 software (Gaussian, Inc., Wallingford, CT, USA) (59).

The models were verified by molecular dynamics (MD) simulations using Amber 20 software (AMBER Software, San Francisco, CA, USA) (60). The influence of the solvent was simulated using the OPC3 water model (61). Rectangular box and periodic boundary conditions were used in the simulation. The space between the models and the periodic box wall

was at least 15 Å. Potassium ions were used to neutralize the negative charge of the DNA backbone. The parameters needed for interatomic energy calculation were taken from the force fields OL15 (62,63) for the DNA and gaff2 for 1,3-diaza-2-oxophenoxazine. At the beginning of the computation, the models with thrombin were optimized in two stages. First, the location of the solvent molecules was optimized by using 1000 steps (500 steps of steepest descent followed by 500 steps of conjugate gradient). At this stage, the mobility of all solute atoms was restrained by a force constant of 500 kcal $\times$ mol$^{-1}\times$Å^{-2}. In the second stage, the optimization was performed without restrictions using 2500 steps (1000 steps of steepest descent and 1500 steps of conjugate gradient). Then, gradual heating to 300 K was carried out for 20 ps. To avoid spontaneous fluctuations at this point, weak harmonic restraints were applied with a force constant of 10 kcal $\times$ mol$^{-1}\times$Å^{-2} for all atoms other than the solvent ones. The SHAKE algorithm was applied to constrain hydrogen-containing bonds, which allowed the use of a 2 fs time step. Scaling of 1–4 non-bonded van der Waals and electrostatic interactions was performed by using the standard Amber values. The cut-off distance for nonbonded interactions was set to 10 Å, and the long-range electrostatics were calculated using the particle mesh Ewald method. The MD simulations in the production phase were performed using constant temperature (T = 300 K) and constant pressure (p = 1 atm) over 80 ns. Each experiment was performed at once; the statistical data were analyzed for overall time lapse. The data are shown in Supplementary materials. To control the temperature, a Langevin thermostat was used with a collision frequency of 1 ps^{-1}. The energy of interaction of bases of cytosines in the in C-C pairs was estimated using the lie command from the cpptraj module with a dielectric constant of 2. Free energy was calculated as the sum of the electrostatic energies (E_q), Van der Waals energies (E_{VDW}), energy of solvation and deformation energy of valence bonds, valence and dihedral angles (U). The energy of solvation was calculated as the sum of the polar and nonpolar contributions. The polar contribution (E_{GB}) was computed using the Generalized Born method and the algorithm developed by Onufriev et al. for calculating the effective Born radii (64). The non-polar contribution to the solvation energy (E_{surf}), which includes solute-solvent van der Waals interactions and the free energy of cavity formation in solvent, was estimated from solvent-accessible surface area. The plots of geometrical parameters and energy of interaction versus time were smoothed using the moving average method (span = 5). The histogram for deviations of the distances between centers of mass of the cytosine bases and angles between normals to the bases from their average values over the trajectory were constructed in the following way. 100 bins(intervals) of equal length between the minimum and maximum values for each considered value were plotted on the abscissa axis, and the number of hits in a particular bin (interval) was plotted on the ordinate axis.

Results

BV42 affinity for hemagglutinin and influenza A virus is pH dependent

To determine the role of iM structure in protein recognition, BV42 affinity for hemagglutinin was assessed at varying pH.

Recombinant hemagglutinin from subtype H7N1 of influenza A virus was immobilized on the sensor, and a complex formation between the aptamer and the protein in PBS at pH 6.0 (folded iM) or 8.0 (unfolded iM) was monitored in real time mode using biolayer interferometry. BV42 exhibited high affinity at pH 6.0 (Figure 2A) with a K_D of 1.4 ± 0.4 nM at pH 6.0. Conversely, at pH 8.0, K_D increased 125-fold to 176 ± 9 nM (Supplementary Figure S1A). To confirm hemagglutinin conformational integrity at pH 8.0, a similar experiment using the G-quadruplex-based hemagglutinin aptamer RHA0385 (65) showed comparable binding behavior at both pH 6.0 and 8.0 (Supplementary Figure S1B). These results confirm the necessity of iM structure for BV42 functional activity.

The binding affinity of the aptamer to influenza A virus also demonstrated pH dependence. The BV42:influenza A virus of subtype H7N1 complex K_D was 0.010 ± 0.002 pM at pH 6.0 (Figure 2B). No affinity was observed for either the off-target NDV at pH 6.0 or influenza A virus of subtype H7N1 at pH 8.0 (Supplementary Figure S1C). The extremely high affinity to virions aligns with the recombinant protein affinity considering each viral particle carries approximately 1000 hemagglutinin molecules and the multipoint binding to the aptamer-functionalized surface enhances avidity (66).

Fluorometric binding assay further corroborated the pH-dependence. Cyanine-3 labeled BV42 exhibited higher affinity for influenza A virus (H3N1) adsorbed onto fetuin-coated plates at pH 6.0 (Supplementary Figure S2). Electrophoretic mobility shift assays with H7N1 influenza A virus at pH 6.0 also supported these findings. Increasing viral concentrations decreased the free labeled BV42 band intensity, with an apparent K_D of approximately $2 \cdot 10^9$ VP/mL (Supplementary Figure S3), equivalent to 3 nM hemagglutinin, considering the viral particle's hemagglutinin content. Repeating the experiment at pH 8.0 showed no concentration dependence of the free aptamer fraction, further confirming the aptamer's pH-dependent activity.

Inhibitory activity of BV42 was then assessed using a hemagglutination assay at pH 6, 7 and 8. RBCs typically sediment, forming a dot in V-shaped or a small ring in U-shaped wells. Influenza A virus disrupts this sedimentation by forming a gel-like lattice where viruses cross-link RBCs. The absence of the RBC dot indicates viral activity, measured in hemagglutination units. Antivirals that disrupt virus-RBC interaction can be evaluated using this assay. Antiviral presence restores the RBC dot due to virus neutralization, allowing IC_{50} calculation based on dot size (67). For this assay, influenza A virus (H3N1) was purified by 40 nm track-etched membrane filtration and RBCs were separately resuspended in PBS (pH 6.0 or 8.0) to control pH and eliminate off-target biomolecules. BV42 (32–2048 nM) was pre-incubated with the virus for 10 min before adding RBCs. At pH 6.0, increasing BV42 concentrations neutralized virus-induced hemagglutination inhibition, and no antiviral activity was observed at pH 8.0 (Supplementary Figure S4). While dots were visible even at 32 nM BV42 at pH 6.0, absorbance measurements revealed an IC_{50} of 130 ± 50 nM at pH 6.0 and >2.0 μM at pH 8.0.

In summary, BV42 demonstrated pH-dependent binding affinity and viral attachment blocking activity.

i-motif formation is necessary for the aptamer function

Typical iMs assemble under slightly acidic conditions (pH 5–6) (18–20). However, CD spectroscopy confirmed BV42 to form a stable iM even at neutral pH. CD spectra of BV42 at pH 6.0–7.0 have a positive band with a maximum at 285 nm and a negative band with a minimum at 263 nm (Figure 2C and Supplementary Figure S5), characteristic of iMs (68). At pH 8.0, the CD spectrum resembles an unstructured oligonucleotide (Supplementary Figure S5). The observed pH_i of 7.3 aligns with iMs containing extended C-tracts, such as the iM from the HIF-1α promoter region ($pK_a = 7.2$) with C3-C5-C4-C4 cytosine tracts (21) or an artificial iM ($pK_a = 7.9$) with C4-C3-C4-C3 cytosine tracts (22).

CD spectroscopy revealed a monotonic decrease in iM melting temperature (T_m) with increasing pH (Supplementary Table S1). The i-motif remained stable even at neutral pH, with T_m of 43.7°C at pH 6.5, 32.4°C at pH 7.0 and 26.1°C at pH 7.3 (Supplementary Figures S6 and S7). UV spectroscopy yielded similar T_m values (Supplementary Figure S8, Supplementary Table S1). This iM exhibits unusually high thermal stability at neutral pH, comparable to iM from the PDGF-A (platelet derived growth factor subunit A) promoter region with C2-C4-C5-C5-C5-C13 cytosine tracts (21). While the duplex region folds independently, the melting curves often obscure the two-stage process, leading to an apparent T_m increase in a UV melting experiments. Duplex melting is clearly seen at pH 8.0, with a $T_m = 61.2 \pm 0.1$°C (Supplementary Figure S9, Supplementary Table S1). Thus, BV42 maintains its iM-duplex structure from slightly acidic to neutral pH.

^{1}H NMR spectroscopy was used to further investigate BV42 structure (Figure 2D and E). The spectra displayed characteristic imino signals of Watson-Crick (12.0–13.5 ppm) and hemiprotonated C:C$^+$ (15.0–16.0 ppm) base pairs. The simultaneous decrease of both groups of imino signals with increasing temperatures indicates concurrent unfolding of the iM and duplex moieties at similar temperatures.

While hemiprotonated C:C$^+$ base pairs are considered as a characteristic mark of iM, they can also occur in C-rich parallel duplexes. To exclude this possibility, we explored the effect of silver cations on the BV42 structure and affinity.

Although Ag$^+$ was previously considered an alternative to protons for iM stabilization (69,70), recent studies demonstrate Ag$^+$ stabilizes parallel cytosine-tract duplexes instead of the iM structure (71,72). This alternative structure was formed in BV42 at pH 7–8 by adding AgNO$_3$. CD spectra confirmed the conformational change for both folded iM (pH 7.0 and 7.3) and unfolded iM into a cytosine parallel duplex (Supplementary Figure S10A–C), characterized by an intense negative maximum at 267 nm and a slight positive maximum at 290 nm at equimolar concentration of Ag$^+$ (72).

This alternative structure exhibited even lower affinity to hemagglutinin than the unfolded iM (Supplementary Figure S10, Supplementary Table S2). Although the association rate constant increased, the dissociation rate constant increased concurrently. These data demonstrate that cytosine tract structuring within Ag$^+$-stabilized duplexes is insufficient for hemagglutinin binding; the iM structure is essential for high-affinity interaction. This experiment confirms the iM structure necessity for aptamer biological activity. A poten-
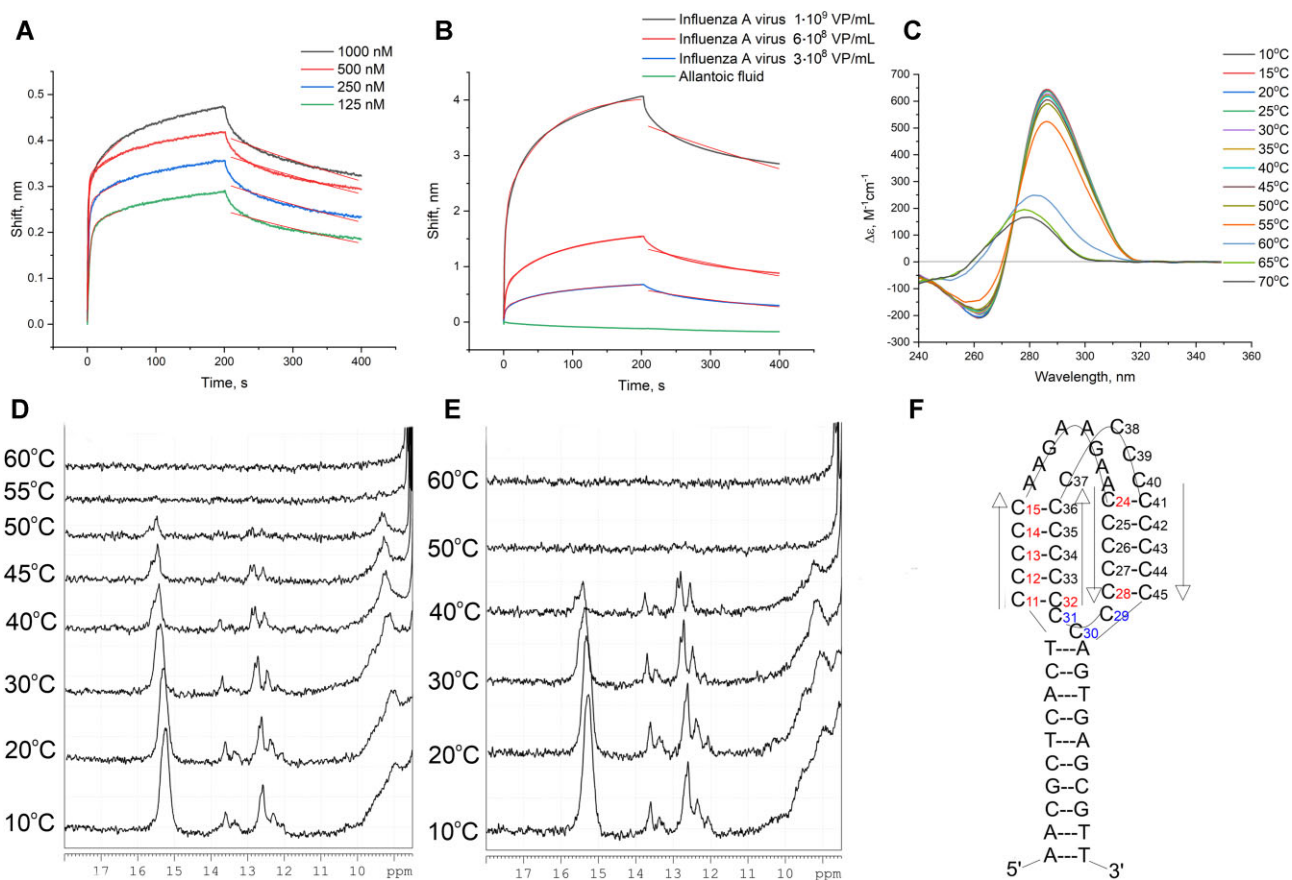
Figure 2. Affinity of aptamer BV42 to hemagglutinin, spectral data and possible aptamer structure. Aptamer binding to immobilized recombinant hemagglutinin measured by biolayer interferometry at pH 6.0 (**A**). Influenza A virus and allantoic fluid binding to immobilized BV42 aptamer at pH 6.0 measured by biolayer interferometry (**B**). The fitting of the curves is shown. CD spectra at pH 6.0 at different temperatures are shown (**C**); the melting curve and its fitting is shown in Supplementary Figure S7A. ^{1}H NMR spectra of BV42 at different temperatures in 0.15 M phosphate-citrate buffers at pH 4.7 (**D**) and 6.0 (**E**). A schematic representation of the possible conformer with sites labeled by tC° modification (**F**), indicating modified cytosines in the iM core and the loop; arrows indicate the polarity of the strand from 5′-end to 3′-end. The modified aptamers contained a single modification at one of the indicated sites.

tial conformer, depicted in Figure 2F, served as a template for introducing fluorophores for further structural elucidation

The iM moiety of BV42 is a mix of several conformers

While the preceding results supported the initial BV42 topology hypothesis, the precise structure of the aptamer remained unclear. Molecular modeling was employed to accurately represent both the i-motif core and loops. Sequence analysis suggested six conformers sharing the same duplex region but differing in their i-motif fragments (Figure 3, Supplementary Figures S11–S16). Models for each conformer were generated using the strategy proposed by Tsvetkov V.B. (57), enabling the creation of complex non-canonical DNA structures. Conformer stability was evaluated using molecular modeling parameters for iM (57). The six conformers exhibited similar free energies (Supplementary Figure S17), indicating BV42 is represented as a set of coexisting iMs sharing a common duplex module. All predicted variants contained ten C:C$^+$ pairs and a conserved AAGAAGAA loop; however, the lengths of the second and third C_n loops varied due to C_{22}-tract sliding. C:C$^+$ pair stability was assessed based on geometric parameters: inter-cytosine mass center distances and an-

gles between normals to the planes of the nucleobases. Deviations from the mean for these values are shown in Figure 3, Supplementary Figures S18–S23, Supplementary Table S3.

This analysis revealed that the conformer with a single cytosine in the second loop has destabilized C11:C30$^+$ and C28:C45$^+$ pairs near the second loop and duplex module. The C11:C30$^+$ pair was entirely disrupted, while the C28:C45$^+$ pair remained coplanar (Figure 3A and D, Supplementary Figure S11). This suggests that the conformer with a single-cytosine second loop is not stable, the iM is partially unfolded, with a free energy 13–31 kcal/mol higher than other conformers. The conformer with two cytosines in the second loop shows only a destabilized (but still coplanar) C11:C30$^+$ pair (Supplementary Figure S12). Conformers with 3–6 cytosines in the second loop maintain ten C:C$^+$ pairs within the iM core and possess similar free energies (Figure 3B, Supplementary Figures S13–S16). Increasing the second loop length beyond three nucleotides increased both the possibility of hydrogen bonding between the loop and the duplex module (Supplementary Figure S24, Supplementary Table S4), and the number of possible loop conformations. Although the six-cytosine second loop model showed the lowest free energy, other conformers differs by only 6–8 kcal/mol, an energy less than that of a single hydrogen bond.
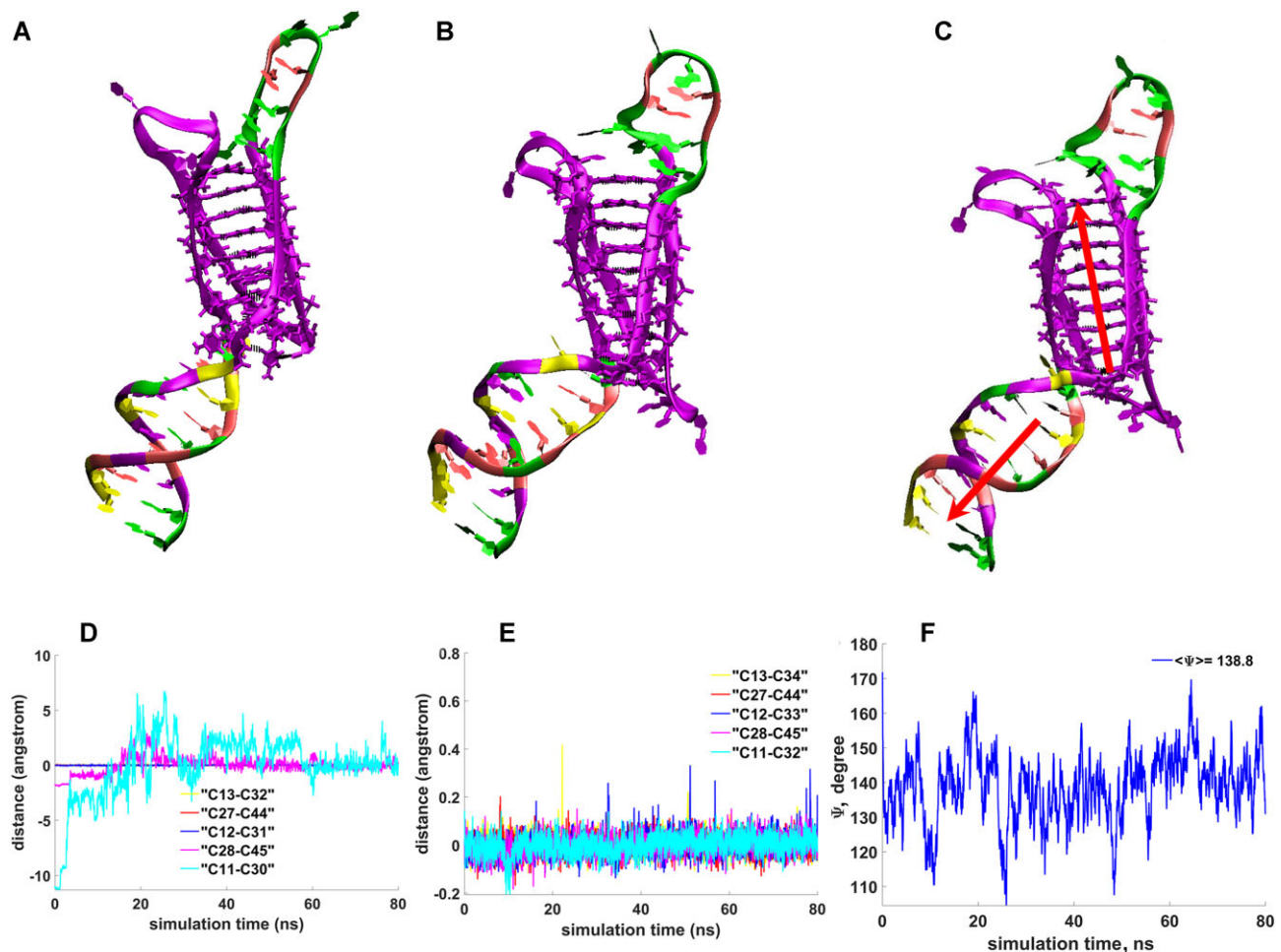
Figure 3. MD of BV42 aptamer with 1 (**A, D**) and 3 (**B, E**) cytosines in the second loop. The conformers of BV42 aptamer at the end of MD simulation are shown (**A, B**). Distances between mass centers of the cytosine bases during the molecular dynamic simulation are shown (**D, E**). Changes in coaxiality of duplex and iM modules (**C**) derived from the MD experiment with BV42 with three cytosines in the second loop (**F**). The angle between two vectors decreased down to 139º immediately and fluctuated within ± 15º.

Notably, molecular modeling revealed disruption of the final A-T base pair in the duplex, adjacent to the i-motif (see Supplementary Figure S13 and supplementary pdb file of the aptamer structure with 3-cytosine second loop). This suggests i-motif formation perturbs the neighboring duplex structure. Further analysis examined the coaxiality of iM and duplex. While initially coaxial, the angle between their respective axis rapidly decreased down to 139º and fluctuated within ± 15º during the simulation (Figure 3C and F), indicating a flexible junction between the two modules.

The conformational polymorphism was corroborated by ^{1}H NMR spectroscopy (Figure 2D and E). While the spectra exhibited characteristic peaks for the duplex module (12.3–13.6 ppm) and the iM module (15.0–15.6 ppm), the iM peaks were not resolved forming a single broad peak, indicative of a mixture of conformers. Therefore, modification of the BV42 aptamer is required to identify the hemagglutinin-binding conformer.

Phenoxazine (tCº) modification for probing the iM structure

A cytosine analogue with an extended aromatic system, 1,3-diaza-2-oxophenoxazine (tCº, Supplementary Figure S25), is

known to maintain the cytosine H-bonding pattern. When tCº is located in the middle of C-tracts, its effect on iM stability and folding is minimal. tC° can be used as a fluorescent analogue of cytosine, enabling the monitoring of the local environment within iMs (73). Fluorescence quenching is characteristic for tCº due to electron coupling interactions with surrounding protonated dC residues. Modification in the middle of a cytosine tract (tCº:C⁺ pair sandwiched between two neighboring C:C⁺ pairs) provides maximal fluorescence quenching. Modification at the end of a cytosine tract (tCº:C⁺ pair stacked only with one neighboring C:C⁺ pair) provides 1.5-times lower fluorescence quenching; whereas tCº in the loop (no tCº:C⁺ pair) leads to slight quenching of the initial fluorescence intensity, presumably due to absence of protonated cytosines and weak stacking with the neighboring C:C⁺ pair and/or loop nucleobases (73).

The modification was inserted into the BV42 aptamer to elucidate the preferred conformations of iM and local stability of C:C⁺ pairs. A set of modified oligonucleotides (Supplementary Table S5) was studied by CD, UV and fluorescent spectroscopies and biolayer interferometry. The first C_5-tract (C11-C15), terminal cytosines of the second C_5-tract (C24 and C28) and nucleotides in the region of the putative

Table 1. Melting temperatures of the tC°-modified BV42 aptamers at pH 6.0–7.3 as determined by fluorescent spectroscopy

Aptamer	pH 6.0	pH 6.5	pH 7.0	pH 7.3
BV42°C11	44.9 ± 0.5	32.0 ± 0.8	<20	<20
BV42°C12	53.0 ± 0.7	44.0 ± 0.3	40.9 ± 1.0	33 ± 4
BV42°C13	>65	47.9 ± 0.7	34.1 ± 0.6	28.8 ± 0.6
BV42°C14	>65	50.4 ± 0.3	37.0 ± 0.3	34.3 ± 0.5
BV42°C15	59.0 ± 0.2	45.0 ± 0.3	28.7 ± 1.6	<20
BV42°C24	>65	55.1 ± 0.7	35.7 ± 0.4	<20
BV42°C28	62 ± 10	56.9 ± 1.0	43.5 ± 0.6	32.7 ± 0.4
BV42°C29	>65	52.6 ± 0.5	45 ± 3	36.5 ± 0.5
BV42°C30	>65	47.5 ± 0.7	34.1 ± 0.3	<20
BV42°C31	>65	66 ± 2	52.5 ± 1.9	37.2 ± 0.3
BV42°C32	>65	57.4 ± 0.9	42.2 ± 0.3	37 ± 2

Table 2. Affinity of tC°-modified BV42 aptamers to recombinant hemagglutinin[1]

Aptamer	K_D, nM (pH 6.0)	k_{ass}, $\mu M^{-1} \cdot s^{-1}$ (pH 6.0)	k_{diss} $\cdot 10^3$, s^{-1} (pH 6.0)	K_D, nM (pH 8.0)
BV42	1.4 ± 0.4	1.3 ± 0.4	1.5 ± 0.2	176 ± 9
BV42°C11	2.2 ± 0.5	0.61 ± 0.16 ↓	1.37 ± 0.12	90 ± 10
BV42°C12	1.6 ± 0.6	1.2 ± 0.6	1.3 ± 0.2	>100
BV42°C13	1.1 ± 0.3	1.1 ± 0.4	1.13 ± 0.12 ↓	>100
BV42°C14	2.1 ± 0.5	0.9 ± 0.2	1.7 ± 0.3	>100
BV42°C15	2.8 ± 0.4 ↑	1.25 ± 0.11	3.0 ± 1.2 ↑	>100
BV42°C24	1.7 ± 0.4	0.56 ± 0.13 ↓	1.3 ± 0.3	>100
BV42°C28	4.8 ± 0.9 ↑	0.45 ± 0.10 ↓	2.2 ± 0.3 ↑	80 ± 20
BV42°C29	1.0 ± 0.4	1.9 ± 0.8	1.68 ± 0.15	80 ± 30
BV42°C31	0.7 ± 0.2 ↓	1.4 ± 0.3	0.90 ± 0.06 ↓	70 ± 30
BV42°C32	0.9 ± 0.3	1.0 ± 0.2	1.06 ± 0.12 ↓	70 ± 15

[1]The data were obtained at pH 6.0 and 8.0. K_D is the equilibrium dissociation constant; k_{ass} is the association rate constant; k_{diss} is the dissociation rate constant.

second C-loop (C29, C30, C31 and C32) were modified individually to identify the cytosines involved in iM formation.

In accordance with the previous report (73), iM stability was affected minimally by a single introduction of tC°. Melting temperatures in CD and UV melting experiments varied slightly (up to ± 10% of the value for the unmodified aptamer) confirming that tC° is an excellent non-perturbating fluorescent cytosine surrogate for probing individual residues in the BV42 aptamer (Supplementary Tables S6 and S7, Supplementary Figures S26–S31).

Based on position-dependent fluorescence quenching, the C11, C15, C24 and C32 nucleotides are involved in the formation of outer C:C+ pairs, whereas the C12, C13 and C14 nucleotides are inside the iM structure at pH 6.0 and 6.5 (Supplementary Table S8). In turn, C-sliding may be responsible for partial relocation of the C28 residue from the outer hemiprotonated pair inside the C-tract, accompanied by a displacement of C29 and C30 residues from the second loop into the iM core and a change in the loop length. Finally, the C31 nucleotide is positioned in loop regardless of C sliding. Cytosines from outer C:C+ pairs (C11, C15 and C24) quit the iM at pH 7.0 leaving C12, C13, C14, C28 and C29 in the shortened iM version. The C28 and C32 nucleotides retain their place in the iM at pH 7.0, that can be possible in the case of total remodeling of the iM during folding and indicates, in addition to conformers existing at pH 6.0, the presence of extra iM variants at pH 7.0. Thus, the fluorescence quenching experiments confirm iM folding and its structural polymorphism, as well as allow mapping some of the residues within the iM structure.

CD and UV melting experiments provide information about overall iM stability. In contrast, fluorescence melting experiments give insight into the local stability of the molecular environment involved in the interactions with fluorescent tC° residue. In this case, melting temperatures are drastically dependent on the site of modification (73,74) and may strongly deviate from those obtained by CD and UV melting (Table 1, Supplementary Figures S29–S31). Fluorescence melting showed negligible differences in the local tC° environment at high iM stability (pH 6.0). Only tC° in the BV42°C11 aptamer quickly dropped out of the stacking interactions with the neighbors. As the pH value increased (pH 6.0→7.3), the highest decrease in melting temperature was observed for the C11, C15, C24 and C30 modifications (Table 1), that may indicate their location in the outer hemiprotonated pairs in accordance with the proposed structure.

Affinity studies of tC°-modified aptamers to hemagglutinin could provide insight into the structural moieties involved in target recognition. The affinity of all tC°-modified aptamers

to the recombinant hemagglutinin was close to that of unmodified BV42 (Table 2, Supplementary Figures S32–S36). The most significant differences were observed for aptamers BV42°C28 and BV42°C31. BV42°C28 has the highest dissociation constant, 3.4 times higher than K_D of BV42. The decrease in affinity was caused by a 3-fold decrease in the association rate constant. BV42°C31 has the lowest dissociation constant with 2-fold decrease compared to BV42. The increase in affinity was achieved by stabilization of the complex due to a 1.5 times lower dissociation rate constant of BV42°C31 compared with BV42. Both modified positions are located near or within the second loop; C28 is a terminal cytosine in the iM core, whereas C31 is located in the second loop. The deviations in the affinity of C28- and C31-modified aptamers suggest the participation of these residues in the binding to hemagglutinin. To test this hypothesis, a control oligonucleotide was tested. BV42_T_loop oligonucleotide has duplex module, i-motif module (pH$_i$ = 7.1, Supplementary Figure S36), but first, second and third loops are replaced with T$_3$. This variant has a vanishingly weak binding to hemagglutinin (Supplementary Figure S36E), supporting the significance of the loops in the protein recognition.

MD confirmed the localization of tC° at the 28th and 31th positions in the iMs. tC° at 28th position forms a stable C:C+ pair with C45 (Supplementary Figure S37, Supplementary Tables S9 and S10). Overall, iM stability was diminished by 1.3%, in agreement with the slight decrease in thermal stability (Supplementary Tables S6 and S7, Supplementary Figure S38). On the contrary, tC° at the 31th position tends to be in the loop instead of a C:C+ pair, having pair occupancy below 50% (Figure 4, Supplementary Figures S39–S41, Supplementary Tables S9 and S10). Again, the overall stability of iM was diminished by 1.2%, in agreement with the slight decrease in thermal stability (Supplementary Tables S6 and S7).

To conclude this part, molecular probing has confirmed the putative iM structure and located the hemagglutinin recognition site near the second loop.

Structural studies: redesigning of iM-based aptamer

As mentioned above, ¹H NMR spectra of BV42, consistent with computational predictions, indicate the simultaneous formation of the iM and duplex moieties. DNA molecules, in which i-motif coexists with B-DNA moieties, have been
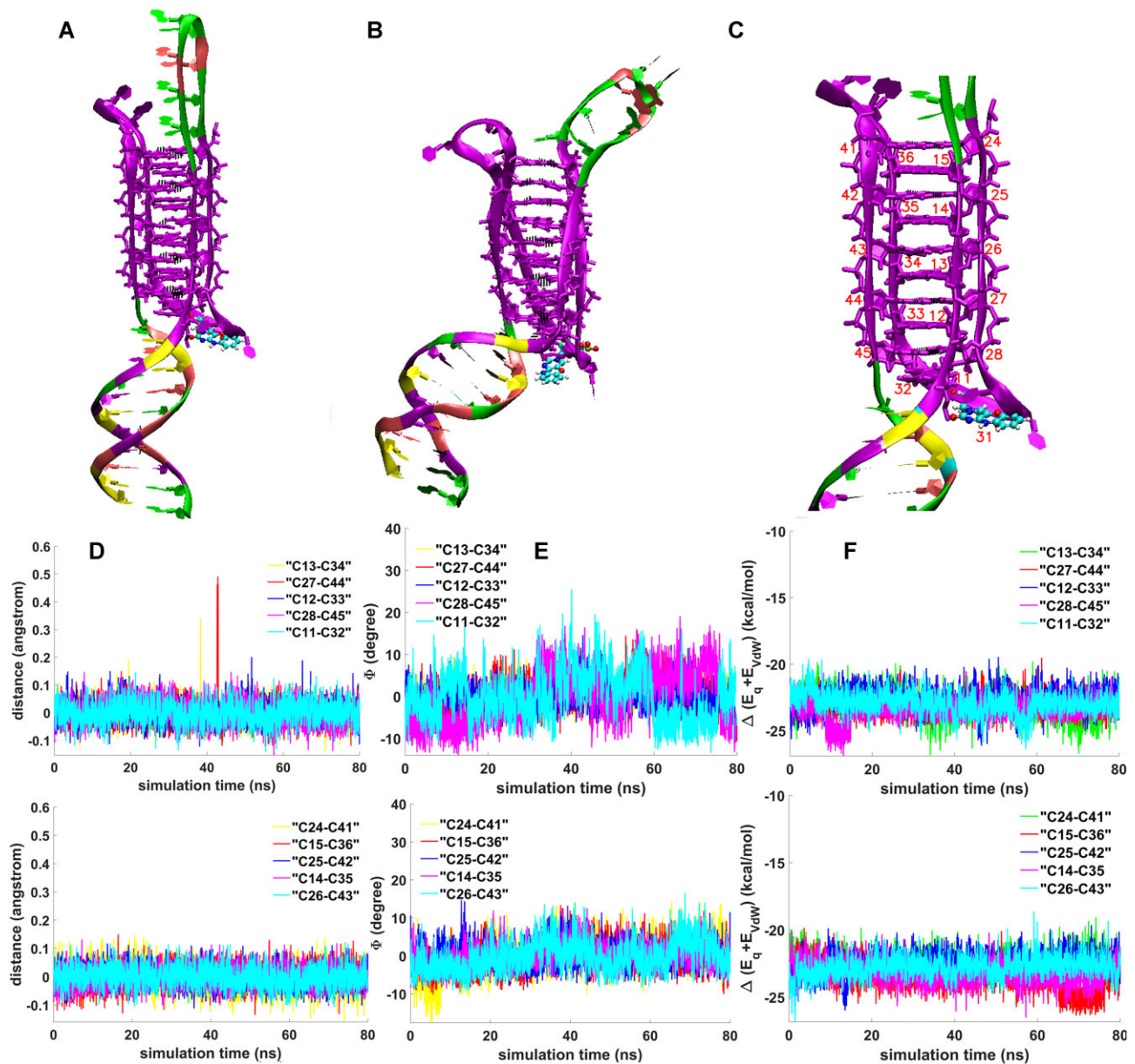
Figure 4. MD simulation results of BV42°C31 with tC° in the second loop. Conformers of BV42 at the start (**A**) an at the end (**B**) of the MD simulation at 80 ns. The enlarged iM structure is shown in subset (**C**). Distances between mass centers of the cytosine bases (**D**); angles between normals to the cytosine bases (**E**); sum of electrostatic (E_q) and Van der Waals (E_{VdW}) of contributions to interaction energy between cytosine bases (**F**).

observed recently at neutral pH (75). To date, the only high-resolution structure of a DNA molecule with an i-motif and a B-DNA part, forming a so-called iM/duplex junction, has been determined by Serrano-Chacón (75). The structure remains stable at neutral pH, in part due to the presence of a minor groove G:T:G:T tetrad capping one edge of the i-motif. Minor groove tetrads can induce significant shifts in the pK_a of the neighboring C:C$^+$ base pair, leading to strong stabilization at neutral pH. The presence of a minor groove tetrad at one side of the iM moiety not only stabilizes the structure but also eliminates potential effects of sliding of a C-tract (C-sliding) with respect to others, resulting in better quality of NMR spectra.

We used this approach to redesign the BV42 aptamer to obtain more comprehensive structural information in the absence of C-sliding. Based on the affinity data of BV42°C28

and BV42°C31, we explored various DNA constructs while preserving the second loop of the iM and replacing the first and third loops (AAGAAGAA and C19-C22) with short sequences capable of forming a G:T:G:T tetrad with two T_2 loops (Supplementary Figure S42). Also, the AT pairs were replaced with GC-pair at the duplex-iM interface in order to stabilize the duplex part in NMR experiments. We compared several constructs with a different number of cytosines in the second loop (1, 3 or 5 cytosines, respectively). The sequences studied (CapTr5C1, CapTr5C3 and CapTr5C5) are shown in Supplementary Table S5.

The NMR spectra of CapTr5C1 and CapTr5C3 exhibited narrower signals than the parent aptamer BV42, suggesting reduced conformational heterogeneity compared to BV42. The imino region spectra are consistent with the formation of C:C$^+$, G:C and A:T Watson-Crick base pairs, and G:T mis-

Table 3. Affinity of the aptamer candidates to recombinant hemagglutinin[1]

Aptamer	K_D, nM	k_{ass}, $\mu M^{-1} \cdot s^{-1}$	$k_{diss} \cdot 10^3$, s^{-1}
BV42	1.4 ± 0.4	1.3 ± 0.4	1.5 ± 0.2
CapTr5C1	1.4 ± 0.2	2.2 ± 0.8	2.9 ± 0.7
CapTr5C2	7 ± 2 ↑	0.39 ± 0.08 ↓	2.5 ± 0.4
CapTr5C3	1.5 ± 0.4	1.7 ± 0.4	2.5 ± 0.2
CapTr5C4	13 ± 5 ↑↑	0.24 ± 0.09 ↓	2.9 ± 0.2
CapTr5C5	>100 ↑↑↑	n.d.	n.d.
CapTr5T3	2.8 ± 0.7 ↑	0.7 ± 0.2	1.8 ± 0.2
CapTr4C1	8 ± 2 ↑	0.37 ± 0.09 ↓	2.8 ± 0.2
CapTr4C3	5.0 ± 1.5 ↑	0.38 ± 0.09 ↓	1.8 ± 0.2
CapTr3C1	52 ± 17 ↑↑	0.12 ± 0.04 ↓↓	5.9 ± 0.3 ↑

[1]The data were obtained at pH 6.0. K_D is the equilibrium dissociation constant; k_{ass} is the association rate constant; k_{diss} is the dissociation rate constant.

matches at neutral and acidic pH (Supplementary Figure S43). Although the size of these systems makes the spectra too complex for complete sequential assignments, the cross-peaks patterns in the NOESY experiments (Supplementary Figures S44 and S45) confirm the formation of duplex and i-motif structures, in agreement with the designed structures (Supplementary Figure S42). In both cases, exchangeable protons signals are visible at high temperature, indicating that the structures are very stable (Supplementary Figure S43). This high stability is confirmed by CD and UV melting experiments (Supplementary Table S11, Supplementary Figures S46 and S47).

Interestingly, CapTr5C1 and CapTr5C3 have the same affinity for hemagglutinin as the parent aptamer BV42 (K_D = 1.4–1.5 nM). In turn, CapTr5C5 exhibited extremely low affinity for hemagglutinin with $K_D > 100$ nM (Table 3, Supplementary Figure S48). The affinity of several other variants was also tested (Table 3, Supplementary Figure S49). CapTr5C2 and CapTr5C4, with two or four cytosines in the second loop, respectively, have 5-fold and 10-fold lower affinity than BV42. Thus, the length of the second loop is important for target recognition. Models of the aptamer with a 1-nucleotide loop are not stable, as can be clearly seen from MD calculations for the BV42 conformer and pH dependence for CapTr5C1 (Figure 3, Supplementary Figure S46). This is consistent with the NMR spectra of CapTr5C1. The low-intensity imino signals observed at low temperature at 15.0 and 15.9 ppm (Supplementary Figure S43) suggest the presence of minor i-motif conformations, most probably affecting the second loop, since the signals from the G:T:G:T tetrad and the Watson–Crick base pairs do not show any indication of multiple species. These signals are not observed in CapTr5C3, suggesting that a one-residue loop is suboptimal, leading to distortion of the structure. The higher molar CD observed for CapTr5C3 versus CapTr5C1 is also in line with this conclusion.

Our approach, consisting of replacing the loops responsible for structural heterogeneity with a G:T:G:T tetrad at the side of the i-motif opposite to the duplex region, provided shorter oligonucleotides with good affinity and stability properties. However, these DNA constructs are still too large to yield NMR spectra of sufficient quality for their complete assignment. To gain more insight into the structure of these aptamers, we explored shorter constructs with a reduced number of C:C$^+$ pairs in the i-motif module and eliminated the terminal AT base pairs in the duplex region. The sequences CapTr4C3 and CapTr2C3 were studied using NMR meth-

ods (Figure 5, Supplementary Figure S50). ^{1}H NMR spectra present similar features to the larger constructs previously studied (Figure 5, Supplementary Figure S51). In the case of CapTr2C3, signals of the C:C$^+$ pair disappear at a relatively lower temperature compared to previous constructs, indicating that the i-motif part of the structures is less stable. Therefore, we focused on the sequence CapTr4C3.

NMR spectra of CapTr4C3 are shown in Figure 5C and Supplementary Figures S51–S57. Complete assignment of the duplex region could be achieved following standard strategies. Characteristic sequential amino-methyl cross-peaks from $5'$ − TC − $3'$ steps of the duplex region served as a starting point to identify spin systems corresponding to T4 and C5. From this point onwards, sequential aromatic-aromatic and sugar-aromatic contacts, together with cross-peaks between exchangeable protons involved in Watson-Crick base pairing, allowed sequential assignment of the B-DNA moiety. Formation of the G8-C36 base pair could not be confirmed, indicating its low stability due to its proximity to the interface. Assignment of the i-motif region was more complicated due to strong overlapping of the cytosine amino signals, but it could be partially achieved. Key contacts between hemiprotonated C residues confirmed the presence of a C:C$^+$ stack capped by G residues involved in the formation of the G:T:G:T tetrad. The $5'$ − C12-G13-T14 − $3'$ and $5'$ − C27-G28-T29 − $3'$ tracks were easily followed by sequential Me/H2'/H2''-H6/H8 and by GH1-CH3$^+$ cross-peaks. G13:G28 minor groove interaction was confirmed on the basis of H1'-H1' cross-peaks. GH1-TH1 cross-peaks were detected at slightly acidic conditions (pH 6), confirming the formation of the G:T:G:T tetrad. H42/H41 protons could not be related to aromatic/sugar protons due to large overlap of the signals. However, CH3$^+$-CH3$^+$ cross-peaks led us to sequentially assign six stacked C:C$^+$ bp as C12-C27$^+$ (15.34 ppm), C17-C32$^+$ (15.69 ppm), C11-C26$^+$ (15.49 ppm), C18-C33$^+$ (15.50 ppm), C10-C25$^+$ (15.64 ppm) and C19-C34$^+$ (15.52 ppm). This stacking order was confirmed by characteristic H3$^+$-H2'/H2'' and H42-H2'/H2'' spatial cross-peaks between cytosines through the major grooves of the structure. The assignment list is shown in Supplementary Table S12. Protons of the residues in the G:T:G:T tetrad and in most of the i-motif regions could be assigned. However, the protons in the C-loop and in the two cytosines close to the duplex region (C20 and C35) were not detected. These results correlate well with the experiments on phenoxazine-modified BV42 whereas the C:C$^+$ pairs near the duplex part disappeared under the increase of pH from 6 to 7 (Supplementary Table S8). The detailed structure of the aptamer is shown in Figure 6, including G:T:G:T tetrad (Figure 6A), duplex-i-motif junction (Figure 6B) and several views of a representative structure (Figure 6C–F).

The lack of exchangeable signals for these residues indicates that they are not involved in base pairs. 234 experimental distance constraints plus additional distances derived from confirmed base pairs (Supplementary Table S13) were used to determine the structure, as described in the 'Materials and methods' section. The structure can be accessed using PDB ID 8S4N and BMRB ID 34906. The ensemble of superimposed structures is shown in Figure 7. The duplex and most of the i-motif regions are well-defined when superimposed independently (see Supplementary Table S12). However, due to the lack of distance constraints in the interface and in the C-loop regions, the relative position of the i-motif and the duplex is not well-defined.
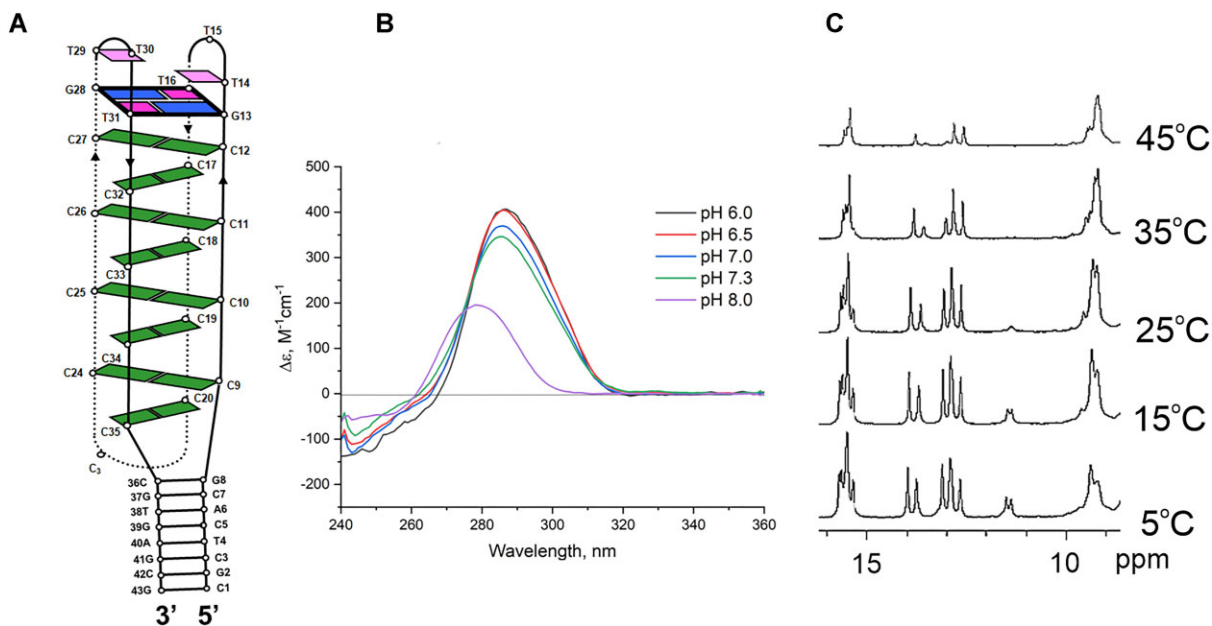
Figure 5. Redesigning of BV42 aptamer. Scheme of the putative structure of CapTr4C3 (**A**). CD spectra of the aptamer in PBS buffer at pH in the range of 6.0–8.0 confirms i-motif structure (**B**). ^{1}H NMR spectra at various temperatures recorded for an 800 μM solution of CapTr4C3 in 10 mM phosphate buffer at pH 7.0 (**C**).
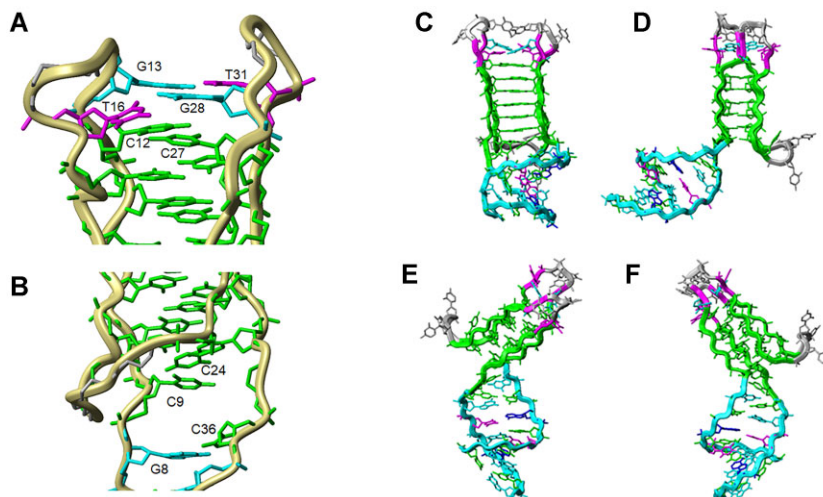

Figure 6. Structure of CapTr4C3. Enlarged junctions between G:T:G:T tetrad and i-motif (**A**) as well as i-motif and duplex junction (**B**) are shown. Views of a representative structure of CapTr4C3: view from the major groove of the i-motif moiety (**C**); view from the i-motif minor groove (**D**); view from the minor groove side of the duplex moiety (**E**); view from the duplex major groove (**F**).

Despite the lack of definition in the interface region, the solution structure of CapTr4C3 at pH 7.0 confirms many of the features suggested by the NMR spectra of previous constructs. The complementary region adopts a well-defined B-form helical structure. Six C:C$^+$ base pairs are clearly intercalated, forming a standard i-motif. Cytosines C9, C20, C24 and C35 are not base-paired, as well as the three cytosines in the second loop. The GC base pair in the duplex close to the C-rich region is probably distorted, but the presence of a guanine imino proton nearly at 13 ppm suggests that these residues are base-paired. This highly flexible interface region contrasts with the well-defined interface observed in the three-dimensional structure of 7O5E (75). It must be noted that the topology of the two interfaces is different since the complementary region in

CapTr4C3 occurs between two terminal segments, whereas in 7O5E it occurs in a loop region connecting two C-tracts.

Although the two redesigned aptamers, CapTr5C1 and CapTr5C3, can recognize hemagglutinin with high affinity (Table 3), the latter presents better thermal and pH stability than that of BV42, with a pH$_i$ of 7.3 (Supplementary Figure S46) and the highest melting temperatures among the aptamers studied (Supplementary Table S11, Supplementary Figure S47). In contrast, the loop bearing three thymine residues (CapTr5T3) had a 2-fold diminished affinity, reflecting the loss of 1–2 polar contacts in the complex. The suboptimal loops decreased the association rate constants, affecting the recognition process. Moreover, shortening the C-tract length by one or two cytosine residues (CapTr4C1
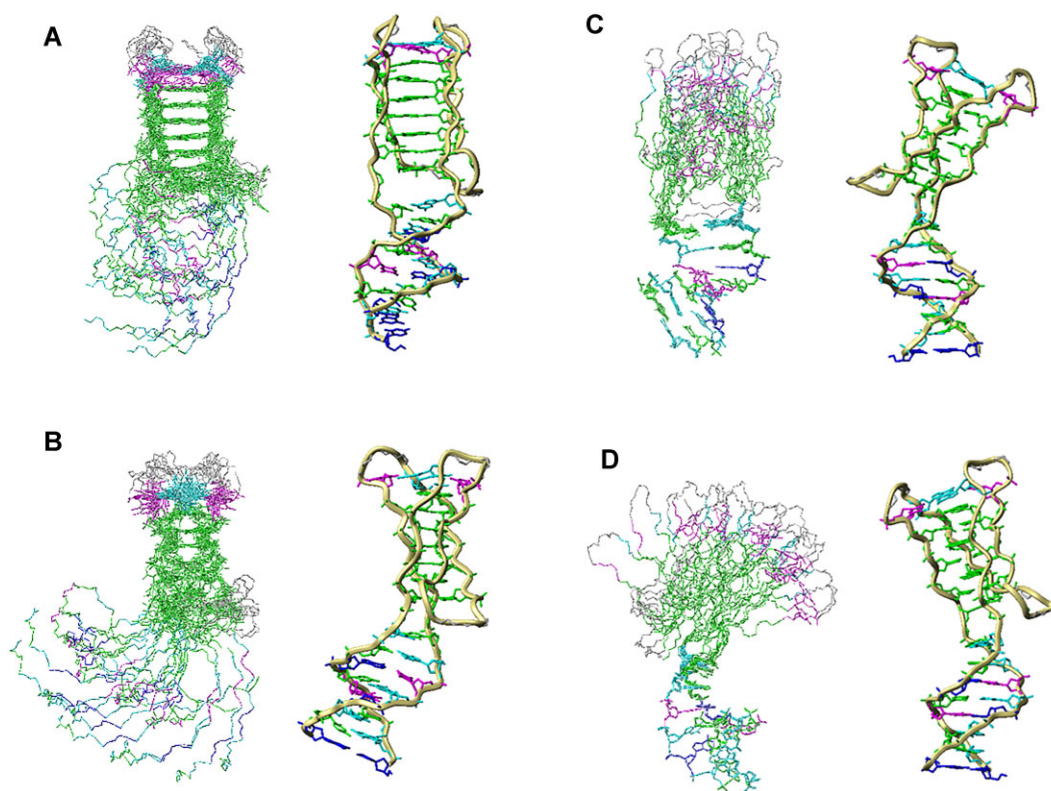
Figure 7. Two views of the ensemble of 10 resulting structures of CapTr4C3. Top: Superpositions considering only cytosines forming C:C$^+$ pairs in the i-motif moiety. Bottom: Superpositions considering only residues involved in the duplex moiety. Only the backbone is shown for residues not considered in the superposition.

and CapTr4C3 versus CapTr3C1) results in a 4–5-fold and 37-fold decrease in affinity, respectively (Table 3). However, for three cytosine tracts, iM stability was diminished significantly, complicating the interpretation of the results (Supplementary Figure S58). On contrary, CapTr4C3 has stable iM with melting temperature $57.4 \pm 0.6°$C at pH 6.0 that is even higher than T_m of BV42 under the same conditions (Supplementary Figure S59). Similarly, pH stability of the redesigned aptamer is high, namely, pH$_i$ = 7.5 that is higher than that of BV42 (pH$_i$ = 7.3, Supplementary Figure S58). The effect of C-tract shortening is obvious at neutral pH providing 3°C lower T_m at pH 7.0 (Supplementary Figure S60). Thus, the redesigned version is a good model of BV42 conformer with three cytosines in the second loop. The functional activity of CapTr4C3 is nearly the same as BV42 aptamer (Supplementary Figures S3 and S4), moreover, these two aptamers bind the same site as it is shown in the experiment with direct competition for influenza A virus (Supplementary Figure S2).

Discussion

iMs are known to be potential genomic regulatory elements influencing transcription and replication (26,76). While folded iM structures have been identified in the genome by immunoassays (77), few iM-binding proteins are currently known (24–26) and recognized structural features remain unclear. Previously, an artificial C-rich oligonucleotide was designed to bind hemagglutinin of influenza A virus and inhibit host cell entry (52). Our study shows BV42 contains an iM module that is crucial for protein recognition. Folded BV42 binds recombinant and viral hemagglutinin at pH 6.0 (Figure 2) with high affinity (K_D 1.4 nM and 10 fM, respectively), while unfolded BV42 at pH 8.0 exhibits negligible binding (Supplementary Figure S1). Conformational aptamer alteration by Ag$^+$ ions abolishes the functional activity that confirm the requirement of iM formation for high affine binding (Supplementary Table S2).

BV42 displays conformational heterogeneity due to C$_{22}$-tract sliding during folding. When experimental methods hinder precise non-canonical DNA or RNA structure topology prediction, molecular modeling using above mentioned strategy (57) can help in comprehensive structural analysis of all expected conformers. MD simulations predicted coexistence of several conformations for BV42 with similar stability, each featuring the second loop bearing 2 to 6 cytosine residues (Figure 3, Supplementary Figures S11–S16). The loop containing only one nucleotide destabilizes the neighboring C:C$^+$ pair, increasing the actual length of the loop. This conformational heterogeneity resulted in unresolvable ^{1}H NMR spectra.

To map iM cytosine inclusion, tC^o nucleotide, originally a flanks-dependent duplex stabilizer (78,79) with duplex-intercalation-dependent fluorescence change (80,81), was introduced into BV42. tC^o impact minimally on thermal stability, pK$_i$ and the folding pathway, due to maintained cytosine H-bonding and undisturbed geometry, allows its use as a fluorescent cytosine surrogate for monitoring individual iM residues (73). The single-nucleotide modification caused slight iM rearrangement, introducing modified nucleotide into the core with minimal overall stability impact regardless of tC^o position (Supplementary Tables S6 and S7). However, fluorescence melting revealed significant tC^o:C$^+$ pair stability dif-

ferences depending on the modification site and folding pH (Table 1), supporting distinct iM arrangement at pH 6.0 and 7.0. Consequently, BV42 structure remains too complicated for definitive structural assignment.

Serrano-Chacón *et al*. proposed an alternative way for iM stabilization at neutral pH using a G:T:G:T tetrad cap to restrict C-tract sliding (75). We adopted this approach for BV42 redesign, recognizing that any minor groove tetrad could serve this function. While G:T:G:T, G:C:G:T and G:C:G:C tetrads can stabilize i-motifs (82), we selected G:T:G:T to facilitate spectral assignment. G:T base pairs are readily identifiable in the NMR spectra, unlike G:C base pairs within a G:C:G:C tetrad, which can be confused with G:C base pairs in the duplex part. Furthermore, G:C:G:C tetrads exhibit pH dependency (83), a characteristic we aimed to avoid in our constructs.

Affinity experiments on tC°-modified BV42 variants revealed significant K_D changes for two aptamers modified with tC° near the second loop (C29-C31) (Table 2), suggesting its involvement in hemagglutinin recognition. Retaining this loop in redesigned sequences, with lengths varying from 1 to 5 nucleotides, and incorporating a G:T:G:T cap in place of the AAGAAGAA and C37-C40 loops, yielded stable iM structures at pH 6.0. However, only CapTr5C1 and CapTr5C3, with 1 and 3 cytosine residues in the second loop, respectively, maintained suitable affinity. These oligonucleotides form remarkably stable structures at neutral pH. MD suggests, the CapTr5C1 loop instability causes the loss of an outer C:C⁺ pair, restoring of 3-nucleotide loop. A 3-cytosine loop may be crucial for the recognition site configuration, as replacing these cytosine residues with thymines, while preserving the 3-nucleotide length, reduced affinity two-fold. Shortening the iM core decreased affinity 5-fold and 37-fold for C_4 and C_3 iM tracts, respectively (Table 3). Therefore, high hemagglutinin affinity depends on both an extended, stable iM core and a 3-cytosine loop.

NMR and MD demonstrate coexisting duplex and iM modules in both parental BV42 and its redesigned truncated version, surprisingly, unstacked and joint by a flexible junction. Given the putative hemagglutinin recognition site location within the second loop near the junction, flexibility may facilitate induced-fit aptamer conformation during binding.

In conclusion, we present the first demonstration of i-motif structures playing a crucial role in protein/aptamer recognition. Specifically, the iM-duplex structure serves as the hemagglutinin binding interface, featuring a prominent short cytosine loop positioned. While MD identified multiple conformers with comparable free energies, hindering complete structural analysis, subsequent redesign produced a configuration with both conformational homogeneity and high hemagglutinin affinity. NMR elucidated the unique i-motif/duplex architecture with a flexible junction. These findings represent a significant advance in rational aptamer design, highlighting the functional potential of DNA i-motifs for specific protein recognition under physiological conditions.

Data availability

The NMR structure of the aptamer is available in RCSB Protein Data Bank at https://www.rcsb.org/structure/8s4n, and can be accessed with PDB ID 8S4N. Also, the NMR structure of the aptamer is available in Biological Magnetic Res-

onance Data Bank at https://bmrb.io/data_library/summary/?bmrbId=34906, and can be accessed with BMRB ID 34906. Other data underlying this article are available in the article and in its online supplementary material.

Supplementary data

Supplementary Data are available at NAR Online.

Conflict of interest statement

None declared.

References

1. Mayer,G. (2009) The chemical biology of aptamers. *Angew. Chem. Int. Ed. Engl.*, **48**, 2672–2689.
2. Chen,J., Zhou,J., Peng,Y., Xie,Y. and Xiao,Y. (2022) Aptamers: a prospective tool for infectious diseases diagnosis. *J. Clin. Lab. Anal.*, **36**, e24725.
3. Zhuo,Z., Yu,Y., Wang,M., Li,J., Zhang,Z., Liu,J., Wu,X., Lu,A., Zhang,G. and Zhang,B. (2017) Recent advances in SELEX technology and aptamer applications in biomedicine. *Int. J. Mol. Sci.*, **18**, 2142.
4. Xu,Y., Jiang,X., Zhou,Y., Ma,M., Wang,M. and Ying,B. (2021) Systematic evolution of ligands by exponential enrichment technologies and aptamer-based applications: recent progress and challenges in precision medicine of infectious diseases. *Front. Bioeng. Biotechnol.*, **9**, 704077.
5. Chandola,C., Kalme,S., Casteleijn,M.G., Urtti,A. and Neerathilingam,M. (2016) Application of aptamers in diagnostics, drug-delivery and imaging. *J. Biosci.*, **41**, 535–561.
6. Zhou,W., Huang,P.J.J., Ding,J. and Liu,J. (2014) Aptamer-based biosensors for biomedical diagnostics. *Analyst*, **139**, 2627–2640.
7. Antipova,O.M., Zavyalova,E.G., Golovin,A.V., Pavlova,G.V., Kopylov,A.M. and Reshetnikov,R.V. (2018) Advances in the application of modified nucleotides in SELEX technology. *Biochemistry (Moscow)*, **83**, 1161–1172.
8. Duo,J., Chiriac,C., Huang,R.Y.-C., Mehl,J., Chen,G., Tymiak,A., Sabbatini,P., Pillutla,R. and Zhang,Y. (2018) Slow off-rate modified aptamer (SOMAmer) as a novel reagent in immunoassay

development for accurate soluble glypican-3 quantification in clinical samples. *Anal. Chem.*, **90**, 5162–5170.

9. Abeydeera,N.D., Egli,M., Cox,N., Mercier,K., Conde,J.N., Pallan,P.S., Mizurini,D.M., Sierant,M., Hibti,F.E., Hassell,T., *et al.* (2016) Evoking picomolar binding in RNA by a single phosphorodithioate linkage. *Nucleic Acids Res.*, **44**, 8052–8064.

10. Davies,D.R., Gelinas,A.D., Zhang,C., Rohloff,J.C., Carter,J.D., O'Connell,D., Waugh,S.M., Wolk,S.K., Mayfield,W.S., Burgin,A.B., *et al.* (2012) Unique motifs and hydrophobic interactions shape the binding of modified DNA ligands to protein targets. *Proc. Natl. Acad. Sci. USA*, **109**, 19971–19976.

11. Yatime,L., Maasch,C., Hoehlig,K., Klussmann,S., Andersen,G.R. and Vater,A. (2015) Structural basis for the targeting of complement anaphylatoxin C5a using a mixed L-RNA/L-DNA aptamer. *Nat. Commun.*, **6**, 6481.

12. Zavyalova,E.G., Legatova,V.A., Alieva,R.S., Zalevsky,A.O., Tashlitsky,V.N., Arutyunyan,A.M. and Kopylov,A.M. (2019) Putative mechanisms underlying high inhibitory activities of bimodular DNA aptamers to thrombin. *Biomolecules*, **9**, 41.

13. Zavyalova,E., Tagiltsev,G., Reshetnikov,R., Arutyunyan,A. and Kopylov,A. (2016) Cation coordination alters the conformation of a thrombin-binding G-quadruplex DNA aptamer that affects inhibition of thrombin. *Nucleic Acid Ther.*, **26**, 299–308.

14. DeRosa,M.C., Lin,A., Mallikaratchy,P., McConnell,E.M., McKeague,M., Patel,R. and Shigdar,S. (2023) In vitro selection of aptamers and their applications. *Nat. Rev. Methods Primers.*, **3**, 54.

15. Kato,K., Ikeda,H., Miyakawa,S., Futakawa,S., Nonaka,Y., Fujiwara,M., Okudaira,S., Kano,K., Aoki,J., Morita,J., *et al.* (2016) Structural basis for specific inhibition of autotaxin by a DNA aptamer. *Nat. Struct. Mol. Biol.*, **23**, 395–401.

16. Jarvis,T.C., Davies,D.R., Hisaminato,A., Resnicow,D.I., Gupta,S., Waugh,S.M., Nagabukuro,A., Wadatsu,T., Hishigaki,H., Gawande,B., *et al.* (2015) Non-helical DNA triplex forms a unique aptamer scaffold for high affinity recognition of nerve growth factor. *Structure*, **23**, 1293–1304.

17. Russo Krauss,I., Spiridonova,V., Pica,A., Napolitano,V. and Sica,F. (2016) Different duplex/quadruplex junctions determine the properties of anti-thrombin aptamers with mixed folding. *Nucleic Acids Res.*, **44**, 983–991.

18. Abou Assi,H., Garavís,M., González,C. and Damha,M.J. (2018) i-motif DNA: structural features and significance to cell biology. *Nucleic Acids Res.*, **46**, 8038–8056.

19. Day,H.A., Pavlou,P. and Waller,Z.A. (2014) i-motif DNA: structure, stability and targeting with ligands. *Bioorg. Med. Chem.*, **22**, 4407–4418.

20. Abdelhamid,M.A.S. and Waller,Z.A.E. (2020) Tricky topology: persistence of folded human telomeric i-motif DNA at ambient temperature and neutral pH. *Front. Chem.*, **8**, 40.

21. Brazier,J.A., Shah,A. and Brown,G.D. (2012) i-motif formation in gene promoters: unusually stable formation in sequences complementary to known G-quadruplexes. *Chem. Commun.*, **48**, 10739–10741.

22. 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.

23. 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.

24. Niu,K., Zhang,X., Deng,H., Wu,F., Ren,Y., Xiang,H., Zheng,S., Liu,L., Huang,L., Zeng,B., *et al.* (2018) BmILF and i-motif structure are involved in transcriptional regulation of BmPOUM2 in Bombyx mori. *Nucleic Acids Res.*, **46**, 1710–1723.

25. Bai,X., Talukder,P., Daskalova,S.M., Roy,B., Chen,S., Li,Z., Dedkova,L.M. and Hecht,S.M. (2017) Enhanced binding affinity for an i-motif DNA substrate exhibited by a protein containing nucleobase amino acids. *J. Am. Chem. Soc.*, **139**, 4611–4614.

26. 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.

27. Xin,B.-G., Huang,L.-Y., Yuan,L.-G., Liu,N.-N., Li,H.-H., Ai,X., Lei,D.-S., Hou,X.-M., Rety,S. and Xi,X.-G. (2024) Structural insights into the N-terminal APHB domain of HrpA: mediating canonical and i-motif recognition. *Nucleic Acids Res.*, **52**, 3406–3418.

28. Zeraati,M., Langley,D.B., Schofield,P., Moye,A.L., Rouet,R., Hughes,W.E., Bryan,T.M., Dinger,M.E. and Christ,D. (2018) I-motif DNA structures are formed in the nuclei of human cells. *Nat. Chem.*, **10**, 631–637.

29. Zanin,I., Ruggiero,E., Nicoletto,G., Lago,S., Maurizio,I., Gallina,I. and Richter,S.N. (2023) Genome-wide mapping of i-motifs reveals their association with transcription regulation in live human cells. *Nucleic Acids Res.*, **51**, 8309–8321.

30. Víšková,P., Ištvánková,E., Ryneš,J., Džatko,Š., Loja,T., Živković,M.L., Rigo,R., El-Khoury,R., Serrano-Chacón,I., Damha,M.J., *et al.* (2024) In-cell NMR suggests that DNA i-motif levels are strongly depleted in living human cells. *Nat. Commun.*, **15**, 1992.

31. 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 inside living cells. *Nat. Nanotechnol.*, **4**, 325–330.

32. Turaev,A.V., Isaakova,E.A., Severov,V.V., Bogomazova,A.N., Zatsepin,T.S., Sardushkin,M.V., Aralov,A.V., Lagarkova,M.A., Pozmogova,G.E. and Varizhuk,A.M. (2021) Genomic DNA i-motifs as fast sensors responsive to near-physiological pH microchanges. *Biosens. Bioelectron.*, **175**, 112864.

33. Ma,W., Yan,L., He,X., Qing,T., Lei,Y., Qiao,Z., He,D., Huang,K. and Wang,K. (2018) Hairpin-contained i-motif based fluorescent ratiometric probe for high-resolution and sensitive response of small pH variations. *Anal. Chem.*, **90**, 1889–1896.

34. Shtork,A., Tsvetkov,V., Slushko,G., Lushpa,V., Severov,V., Kamzeeva,P., Varizhuk,A. and Aralov,A. (2024) Diversifying i-motif-based pH sensors: labeling patterns tune the intracellular localization. *Sens. Actuators B Chem*, **411**, 135747.

35. Li,L., Jiang,Y., Cui,C., Yang,Y., Zhang,P., Stewart,K., Pan,X., Li,X., Yang,L., Qiu,L., *et al.* (2018) Modulating aptamer specificity with pH-responsive DNA bonds. *J. Am. Chem. Soc.*, **140**, 13335–13339.

36. Lei,Y., He,X., Tang,J., Shi,H., He,D., Yan,L., Liu,J., Zenga,Y. and Wang,K. (2018) Ultra-pH-responsive split i-motif based aptamer anchoring strategy for specific activatable imaging of acidic tumor microenvironment. *Chem. Commun.*, **54**, 10288–10291.

37. Tian,Z., Peng,P., Wang,H., Zheng,J., Shi,L. and Li,T. (2020) Aptamer-braked multi-hairpin cascade circuits for logic-controlled label-free in situ bioimaging. *Anal. Chem.*, **92**, 10357–10364.

38. Ma,W., Sun,H., Chen,B., Jia,R., Huang,J., Cheng,H., He,X., Huang,M. and Wang,K. (2021) Engineering a facile aptamer "molecule-doctor" with hairpin-contained i-motif enables accurate imaging and killing of cancer cells. *Anal. Chem.*, **93**, 14552–14559.

39. 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.

40. Son,S., Nam,J., Kim,J., Kim,S. and Kim,W.J. (2014) i-motif-driven Au nanomachines in programmed siRNA delivery for gene-silencing and photothermal ablation. *ACS Nano*, **8**, 5574–5584.

41. Han,X., Xu,X., Wu,Z., Wu,Z. and Qi,X. (2021) Synchronous conjugation of i-motif DNA and therapeutic siRNA on the vertexes of tetrahedral DNA nanocages for efficient gene silence. *Acta Pharm. Sin. B.*, **11**, 3286–3296.

42. Narum,S., Deal,B., Ogasawara,H., Mancuso,J.N., Zhang,J. and Salaita,K. (2024) An endosomal escape trojan horse platform to improve cytosolic delivery of nucleic acids. *ACS Nano*, **18**, 6186–6201.

43. Lee,Y., Nam,K., Kim,Y.M., Yang,K., Kim,Y., Oh,J.W. and Roh,Y.H. (2024) Functional polymeric DNA nanostructure-decorated

cellulose nanocrystals for targeted and stimuli-responsive drug delivery. *Carbohydr. Polym.*, **340**, 122270.

44. Miao,D., Yu,Y., Chen,Y., Liu,Y. and Su,G. (2020) Facile construction of i-motif DNA-conjugated gold nanostars as near-infrared and pH dual-responsive targeted drug delivery systems for combined cancer therapy. *Mol. Pharm.*, **17**, 1127–1138.

45. Kim,J.Y., Song,J., Jung,H. and Mok,H. (2018) i-motif-coated exosomes as a pH-sensitive carrier for anticancer drugs. *Appl. Biol. Chem.*, **61**, 599–606.

46. Yang,T., Peng,S., Zeng,R., Xu,Q., Zheng,X., Wang,D., Zhou,X. and Shao,Y. (2022) Visible light-driven i-motif-based DNAzymes. *Spectrochim. Acta A Mol. Biomol. Spectrosc.*, **270**, 120845.

47. Guo,Y., Tong,Z., Huang,Y., Tang,J., Xue,X., Yang,D. and Yao,C. (2024) Dynamic assembly of DNA nanostructures in cancer cells enables the coupling of autophagy activating and real-time tracking. *Nano Lett.*, **24**, 3532–3540.

48. Gao,Z.F., Chen,D.M., Lei,J.L., Luo,H.Q. and Li,N.B. (2015) A regenerated electrochemical biosensor for label-free detection of glucose and urea based on conformational switch of i-motif oligonucleotide probe. *Anal. Chim. Acta*, **897**, 10–16.

49. Hu,Y., Gao,S., Lu,H. and Ying,J.Y. (2022) Acid-resistant and physiological pH-responsive DNA hydrogel composed of A-motif and i-motif toward oral insulin delivery. *J. Am. Chem. Soc.*, **144**, 5461–5470.

50. Hu,Y., Ke,Y. and Willner,I. (2023) A pH-cascaded DNA hydrogel mediated by reconfigurable A-motif duplex, i-Motif quadruplex, and T·A-T triplex structures. *Adv. Funct. Mat.*, **33**, 2304966.

51. Hu,Y., Gao,S., Lu,H., Tan,S., Chen,F., Ke,Y. and Ying,J.Y. (2023) A self-immolative DNA nanogel vaccine toward cancer immunotherapy. *Nano Lett.*, **23**, 9778–9787.

52. Musafia,B., Oren-Banaroya,R. and Noiman,S. (2014) Designing anti-influenza aptamers: novel quantitative structure activity relationship approach gives insights into aptamer – virus interaction. *PLoS One*, **9**, e97696.

53. Zavyalova,E. and Kopylov,A. (2016) G-quadruplexes and i-motifs as scaffolds for molecular engineering of DNA aptamers. In: *G-quadruplex structures, formation and roles in biology*. Nova SciencePublishers, Hauppauge, NY, USA, pp. 53–80.

54. Gardarsson,H., Kale,A.S. and Sigurdsson,S.T. (2011) Structure–function relationships of phenoxazine nucleosides for identification of mismatches in duplex DNA by fluorescence spectroscopy. *ChemBioChem*, **12**, 567–575.

55. Killian,M.L. (2014) Hemagglutination assay for influenza virus, in animal influenza virus. In: Spackman,E. (ed.) *Methods in Molecular Biology (Methods and Protocols)*. Humana Press, New York, NY, Vol. **1161**, pp. 3–9.

56. Masson,L., Mazza,A. and De Crescenzo,G. (2000) Determination of affinity and kinetic rate constants using surface plasmon resonance. *Methods Mol. Biol.*, **145**, 189–201.

57. Tsvetkov,V.B. (2022) Modeling of possible quadruplexes and i-motifs formed during DNA contacts: strategy, classification, most probable shapes, origami based on quadruplexes. bioRxiv doi: https://doi.org/10.1101/2022.10.10.511558, 10 October 2022, preprint: not peer reviewed.

58. Zhao,Y. and Truhlar,D.G. (2008) The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Account.*, **120**, 215–241.

59. Frisch,M.J., Trucks,G.W., Schlegel,H.B., Scuseria,E., Robb,M.A., Cheeseman,J.R., Scalmani,G., Barone,V., Petersson,G.A., Nakatsuji,H., *et al.* (2016) Gaussian 16, Revision C.01. https://gaussian.com/relnotes/.

60. Case,D.A., Belfon,K., Ben-Shalom,I.Y., Brozell,S.R., Cerutti,D.S., Cheatham,T.E. III, Cruzeiro,V.W.D., Darden,T.A., Duke,R.E., Giambasu,G., *et al.* (2020) *AMBER 2020*. University of California, San Francisco.

61. Izadi,S. and Onufriev,A.V. (2016) Accuracy limit of rigid 3-point water models. *J. Chem. Phys.*, **145**, 074501.

62. Zgarbová,M., Luque,F.J., Šponer,J., Cheatham,T.E., 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.

63. Zgarbová,M., Sponer,J., Otyepka,M., Cheatham,T.E. 3rd, 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.

64. Onufriev,A., Case,D.A. and Bashford,D. (2002) Effective Born radii in the generalized Born approximation: the importance of being perfect. *J. Comput. Chem.*, **23**, 1297–1304.

65. Bizyaeva,A.A., Bunin,D.A., Moiseenko,V.L., Gambaryan,A.S., Balk,S., Tashlitsky,V.N., Arutyunyan,A.M., Kopylov,A.M. and Zavyalova,E.G. (2021) The functional role of loops and flanking sequences of G-quadruplex aptamer to the hemagglutinin of influenza A virus. *Int. J. Mol. Sci.*, **22**, 2409.

66. Einav,T., Gentles,L.E. and Bloom,J.D. (2020) SnapShot: influenza by the numbers. *Cell*, **182**, 532.

67. Ustinov,N.B., Zavyalova,E.G., Smirnova,I.G. and Kopylov,A.M. (2017) The power and limitations of influenza virus hemagglutinin assays. *Biochemistry (Mosc.)*, **82**, 1234–1248.

68. Kypr,J., Kejnovská,I., Renciuk,D. and Vorlíčková,M. (2009) Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res.*, **37**, 1713–1725.

69. Day,H.A., Huguin,C. and Waller,Z.A.E. (2013) Silver cations fold i-motif at neutral pH. *Chem. Commun.*, **49**, 7696.

70. Shi,Y., Sun,H., Xiang,J., Yu,L., Yang,Q., Li,Q., Guan,A. and Tang,Y. (2015) i-motif-modulated fluorescence detection of silver(I) with an ultrahigh specificity. *Anal. Chim. Acta*, **857**, 79–84.

71. Huard,D.J.E., Demissie,A., Kim,D., Lewis,D., Dickson,R.M., Petty,J.T. and Lieberman,R.L. (2019) Atomic structure of a fluorescent Ag8 cluster templated by a multistranded DNA scaffold. *J. Am. Chem. Soc.*, **141**, 11465–11470.

72. Kohl,F.R., Zhang,Y., Charnay,A.P., Martínez-Fernández,L. and Kohler,B. (2020) Ultrafast excited state dynamics of silver ion-mediated cytosine–cytosine base pairs in metallo-DNA. *J. Chem. Phys.*, **153**, 105104.

73. 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.

74. 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.

75. Serrano-Chacón,I., Mir,B., Escaja,N. and González,C. (2021) Structure of i-motif/duplex junctions at neutral pH. *J. Am. Chem. Soc.*, **143**, 12919–12923.

76. Kendrick,S., Muranyi,A., Gokhale,V., Hurley,L.H. and Rimsza,L.M. (2017) Simultaneous drug targeting of the promoter MYC G-quadruplex and BCL2 i-motif in diffuse large B-cell lymphoma delays tumor growth. *J. Med. Chem.*, **60**, 6587–6597.

77. Brown,S.L. and Kendrick,S. (2021) The i-motif as a molecular target: more than a complementary DNA secondary structure. *Pharmaceuticals (Basel)*, **14**, 96.

78. Lin,K.-Y., Jones,R.J. and Matteucci,M. (1995) Tricyclic 2'-deoxycytidine analogs: syntheses and incorporation into oligodeoxynucleotides which have enhanced binding to complementary RNA. *J. Am. Chem. Soc.*, **117**, 3873–3874.

79. 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.

80. Sandin,P., Börjesson,K., Li,H., Mårtensson,J., Brown,T., Wilhelmsson,L.M. and Albinsson,B. (2008) Characterization and

use of an unprecedentedly bright and structurally non-perturbing fluorescent DNA base analogue. *Nucleic Acids Res.*, **36**, 157–167.

81. Schönrath,I., Tsvetkov,V.B., Zatsepin,T.S., Aralov,A.V. and Müller,J. (2019) Silver(I)-mediated base pairing in parallel-stranded DNA involving the luminescent cytosine analog 1,3-diaza-2-oxophenoxazine. *J. Biol. Inorg. Chem.*, **24**, 693–702.

82. Mir,B., Serrano,I., Buitrago,D., Orozco,M., Escaja,N. and González,C. (2017) Prevalent sequences in the human genome can form mini i-motifs structures at physiological pH. *J. Am. Chem. Soc.*, **139**, 13985–13988.

83. Serrano-Chacón,I., Mir,B., Cupellini,L., Colizzi,F., Orozco,M., Escaja,N. and González,C. (2023) pH-dependent capping interactions induce large-scale structural transitions in i-motifs. *J. Am. Chem. Soc.*, **145**, 3696−3705.